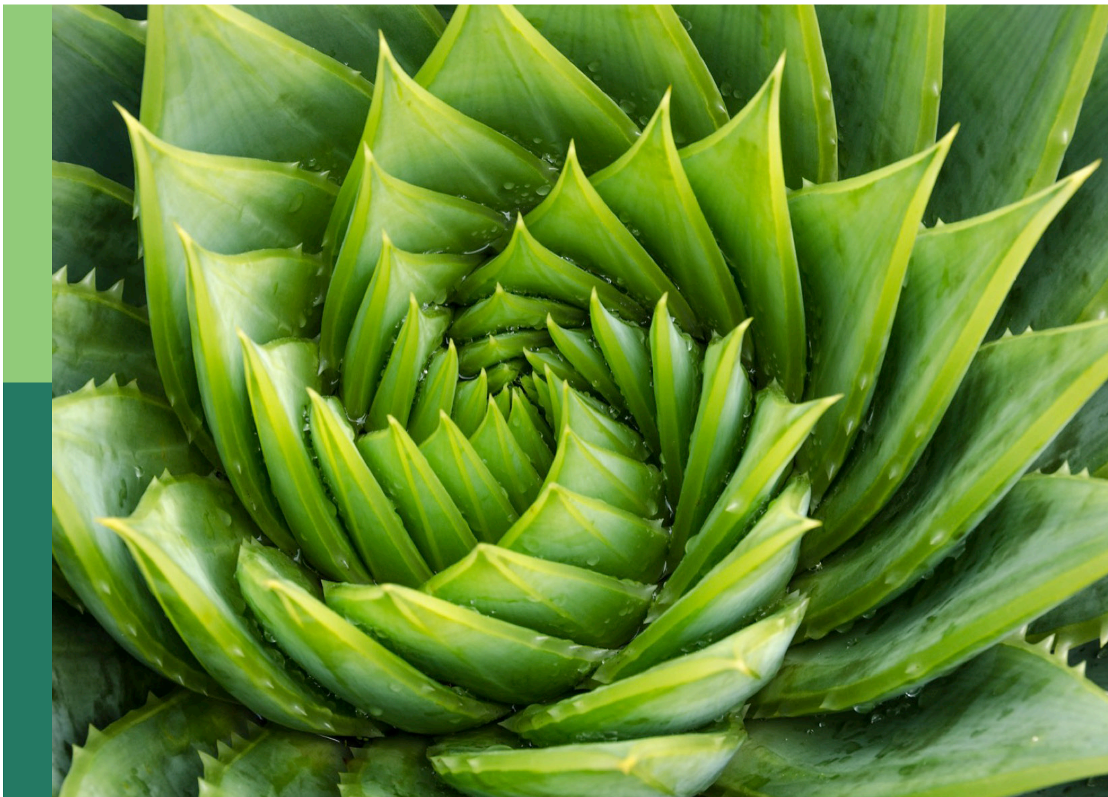
Research Topics

Plant natural products: biosynthesis, regulation, and function

Edited by
Xin Fang, Yongjun Wei, Zongxia Yu and
Boyang Ji

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-6717-3
DOI 10.3389/978-2-8325-6717-3

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Plant natural products: biosynthesis, regulation, and function

Topic editors
Xin Fang — Kunming Institute of Botany, Chinese Academy of Sciences (CAS), China
Yongjun Wei — Zhengzhou University, China
Zongxia Yu — Lushan Botanical Garden (CAS), China
Boyang Ji — BioInnovation Institute (BII), Denmark

Citation
Fang, X., Wei, Y., Yu, Z., Ji, B., eds. (2025). *Plant natural products: biosynthesis, regulation, and function*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-6717-3

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 31 July 2025
DOI 10.3389/fpls.2025.1661170

OPEN ACCESS

EDITED AND REVIEWED BY
Laigeng Li,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Xin Fang:
xinfang@mail.kib.ac.cn

RECEIVED 07 July 2025
ACCEPTED 20 July 2025
PUBLISHED 31 July 2025

CITATION
Yu Z, Wei Y, Ji B and Fang X (2025)
Editorial: Plant natural products:
biosynthesis, regulation, and function.
Front. Plant Sci. 16:1661170.
doi: 10.3389/fpls.2025.1661170

Editorial: Plant natural products: biosynthesis, regulation, and function

Zongxia Yu[1], Yongjun Wei[2], Boyang Ji[3] and Xin Fang[4]*

[1]Jiangxi Key Laboratory for Sustainable Utilization of Chinese Materia Medica Resources, Lushan Botanical Garden, Jiangxi Province and Chinese Academy of Sciences, Jiujiang, China, [2]School of Pharmaceutical Sciences, State Key Laboratory of Antiviral Drugs, Pingyuan Laboratory, Zhengzhou University, Zhengzhou, China, [3]BioInnovation Institute, Copenhagen, Denmark, [4]State Key Laboratory of Phytochemistry and Natural Medicines in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

KEYWORDS

plant specialized metabolites, biosynthesis, metabolomics, medicinal plants, phytochemistry

Editorial on the Research Topic
Plant natural products: biosynthesis, regulation, and function

Plant natural products, also referred to as plant specialized metabolites, are small molecules synthesized by plants. Throughout history, human have harnessed these compounds for diverse applications, including development of medicines, food supplements, and dyes. In their host organisms, these molecules perform diverse functions, such as mediating pollinators and mycorrhizal fungi interactions, defensing against biotic stresses like herbivores and pathogens, and protecting against abiotic stresses like UV-B radiation, frost, and drought (Dixon and Paiva, 1995).

Traditional phytochemical approaches, which typically involve isolating individual metabolites followed by targeted bioassays, have long provided foundational insights into the functions of these compounds. However, given the immense chemical diversity of plant natural products, estimated to exceed one million distinct structures (Afendi et al., 2012), these conventional methods are ill-suited for exploring the full metabolic landscape. The emergence of metabolomics has transformed this paradigm by enabling comprehensive, high-throughput analysis of hundreds to thousands of metabolites simultaneously, thereby offering unprecedented capacity to decipher metabolic functions. Moreover, these compounds are often minute quantities in planta, elucidating the biosynthetic pathways and regulatory networks that govern specialized metabolites is critical. Such understanding facilitates the development of synthetic biology strategies for heterologous production of these compounds in engineered organisms.

This Research Topic, entitled *“Plant Natural Products: Biosynthesis, Regulation, and Function,”* features 14 original research articles, one review, and one opinion piece from authors worldwide. Collectively, these contributions highlight cutting-edge discoveries across all aspects of plant specialized metabolism, from molecular mechanisms to ecological significance.

Metabolomics has emerged as a powerful tool for deciphering the dynamic chemical diversity of plants, offering insights into species-specific metabolites, developmental stage variations, and environmental responses. This approach is crucial due to the high specificity and contextual variability of plant natural products, necessitating comprehensive analyses across diverse species, ontogenetic stages, and environmental conditions. Pannequin et al. performed a comparative metabolome study of a large collection of bryophyte plants including 60 species, 15 orders, and 41 families, which provides a comprehensive overview of bryophyte chemodiversity and in-depth chemical feature of certain species. Liu et al. combined widely targeted metabolomics with biological activities assay, and sensory flavor analyses, revealing that the buds of *Lonicera japonica* Thunb. var. *chinensis* had great exploitation potentials in pharmaceuticals, beverages, and nutraceuticals. Yang et al. identified 154 common differential metabolites across diverse ripening stages of *Baccaurea ramiflora* Lour. fruit based on non-targeted metabolomics analysis, with L-sorbose and 5-hydroxyindole-3-acetic acid as taste biomarkers. Environmental stimuli frequently trigger plant natural products biosynthesis. Wang et al. investigated the effects of geography, soil and climatic factors on the two main secondary metabolites contents in the roots of *Rubia cordifolia* L. Their findings revealed that annual precipitation negatively correlated with the contents of purpurin and mollugin.

Notably, metabolomics combined with other omics approaches can provide an unbiased view of biological processes, thus facilitating the elucidation of underlying molecular mechanism. By integrated analyses of transcriptomic and metabolomic data from plant samples across various developmental stages and distinct locations, Chen et al. and Tian et al. uncovered the seasonal and spatial dynamics, as well as molecular regulation of flavonoid biosynthesis in *Cyclocarya paliurus* and *Epimedium sagittatum*, respectively. In another study, Yan et al., employed proteomic and metabolome analyses to reveal the reconfiguration of energy metabolism and terpenoid biosynthesis in cigar tobacco under low-light conditions, identifying 254 significantly differentially expressed metabolites and 780 significantly differentially expressed proteins.

Despite the rapid development of metabolomic technology, traditional phytochemical method remains playing a critical role in elucidating the functions of novel natural products. Liu et al. isolated six eudesmane-type sesquiterpenoids from *Laggera pterodonta*, and these compounds exhibited varying degrees of inhibitory effects against six plant pathogenic fungi, including *Phytophthora nicotianae*, *Fusarium oxysporum*, *Alternaria alternata*, *Gloeosporium frucrigenum* Berk, *Colletotrichum fructicola*, and *Botrytis cinerea*. Ruan et al. characterized nine new ent-atisane-type diterpenoids from *Euphorbia fischeriana* Steud, one of which exhibited significant antiviral activity against COVID-19 by directly binding to the virus's RNA-dependent RNA polymerase. Additionally, Neel et al. reported that two major metabolites, gymnemic acid IV and gymnestrogenin from the leaves of *Gymnema sylvestre* were effective against *Penicillium digitatum* 6952, *Penicillium expansum* 2995, and *Aspergillus flavus* 6678.

In recent years, with the maturity of high-throughput sequencing technology and advancements in bioinformatics analysis methods, researches on the biosynthesis and regulation of natural products have expanded from model plants to medicinal plants. Wu et al. compared the function and sequences of key oil biosynthetic genes across different cultivars of *Perilla frutescens*, discovering crucial amino acid residues responsible for the catalytic activity of Δ12 fatty acid desaturases. Zeng et al. unveiled that the influence of sugar metabolism on oil synthesis varies throughout distinct fruit development stages in oil plant *Symplocos paniculata*. Senevirathne et al. summarized the mechanism of oxylipin signaling molecules to regulate phytocannabinoids production in *Cannabis sativa*. MYB transcription factors (TFs) are a class of important TFs that regulate the synthesis of various plant secondary metabolites. Wang et al. conducted a genome-wide analysis of MYB TFs in four *Rheum* L. plants, identifying 1054 MYB genes, with 12 characterizing a key role in anthraquinones biosynthesis. Virus-induced gene silencing (VIGS) is a powerful tool to detect gene functions *in vivo*, particularly in non-model plants. Yu et al. developed an easy and effective VIGS approach, which could facilitate endogenous gene studies in two *Nepeta* species. Finally, considering the complexity of plant secondary metabolism biosynthesis, Yin and Yang discussed an alternative chemoproteomic approach that uses affinity probes to identify active enzymes to elucidate biosynthetic pathway.

To date, our knowledge of the biosynthesis and regulation of specialized metabolites across various plant species remains limited. This Research Topic significantly advances the knowledge framework for integrated approach to understand the production and functions of plant natural products. We hope this Research Topic can serve as a landmark reference for future research in the field.

Author contributions

ZY: Writing – original draft. YW: Writing – review & editing. BJ: Writing – review & editing. XF: Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. We thank the funding from the National Natural Science Foundation of China (No. 32360111 and 32371485), Yunnan Revitalization Talent Support Program "Top Team" Project (202305AT350001), Science Fund for Distinguished Young Scholars of Jiangxi Province (No. 20224ACB215004), Double Thousand Plan of Jiangxi Province (No. jxsq2023101106), Key Research and Development Program of

Jiangxi Province (No. 20243BBI91010), Henan Province Outstanding Youth Foundation (No. 242300421097), and the Program for Scientific and Technological Innovation Talents in Universities of Henan Provincial Education Department (25HASTIT057).

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

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

Publisher's note

References

Afendi, F. M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., et al. (2012). KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.* 53, e1. doi: 10.1093/pcp/pcr165

Dixon, R. A., and Paiva, N. L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell* 7, 1085–1097. doi: 10.1105/tpc.7.7.1085

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 11 June 2024
DOI 10.3389/fpls.2024.1424956

OPEN ACCESS

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Yaolei Mi,
China Academy of Chinese Medical Sciences, China
PeiWei Liu,
Chinese Academy of Medical Sciences and Peking Union Medical College, China

*CORRESPONDENCE
Yanqin Xu
✉ xuyanqin927@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 29 April 2024
ACCEPTED 31 May 2024
PUBLISHED 11 June 2024

CITATION
Tian S, Lv X, Li M, Tang Q, Huang H, Hu S, Li F and Xu Y (2024) Metabolomic and transcriptomic analysis of the flavonoid biosynthesis pathway in *Epimedium sagittatum* (Sieb. et Zucc.) Maxim. from distinct locations.
Front. Plant Sci. 15:1424956.
doi: 10.3389/fpls.2024.1424956

Metabolomic and transcriptomic analysis of the flavonoid biosynthesis pathway in *Epimedium sagittatum* (Sieb. et Zucc.) Maxim. from distinct locations

Shuyun Tian†, Xingbin Lv†, Mengxue Li, Qin Tang, Huilian Huang, Shengfu Hu, Fengqin Li and Yanqin Xu*

College of Pharmacy, Jiangxi University of Chinese Medicine, Nanchang, China

Epimedium sagittatum (Sieb. et Zucc.) Maxim. (ESM) which accumulates several principal flavonoid compounds including epimedin A, B, C and icariin, is extensively utilized in traditional herbs for sexual dysfunction, osteoporosis etc. In China, ESM has a wealth of wild plant resources and characterized by significant variability in medicinal compounds accumulation. Understanding the diversity of ESMs can lead to better utilization of these plant resources. In this study, we integrated the metabolomic and transcriptomic analysis of three ESMs that originated in Anhui, Hubei and Jiangxi in China. Results showed that the flavonoid biosynthesis as well as the related gene expression in these ESMs revealed substantial differences. For example, the epimedin A, B, C and icariin as well as some related gene expression in ESMs from Anhui are significantly lower than those of in others. These results suggested that the ESMs from wild population without quality checkout may not be suitable for directly use as the materials for preparation of Chinese medicine and ESMs with different accumulation of metabolites could be used for distinct applications.

KEYWORDS

Epimedium sagittatum **(Sieb. et Zucc.) Maxim, medicinal plant resource, flavonoid biosynthesis, gene expression, diversity**

Introduction

Epimedium sagittatum (Sieb. et Zucc.) Maxim (ESM), as one of the *Epimedii Folium* (named as Yin-yang-huo in Chinese) in the Chinese Pharmacopoeia (Pharmacopoeia of the People's Republic of China, 2020), contains specific medicinal flavonoid compounds. It is extensively utilized in traditional Chinese medicine for treating sexual dysfunction

(Jiang et al., 2016), osteoporosis (Shi et al., 2022), tumors (Li et al., 2015) and cardiovascular disease (Zeng et al., 2022). Notably, the Icaritin Soft Capsule, derived from Epimedii Folium, was launched as a pioneering new drug in China in 2022 (The National Medical Products Administration of China, 2022), yielding significant economic benefits.

To date, over 300 compounds have been identified within the *Epimedium* genus, including four specific compounds: epimedin A, B, C and icariin (Ren et al., 2018), which serve as bioactive markers for *Epimedium*'s quality control (Xie et al., 2010; Ma et al., 2011). China has abundant wild resources of *Epimedium*, with considerable variation in the levels of these four compounds, impacting the raw material's final quality (Li et al., 2021). Although several studies have been reported the accumulation of these main compounds in *Epimedium*, there are still lacking reports for some populations of wild *Epimedium*, such the ESMs. 4

Here, we conducted an integrated metabolomic and transcriptomic analysis of three ESMs from Anhui, Hubei, and Jiangxi, China (Figures 1A, B). This comprehensive analysis revealed distinct metabolomic and gene expression profiles among the three ESMs, expanding our understanding of metabolites accumulation in ESMs and supporting their further potential applications in Chinese medicine preparation.

Materials and methods

Plant materials

Three-year-old plants of ESMs were collected from Anhui (Tangkou Town, Huangshan City, N 118° 10′ 21.3”, E 30°06′ 3.72”), Hubei (Sanlifan Town, Huanggang city, N 115° 17′ 59.88”, E 30° 54′ 21.18”) and Jiangxi (Songxi Town, Jiujiang City, N 115° 02′ 0.72”, E 29°19′ 37.38”) provinces in 2012, and then were grown in the green garden for *Epimedium* cultivation (N 28° 40′ 33.8″, E 115° 44′ 21.4″) at Jiangxi University of Chinese Medicine for nine years to stabilize the active components *in vivo*. Mature trifoliolate leaves (Figure 1A) from three ESMs were collected and flash-frozen in liquid nitrogen, freeze-dried under vacuum and stored at -80 °C for subsequent experiments.

Metabolite extraction

Mature trifoliolate leaves from three ESMs were sampled and lyophilized in a vacuum freeze-dryer (Scientz-100F) and ground into powder with zirconia beads (MM 400, Retsch) for 1.5 min at 30 Hz. Then, 100 mg of powders were suspended in 1.2 mL of 70% aqueous

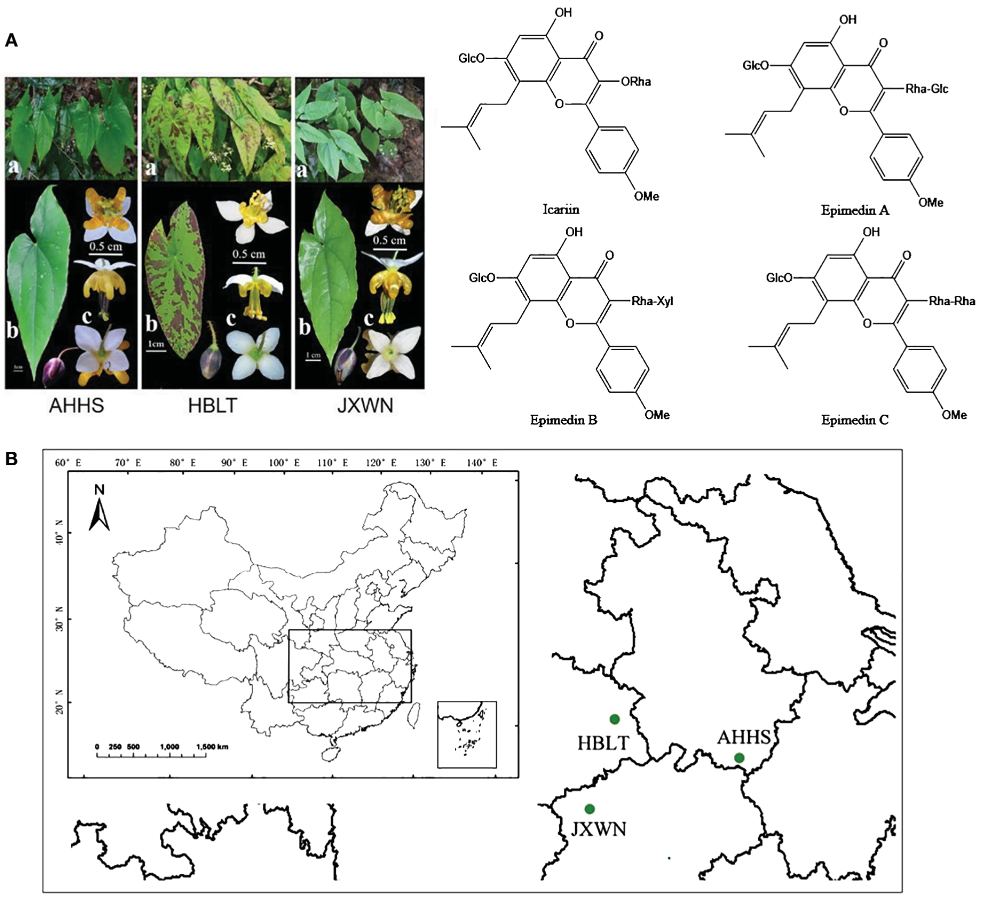


FIGURE 1
The resource of ESMs in China. **(A)** The phenotype of three ESMs (left panel) and main flavonoids compounds (right panel) of ESM. a, whole plants of three ESMs; b, leaf; c, flower; **(B)** The collecting locations of three ESMs including HBLT, AHHS, JXWN in China.

methanol solution. The solution was vortexed for 30 minutes and placed in a refrigerator at 4°C overnight for metabolite extraction. Then, the samples were centrifuged at 12,000 rpm for 10 min and the supernatants were filtered with the 0.22 μm filter (SCAA-104, ANPEL) before UPLC-MS/MS analysis.

UPLC-MS/MS analysis

The extracts were analyzed using a UPLC-ESI-MS/MS system (UPLC, SHIMADZU Nexera X2; MS, Applied Biosystems 4500 Q TRAP) equipped with a Agilent SB-C18 UPLC column (1.8 μm, 2.1 mm × 100 mm). Mobile phase A consisted of 0.1% formic acid in acetonitrile, while mobile phase B consisted of 0.1% formic acid in water. The column was eluted in the running program at a flow rate 0.35 ml/min: 95: 5 (v/v) at 0 min, 5:95 (v/v) at 9.0 min, 5:95 v/v at 10.00 - 11.10 min, 95:5 v/v at 14 min. The elution was then detected by an ESI-triple quadrupole-linear ion trap (QTRAP) - MS (AB4500 Q TRAP UPLC/MS/MS System) which equipped with an ESI Turbo Ion-Spray interface. The ESI source operation parameters were set as followed: turbo spray ion source; source temperature 550°C; ion spray voltage (IS) 5500 V (positive ion mode) - 4500 V (negative ion mode); the pressure of ion source gas I (GSI), gas II (GSII), curtain gas (CUR) was set at 50, 60, and 25.0 psi, respectively; the collision-activated dissociation (CAD) was set to high. Instrument tuning and mass calibration were performed with 10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired in MRM experiments with collision gas (nitrogen) set to medium. DP and CE for individual MRM transitions were done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period.

Qualitative analysis of primary and secondary MS data was conducted using the self-compiled database MWDB (MetWare biological science and Technology Co., Ltd., Wuhan, China). Repeated signals of K^+, Na^+, $NH4^+$, and other small molecular weight compounds were eliminated during identification. The quantitative analysis of metabolites were based on the MRM mode. The characteristic ions of each metabolite were screened through the QQQ mass spectrometer to obtain the signal strengths. The mass spectrometry data were processed using Analyst 1.6.3 software (AB SCIEX, Ontario, Canada). Metabolites with VIP ≥1, $|\log_2 \text{(fold change)}| \geq 1$ and $p < 0.05$ were considered as the differential metabolites.

Transcriptome library preparation

Mature trifoliolate leaves were sampled and mixed for total RNA extraction using the RNAprep Pure Plant Kit (Tiangen, China) following the manufacturer's instructions. The mRNA was pulled down using 5 μg total RNA using the magnetic Dynabeads Oligo $(dT)_{25}$ (Invitrogen, USA). The transcriptome libraries were prepared using the VAHTS Universal DNA Library Prep for Illumina V2 Kit (Vazyme, China) and sent to Illumina HiSeq 3000 for sequencing in paired-end mode (PE150).

Transcriptome *de novo* assembly, gene annotation, expression calculation and differentially analysis

The raw data of RNA-seq were firstly trimmed to obtain clean reads with high sequencing quality (Phred score ≥ 30) and length longer than 50 bases with Trim Galore software (https://github.com/FelixKrueger/TrimGalore). The clean data were then aligned to the plant rRNA sequences database (https://www.plantrdnadatabase.com) for removing rRNA reads. The filtered RNA-seq reads were merged for *de novo* assembly using the Trinity software with min_kmer_cov set to 4 and other parameters in default values (https://github.com/trinityrnaseq). TransDecoder software was employed for the prediction of open reading frames (ORFs) (https://github.com/TransDecoder) and CD-Hit (https://sites.google.com/view/cd-hit) was used to remove the redundancy of protein sequences by the parameters set as "-c 0.8 -n 2". For the gene annotation, the predicted protein sequences were submitted to the EGGNOG-mapper (http://eggnog-mapper.embl.de/) by the default parameter for GO, KEGG annotation etc. The transcription factors were identified by the iTAK pipeline (http://itak.feilab.net/cgi-bin/itak/index.cgi). Salmon software was employed to measure to expression level (TPM, Transcripts Per Kilobase of exonmodel per Million mapped reads) for each unigenes and DEseq2, a R package was used for identification of the differentially expression genes (DEGs). The DEGs with the fold change > 2 and the FDR < 0.05 were further extracted for the downstream analysis.

Quantitative PCR

Total RNA was isolated from the mature trifoliolate leaves of ESMs using the UNIQ-10 column TRIzol Total RNA Extraction Kit (Sangon Biotech), following the manufacturer's instructions. The integrity and concentration of the RNA were assessed via gel electrophoresis and quantification with an SMA 4000 microspectrophotometer (Merinton Instrument). Subsequently, cDNA synthesis was conducted using 1 μg of total RNA and the Maxima Reverse Transcriptase (Thermo Scientific). Quantitative PCR (qPCR) analyses were performed with the SYBR Premix ExTaq Mix (Takara, Japan) on a LightCycler 480 II Real-Time PCR System (Roche), employing a thermal cycling protocol of an initial denaturation at 95°C for 3 min, followed by 45 cycles of 95°C for 15 s and 60°C for 30 s. The relative gene expression level was normalized to an internal control, *EsActin* (Huang et al., 2015). Primers used are listed in Supplementary Table S1.

Results

Metabolic differences of three ESMs

In this study, we focused on three wild ESMs (Figure 1A) collected from Anhui, Jiangxi, and Hubei provinces, which were subsequently named as AHHS, JXWN, and HBLT (Figure 1B), respectively. To compare the metabolite profiles among AHHS, JXWN, and HBLT, leaf samples were analyzed using UPLC-MS/MS based on the total ion chromatogram (TIC outline) (Supplementary Figure S1).

In total, thousands of metabolites were identified, including 151 flavonoids that were the focus of this study (Supplementary Table S2). Principal component analysis (PCA) revealed that 58.64% of the total variance among the samples could be attributed to PC1 (38.31%) and PC2 (20.33%) (Figure 2A), demonstrating significant variation in flavonoid compositions among the leaf samples from AHHS, JXWN, and HBLT.

Additionally, hierarchical cluster analysis (HCA) was performed on these metabolites. Generally, the 151 flavonoid-related compounds were categorized into eight clusters (Figure 2B; Supplementary Table S3). Flavonoid compounds in Cluster 7 were more abundant in AHHS compared to HBLT and JXWN, whereas those in Cluster 1 were less abundant in AHHS. We also examined twelve compounds associated with the biosynthesis of icariin and epimedins A, B, C, and the result indicated that most of these compounds, such as the epimedins, were present at higher levels in JXWN and HBLT than in AHHS (Figure 2C).

These findings suggested significant differences in flavonoid profiles among these ESMs from different locations. Notably, ESM resources like AHHS, which contain lower levels of icariin and epimedins A, B, C, may not be directly suitable for medical preparation.

De novo transcriptome assembly of ESMs

Although the complete biosynthetic pathways of icariin and epimedins A, B, C in ESMs remain to be elucidated, as indicated in Figure 2C, several genes implicated in flavonoid biosynthesis have been identified. These genes provide a partial representation of the gene expression profiles across the three ESM s. Consequently, we conducted a transcriptomic analysis of these ESMs. Nine RNA-seq libraries were constructed from leaf samples of AHHS, JXWN, and HBLT (three biological replicates for each). Approximately 43.8 to

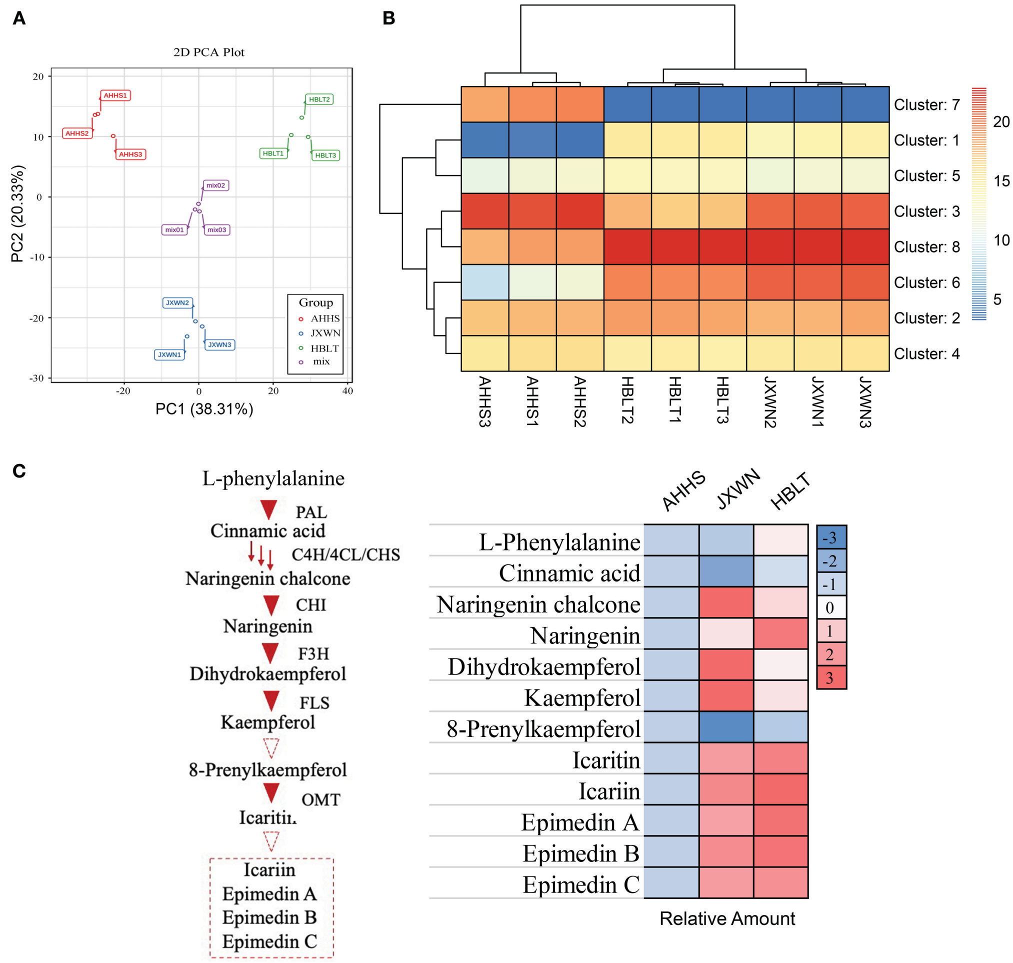


FIGURE 2
The measurement of metabolites in three ESMs including AHHS, JXWN and HBLT. **(A)** Principal component analysis (PCA) of the flavonoids compounds among AHHS, JXWN and HBLT. **(B)** Heatmap showed the eight clusters (by K-mean algorithm) of these flavonoids compounds among AHHS, JXWN and HBLT. **(C)** Twelve flavonoids compounds related to the Icariin or Epimedin biosynthesis was selected to showed the different level among AHHS, JXWN and HBLT. Left panel, the deduced biosynthesis pathway of Icariin or Epimedins. The related genes also showed besides the triangles. Right panel, the relative amount of these twelve flavonoids compounds among AHHS, JXWN and HBLT. The mean amount was normalized to that of AHHS.

52.7 million raw reads were generated for each library (Supplementary Table S4).

Given the absence of a reference genome for ESM, we employed a *de novo* transcriptomic assembly approach to identify the full-length sequences of flavonoid-related genes. Initially, low-quality reads (base accuracy below 99.9%) and reads shorter than 50 bp were discarded. Subsequently, rRNA sequences were removed from the filtered reads by alignment to plant rRNA sequences from the Plant rDNA Database. The remaining reads were then assembled using the Trinity software (Haas et al., 2013). The assembled unigenes were analyzed for open reading frame (ORF) prediction using TransDecoder, and proteins shorter than 150 amino acids were excluded, resulting in 68,413 unigenes. These unigenes were further deduplicated using the CD-Hit program, yielding 31,945 non-redundant unigenes for further analysis (Supplementary Table S5). The contig N50 value, average contig length, and GC content of these unigenes were 1,344 bp, 1,081.2 bp, and 44.06%, respectively.

For annotation, the protein sequences of the 31,945 unigenes were aligned to multiple databases, including COG, Pfam, GO, and KEGG, using the EGGNOG-mapper pipeline. A total of 24,322 unigenes (76.14%) were successfully annotated (Supplementary Table S6). Specifically, 22,960 (71.9%), 18,061 (56.5%), and 7,746 (24.2%) unigenes were annotated in the COG, Pfam, and KEGG pathway databases, respectively. Additionally, 12,938 unigenes (40.5%) were categorized into the three principal GO categories, including Molecular Function, Cellular Component, and Biological Process (Figure 3; Supplementary Table S6). Moreover, 1,231 unigenes (3.9%) were predicted to be transcription factors using the iTAK pipeline and were classified into 65 families (Supplementary Table S7). The five largest transcription factor families were bHLH, AP2/ERF, MYB, C2H2, and C3H, with approximately 68 to 88 unigenes in each category, which is consistent with the distribution observed in other plant species according to the Plant TFDB database (Guo et al., 2008).

Identified differentially expressed genes related to flavonoid biosynthesis

Observations revealed a unique accumulation pattern of flavonoid compounds across three ESMs. This prompted an investigation into the expression of genes associated with flavonoid biosynthesis within these ESMs. To delineate differentially expressed genes (DEGs), we initially quantified the Transcripts per Million (TPM) for each unigene using the Salmon software (Patro et al., 2017). Subsequently, DEGs (HBLT vs. AHHS and JXWN vs. AHHS) were identified employing the DEseq2 program (Love et al., 2014). A comprehensive analysis yielded

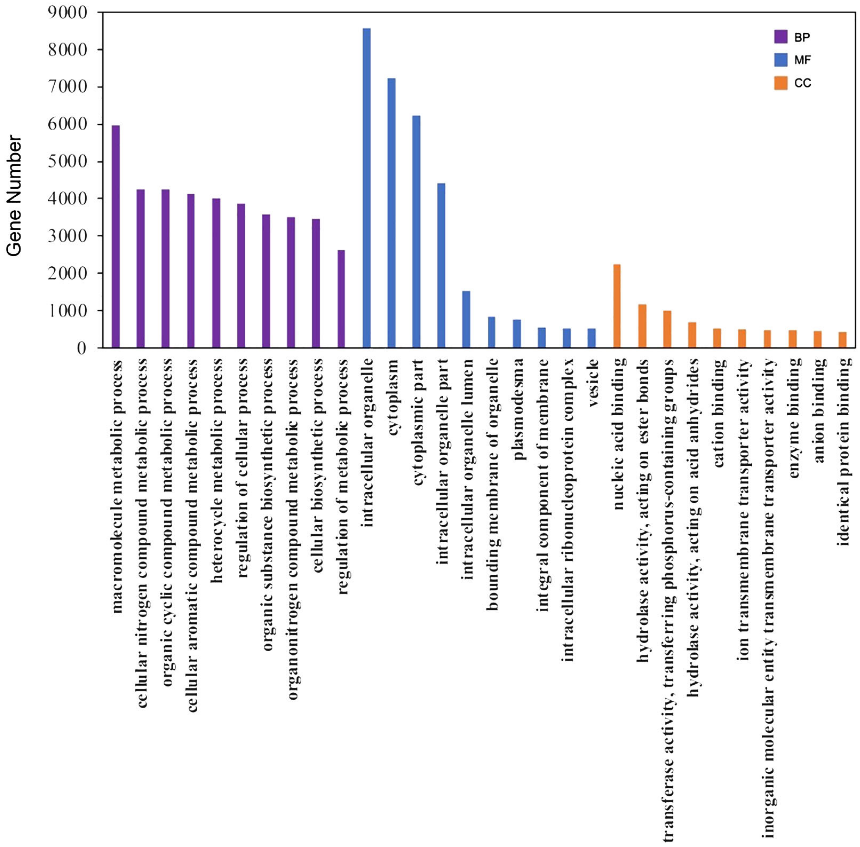


FIGURE 3
GO annotation of the unigenes. The top 10 ranked GO terms in the biological process (BP), molecular function (MF) and cellular component (CC) categories were shown as the examples here.

7,321 DEGs in HBLT relative to AHHS, comprising 3,942 down-regulated and 3,379 up-regulated genes. In contrast, JXWN exhibited 4,504 DEGs compared to AHHS, with 1,771 down-regulated and 2,733 up-regulated genes (Figures 4A, B; Supplementary Table S8). Notably, genes pivotal to flavonoid biosynthesis, including *phenylalanine ammonia-lyase* (*PAL*), *cinnamate 4-hydroxylase* (*C4H*), *4-coumarat: CoA ligases* (*4CL*), *chalcone synthase* (*CHS*), *flavanone 3-hydroxylase* (*F3H*), and *anthocyanin O-methyltransferase* (*OMT*), demonstrated differential expression in HBLT or JXWN relative to AHHS. This differential gene expression correlates with the distinct flavonoid metabolite profiles observed among the three ESMs.

Co-expression analysis to identify potential TFs involved in regulation of DEGs

Transcription factors (TFs) play a crucial role in regulating flavonoid biosynthesis, as previous studies (Ravaglia et al., 2013; Wang et al., 2018). To elucidate the TFs potentially regulating the differentially expressed genes (DEGs) associated with flavonoid biosynthesis, we conducted a co-expression analysis, by calculating the Pearson correlation coefficient (PCC) between TFs and DEGs involved in flavonoid biosynthesis. Gene pairs exhibiting a PCC value over 0.9 and p value < 0.05 were considered as co-expression partners. This analysis identified 1,148 co-expressed DEG - TF pairs (Supplementary Table S9), suggesting a complex regulatory network influencing flavonoid biosynthesis. Notably, two genes, *F3'H* (ESM_Unigene_26659) and *AoMT* (ESM_Unigene_31764) correlated with the epimedins accumulation (Supplementary Table S10) and their co-expressed TFs (Figure 5A), significantly upregulated in HBLT and JXWN compared to AHHS (Figure 5B). These findings were further validated through qPCR (Figure 5C). This supports the hypothesis that these TFs may play an important role in the regulation of DEGs involved in flavonoid biosynthesis.

Discussion

In this study, we collected three ESM s from distinct locations in China, including Hubei, Jiangxi and Anhui (Figures 1A, B). Metabonomic and transcriptomic analyses revealed that the levels of flavonoids, such as icariin and epimedin A, B, C, as well as the expression levels of related genes, varied significantly among these ESMs (Figures 2, 4). Additionally, we identified several transcription factors (TFs) that may play crucial roles in the

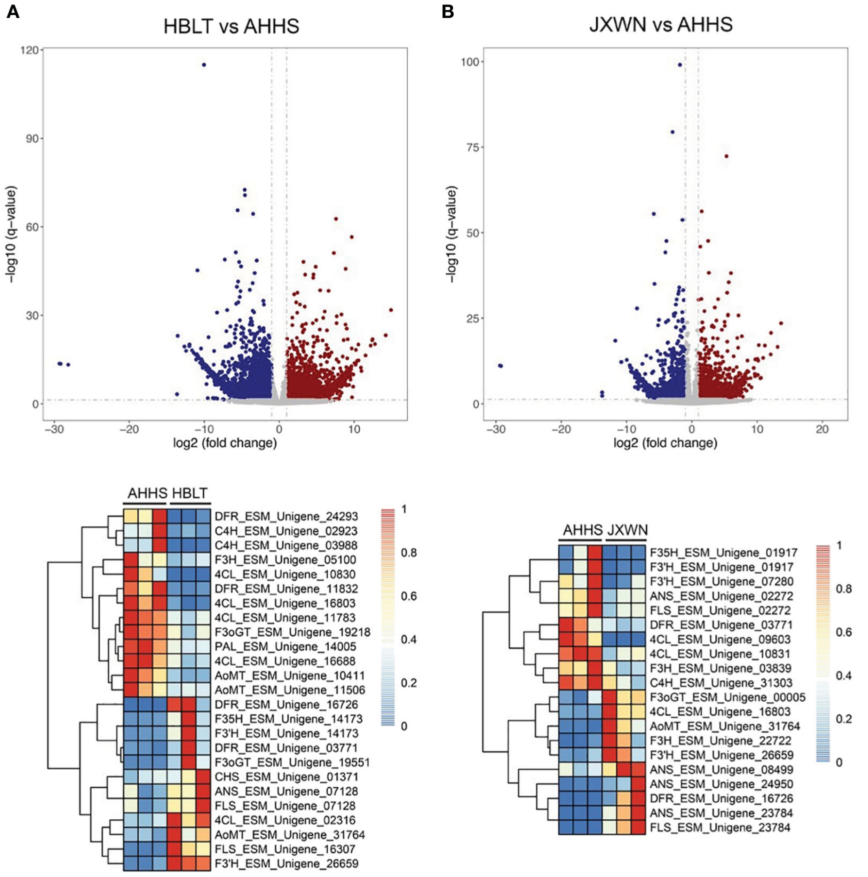


FIGURE 4
Identification of the differentially expressed genes (DEGs) among the ESMs. **(A)** Upper panel: The DEGs identified by comparing HBLT to AHHS; Lower panel: Heatmap showed the TPM of DEGs related to the flavonoid biosynthesis. **(B)** The DEGs identified by comparing JXWN to AHHS. Lower panel: Heatmap showed TPM of the DEGs related to the flavonoid biosynthesis. DEGs was defined as the genes with two-fold change and the FDR < 0.05.

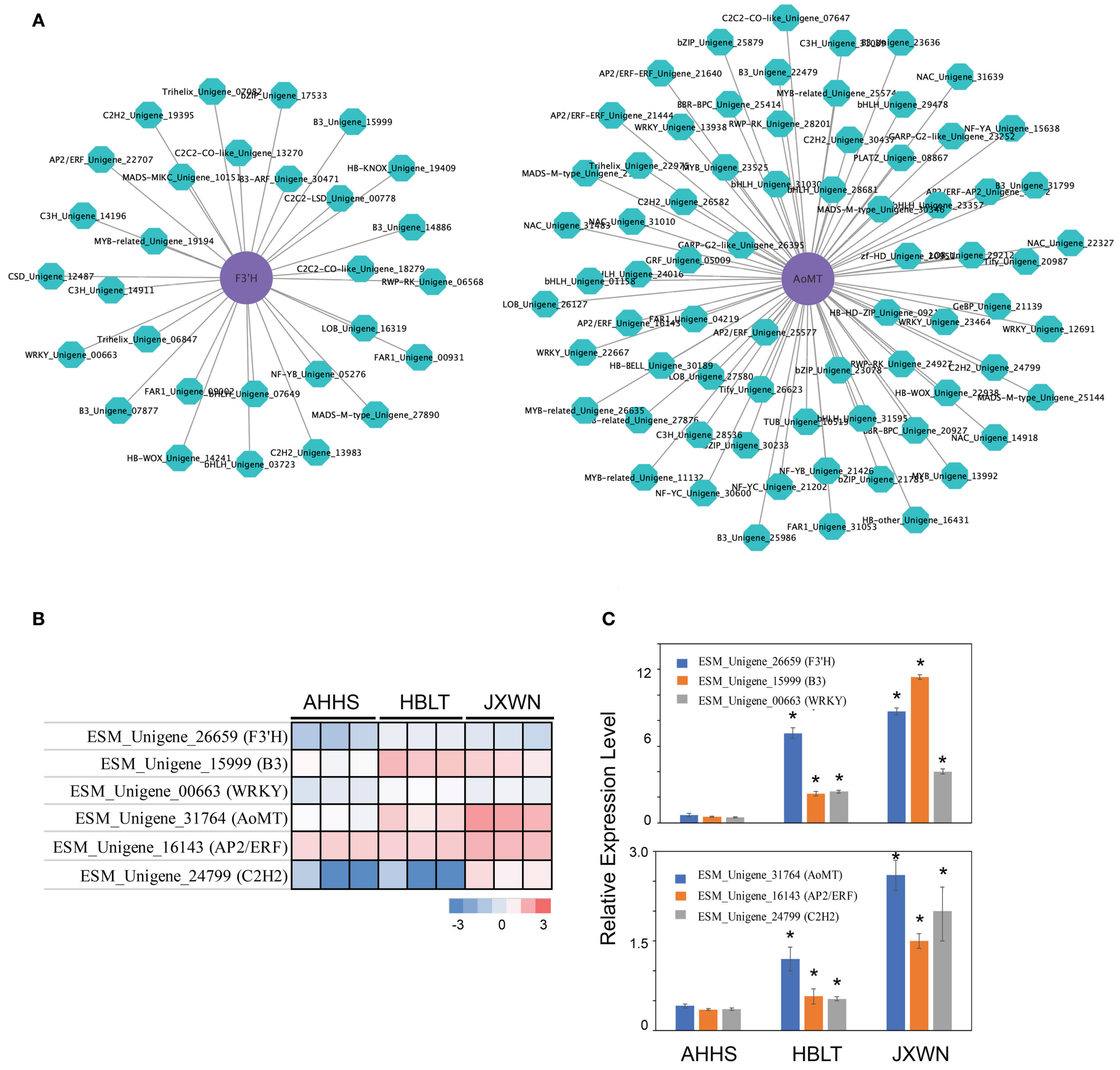


FIGURE 5
Identification of the potential key TFs involved in regulation of differentially expressed flavonoid biosynthesis related genes. **(A)** The co-expression network between the TF and the flavonoid biosynthesis related genes. The *AoMT* and *F3'H* (purple cycle) were used as the examples to show the potential regulators (green cycle). **(B)** The expression level of *AoMT*, *F3'H* as well as their potential regulators were validated by the qPCR. The heatmap (left panel) showed the TPM from RNA-seq of each gene. The qPCR validation of each gene were showed in right panel. **(C)** It represents the validation results of AoMT, F3'H, and their potential regulators via qPCR. "*" indicated the significant differences observed in HBLT, JXWN compared to AHHS ($p < 0.05$, by student's test).

regulation of DEGs involved flavonoid biosynthesis, as suggested by the co-expression network analysis (Figure 5). Collectively, these findings provide a valuable data resource for ESMs.

As an important Chinese medicinal plant, the quality of raw ESM materials has garnered widespread attention (Xie et al., 2010; Ma et al., 2011). Although numerous ESM populations originate from China, our results indicate that ESMs from different locations may exhibit variations in metabolite concentrations, particularly those with medicinal properties, such as icariin and epimedin (Figure 2). For instance, AHHS has the lowest flavonoid content, is not suitable for direct use in the preparation of Chinese medicine or for the extraction of active compounds. These findings underscore the necessity of quantifying the medicinal activity of raw ESM materials prior to market distribution.

In this study, we observed distinct gene expression profiles among three ESMs from Hubei, Jiangxi, and Anhui, with particular attention to the AHHS from Anhui, which exhibited the lowest flavonoid levels (Figure 2C). Despite identifying differential gene expression in AHHS compared to the other two ESMs, the majority of flavonoid biosynthesis-related genes detected did not elucidate the precise causes of the varied metabolite accumulation (Zou et al., 2016; Liu et al., 2021; Lou et al., 2021). Two primary reasons are posited: (1) the complete biosynthetic pathways of major flavonoids such as icariin and epimedin remain elusive to date (Huang et al., 2015; Wang et al., 2021), precluding the direct detection of gene expression responsible for their biosynthesis; (2) the absence of a reference genome for ESMs hinders the examination of genomic variations among these samples. Nevertheless, the finding that

AHHS containing the lowest flavonoid levels has piqued interest, potentially serving as a critical genetic resource for future studies, such as constructing a recombinant inbred line for Quantitative Trait Locus (QTL) analysis (Mauricio, 2001).Taken together, these ESMs with varying flavonoid contents represent significant plant resources for further investigation.

Data availability statement

All the raw data in this study can be found in the China National Center for Bioinformation (https://www.cncb.ac.cn) under the BioProject accession number PRJCA024022.

Author contributions

ST: Writing – original draft, Writing – review & editing. XL: Writing – original draft, Writing – review & editing. ML: Writing – review & editing. QT: Writing – review & editing. HH: Writing – review & editing. SH: Writing – review & editing. FL: Writing – review & editing. YX: Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by National Natural Science Foundation of China (82060684) and Jiangxi University of Chinese Medicine Science and Technology Innovation Team Development Program (CXTD22002).

Acknowledgments

We are thankful to the Wuhan Metware Biotechnology Co., Ltd. for its technical support in metabolite detection and RNA sequencing.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1424956/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The raw TIC of mass spectrometry for three ESMs. **(A)** ion in positive mode and **(B)** anion mode. The blue, red and green lines are indicated the samples from JXWN, HBLT and AHHS.

References

Guo, A., Chen, X., Gao, G., Zhang, H., Zhu, Q. H., Liu, X., et al. (2008). PlantTFDB: A comprehensive plant transcription factor database. *Nucleic Acids Res.* 36, D966–D969. doi: 10.1093/nar/gkm841

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084

Huang, W., Zeng, S., Xiao, G., Wei, G., Liao, S., Chen, J., et al. (2015). Elucidating the biosynthetic and regulatory mechanisms of flavonoid-derived bioactive components in *Epimedium sagittatum*. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00689

Jiang, J., Zhao, B., Song, J., and Jia, X. (2016). Pharmacology and clinical application of plants in *Epimedium* L. *Chin. Herbal Medicines* 8, 12–23. doi: 10.1016/S1674-6384(16)60003-5

Li, B., Lima, M. R. M., Nie, Y., Xu, L., Liu, X., Yuan, H., et al. (2021). HPLC-DAD fingerprints combined with multivariate analysis of Epimedii Folium from major producing Areas in Eastern Asia: effect of geographical origin and species. *Front. Pharmacol.* 12. doi: 10.3389/fphar.2021.761551

Li, C., Li, Q., Mei, Q., and Lu, T. (2015). Pharmacological effects and pharmacokinetic properties of icariin, the major bioactive component in Herba Epimedii. *Life Sci.* 126, 57–68. doi: 10.1016/j.lfs.2015.01.006

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

Lou, H., Hu, L., Lu, H., Wei, T., and Chen, Q. (2021). Metabolic engineering of microbial cell factories for biosynthesis of flavonoids: A Review. *Molecules* 26, 4522. doi: 10.3390/molecules26154522

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Ma, H., He, X., Yang, Y., Li, M., Hao, D., and Jia, Z. (2011). The genus *Epimedium*: An ethnopharmacological and phytochemical review. *J. Ethnopharmacology* 134, 519–541. doi: 10.1016/j.jep.2011.01.001

Mauricio, R. (2001). Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* 2, 370–381. doi: 10.1038/35072085

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197

Pharmacopoeia of the People's Republic of China (2020). *Chinese pharmaopoeia* (Beijing: China Medical Science Press), 340.

Ravaglia, D., Espley, R. V., Henry-Kirk, R. A., Andreotti, C., Ziosi, V., Hellens, R. P., et al. (2013). Transcriptional regulation of flavonoid biosynthesis in nectarine (*Prunus persica*) by a set of R2R3 MYB transcription factors. *BMC Plant Biol.* 13, 68. doi: 10.1186/1471-2229-13-68

Ren, F. C., Jiang, X. J., Wen, S. Z., Wang, L. X., Li, X. M., and Wang, F. (2018). Prenylated 2-Phenoxychromones and flavonoids from *Epimedium* brevicornum and revised Structures of Epimedonins A and B. *J. Nat. Prod.* 81, 16–21. doi: 10.1021/acs.jnatprod.7b00514

Shi, S., Wang, F., Huang, Y., Chen, B., Pei, C., Huang, D., et al. (2022). *Epimedium* for osteoporosis based on western and eastern medicine: an updated systematic review and meta-analysis. *Front. Pharmacol.* 13. doi: 10.3389/fphar.2022.782096

The National Medical Products Administration of China (2022) *The National Medical Products Administration of China: State Drug Administration conditionally approves icaritin softgels for marketing*. Available online at: https://www.nmpa.gov.cn/zhuanti/cxylqx/cxypxx/20220110191736178.html.

Wang, P., Li, C., Li, X., Huang, W., Wang, Y., Wang, J., et al. (2021). Complete biosynthesis of the potential medicine icaritin by engineered Saccharomyces cerevisiae and Escherichia coli. *Sci. Bull.* 66, 1906–1916. doi: 10.1016/j.scib.2021.03.002

Wang, W. L., Wang, Y. X., Li, H., Liu, Z. W., Cui, X., and Zhuang, J. (2018). Two MYB transcription factors (CsMYB2 and CsMYB26) are involved in flavonoid biosynthesis in tea plant [*Camellia sinensis* (L.) O. Kuntze]. *BMC Plant Biol.* 18, 288. doi: 10.1186/s12870-018-1502-3

Xie, P. S., Yan, Y. Z., Guo, B. L., Lam, C. W. K., Chui, S. H., and Yu, Q. X. (2010). Chemical pattern-aided classification to simplify the intricacy of morphological taxonomy of *Epimedium* species using chromatographic fingerprinting. *J. Pharm. Biomed. Anal.* 52, 452–460. doi: 10.1016/j.jpba.2010.01.025

Zeng, Y., Xiong, Y., Yang, T., Wang, Y., Zeng, J., Zhou, S., et al. (2022). Icariin and its metabolites as potential protective phytochemicals against cardiovascular disease: From effects to molecular mechanisms. *Biomedicine Pharmacotherapy* 147, 112642. doi: 10.1016/j.biopha.2022.112642

Zou, L. Q., Wang, C. X., Kuang, X. J., Li, Y., and Sun, C. (2016). Advance in flavonoids biosynthetic pathway and synthetic biology. *Zhongguo Zhong yao za zhi= Zhongguo Zhongyao Zazhi= China J. Chin. Materia Med.* 41, 4124–4128. doi: 10.4268/cjcmm20162207

Frontiers in Plant Science

TYPE Original Research
PUBLISHED 11 June 2024
DOI 10.3389/fpls.2024.1419392

OPEN ACCESS

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Jingang Wang,
Northeast Agricultural University, China
Dr. Pawan Kumar,
Agricultural Research Organization (ARO), Israel

*CORRESPONDENCE
Wei Chen
✉ chenwei@iae.ac.cn
Shuai Yu
✉ yushuai@iae.ac.cn

RECEIVED 18 April 2024
ACCEPTED 27 May 2024
PUBLISHED 11 June 2024

CITATION
Wang Y, Liu H, Yu S, Zhang Y, Huang Y, He X and Chen W (2024) Effects of geographical, soil and climatic factors on the two marker secondary metabolites contents in the roots of *Rubia cordifolia* L. *Front. Plant Sci.* 15:1419392. doi: 10.3389/fpls.2024.1419392

Effects of geographical, soil and climatic factors on the two marker secondary metabolites contents in the roots of *Rubia cordifolia* L.

Yanlin Wang[1,2,3], Huanchu Liu[1,3], Shuai Yu[1,3]*, Yue Zhang[1,3,4], Yanqing Huang[1,3,4], Xingyuan He[1,2,3,4] and Wei Chen[1,2,3,4]*

[1]CAS Key Laboratory of Forest Ecology and Silviculture, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]Liaoning Shenyang Urban Ecosystem National Observation and Research Station, Shenyang, China, [4]Shenyang Arboretum, Chinese Academy of Sciences, Shenyang, China

The growth and quality of medicinal plants depend heavily on environmental variables. The quality of *Rubia cordifolia*, an important medicinal plant, is determined by the two main secondary metabolites of the root, purpurin and mollugin. However, their relationship with environmental factors has not been studied. In this study, the purpurin and mollugin contents of *R. cordifolia* roots from different sampling sites in China were measured using ultra-high-performance liquid chromatography, and the correlations between the two secondary metabolites and environmental variables were analyzed. The results showed that there were significant differences in the contents of purpurin and mollugin in the roots of *R. cordifolia* at different sampling points. The content of purpurin ranged from 0.00 to 3.03 mg g^{-1}, while the content of mollugin ranged from 0.03 to 10.09 mg g^{-1}. The quality of *R. cordifolia* in Shanxi, Shaanxi and Henan border areas and southeastern Liaoning was higher. Liaoning is expected to become a *R. cordifolia* planting area in Northeast China. Correlation and regression analysis revealed that the two secondary metabolites were affected by different environmental factors, the two secondary metabolites contents were positively correlated with longitude and latitude, and negatively correlated with soil nutrients. In addition, higher temperature and shorter sunshine duration facilitated the synthesis of purpurin. Annual precipitation might be the main factor limiting the quality of *R. cordifolia* because it had opposite effects on the synthesis of two major secondary metabolites. Therefore, this study is of great significance for the selection of *R. cordifolia* planting areas and the improvement of field planting quality.

KEYWORDS

***Rubia cordifolia* L., purpurin, mollugin, environmental factors, planting**

1 Introduction

Traditional Chinese Medicine (TCM) has been extensively used to treat various illnesses in China for thousands of years (Wang et al., 2021a). TCM continues to play a key role in safeguarding the health of Chinese people and is attracting increasing global attention (Xue et al., 2012). The idea of integrating Chinese and Western medicines to treat complex disorders has also been adopted by the modern medical community, underscoring the significance of TCM. Since the outbreak of the novel coronavirus pneumonia in 2019, TCM has demonstrated significant promise for the management of COVID-19 (Du et al., 2020; Lyu et al., 2021; Zhao et al., 2021).

The continual development of TCM requires support for further studies on medicinal plants. Only high-quality medicinal plants can be used to treat various diseases. The quality of medicinal plants is not only related to the clinical effect of Chinese medicine but also plays a key role in developing the Chinese medicine industry (Qin et al., 2018). The active components of Chinese medicine are the basis for the efficacy of Chinese medicine, and these innovative drugs are mainly extracted from medicinal plants (Ma et al., 2022). For example, the well-known antimalarial drug artemisinin is extracted from the medicinal plant *Artemisia annua* L (Tu, 1999).

Rubia cordifolia L. is a medicinal plant whose roots can be used to treat various diseases owing to its anti-inflammatory, anticancer, and anti-platelet aggregation properties as various secondary metabolites, including anthraquinones and naphthoquinones, are present in its roots (Figure 1) (Wen et al., 2022). The Chinese Pharmacopoeia usually specifies standards for Chinese herbal medicines, which stipulate that the two main compounds in the roots of *R. cordifolia*, that is, the purpurin content is not less than 0.1% and the mollugin content is not less than 0.4% (Chinese Pharmacopeia Commission, 2020). Purpurin has anticancer, antibacterial, and neuromodulatory effects owing to its strong antioxidant activity (Singh et al., 2021). Similarly, mollugin has a variety of pharmacological effects, including neuroprotective, anti-inflammatory, anticancer, and antiviral effects (Kim et al., 2013; Wang et al., 2017; Li et al., 2020). Therefore, the contents of both compounds are essential for the quality of *R. cordifolia*.

Plant secondary metabolites are not only important for humans but also vital for themselves. Plants synthesize and accumulate various secondary metabolites to resist possible stresses caused by the external environment (Selmar and Kleinwächter, 2013; Loreto et al., 2014). For example, psoralen, the main secondary metabolite in *Psoralea corylifolia* L., increases under chromium and UV-B stress (Pandey et al., 2023). By contrast, a short drought stress period increased the amount of saponins in the roots of *Bupleurum chinense* DC (Yang et al., 2019a). However, an unfavorable growth environment may contribute to the synthesis and accumulation of secondary metabolites but may not be conducive to plant growth, resulting in reduced yield. Further research has shown that medicinal plants grow more effectively and to a higher quality in their appropriate habitats, which is strongly related to the environment in the habitat (Zhan et al., 2022; Zheng et al., 2022).

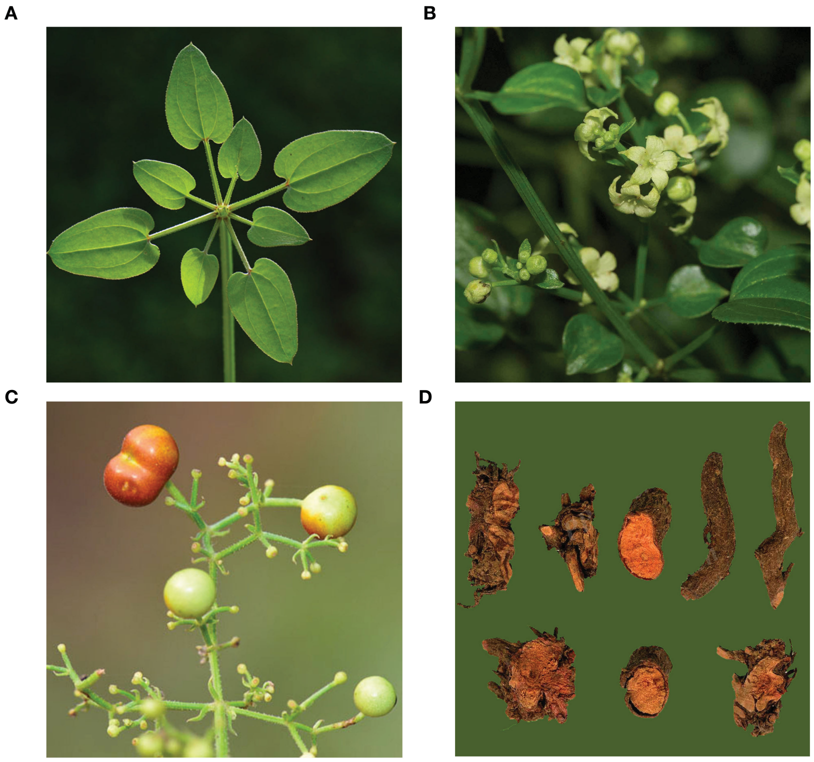


FIGURE 1
Photographs of *R. cordifolia*. **(A)** Leaf; **(B)** Flower; **(C)** Fruit; **(D)** Root. **(A–C)** were obtained from iNaturalist (https://www.inaturalist.org/), and **(D)** was obtained from iPlant (http://www.iplant.cn/).

The synthesis and accumulation of secondary metabolites in medicinal plants are affected by a combination of factors.

Although *R. cordifolia* has long been used as a medicinal plant, research on *R. cordifolia* has mainly focused on the discovery of its secondary metabolites and associated pharmacological effects (Shan et al., 2016; Wen et al., 2022). The quality and productivity of medicinal plants can be improved by understanding the relationship between environmental conditions and secondary metabolite content. Environmental factors may also influence the content of secondary metabolites in *R. cordifolia*; however, little work has been conducted on *R. cordifolia* in this area. Therefore, the objectives of this study were to (1) collect *R. cordifolia* throughout China and determine the content of two major secondary metabolites in the roots at each location; (2) analyze the environmental conditions of the sampling sites, including soil, climate, and topography; and (3) analyze the relationship between secondary metabolite content and environmental factors at each location to explore the primary environmental variables affecting the synthesis and accumulation of *R. cordifolia* root secondary metabolites.

2 Materials and methods

2.1 Sampling

Based on the distribution of *R. cordifolia* and its origin, as described in the literature (Xue et al., 2009; Yu et al., 2017; Yan, 2021), Liaoning, Hebei, Henan, Shanxi, Shaanxi, Gansu, and Sichuan were selected for our sampling (Figure 2). Samples were collected from July to August in 2021 and 2022, and 3–7 whole underground parts of *R. cordifolia* were collected from each sample point, while trying to ensure that the samples had the same plant age. Soil at 0–30 cm depth in the roots was collected for subsequent analysis. The coordinates of each sampling point were recorded for the subsequent acquisition of meteorological and topographical factors, as shown in Table 1.

2.2 Extraction and detection of secondary metabolites

The *R. cordifolia* root powder (0.5 g) was precisely weighed and placed in a stoppered conical flask. Methanol (100 mL) was added to the flask and stored it overnight. Secondary metabolites were extracted using ultrasound (power 250W, frequency 40kHz, 30 minutes). After the ultrasound was completed, the sample was cooled to room temperature, shaken well, and filtered. The filtrate (50 mL) was measured precisely, and evaporate to dryness in a 60°C water bath. The residue was dissolved by adding 20 mL methanol:25% hydrochloric acid (4:1). The solution was heated and hydrolyzed in an 80°C water bath for 30 minutes. The solution obtained after hydrolysis was cooled immediately, 3 ml of triethylamine was added, mixed, transferred to a 25 ml measuring flask, and diluted to volume with methanol. The solution was filtered through a microporous membrane and used for detection (Chinese Pharmacopeia Commission, 2020).

Ultra-high-performance liquid chromatography (UHPLC) (Ultimate 3000 RSLC, Thermo Fisher Scientific, MA, USA) was used to determine the mollugin and purpurin contents in the extract. Octadecylsilane bonded silica gel was used as the filler of chromatographic column, and the 25:50:25 mixture of methanol, acetonitrile, and 0.2% phosphoric acid was used as the mobile phase The detection wavelength is 250 nm. According to the calculations of mollugin and purpurin peaks, the theoretical plate number should not be less than 4000.

2.3 Soil sample processing and analysis

Soil samples were taken back to the laboratory and sieved to remove stones, fine roots, and other debris. The soil was naturally dried and ground, and part of the ground samples were passed through a 0.25 mm sieve to determine the organic matter content of the soil. The remaining samples were sieved through a 1 mm sieve, and the soil pH, available nitrogen, available phosphorus, and available potassium were determined using the methods described by Hu and Wang (2020). Soil pH was determined by the potentiometric method, organic matter content was determined by the potassium dichromate volumetric method, available nitrogen content was determined by the alkaline hydrolysis diffusion method, available phosphorus content was determined by the $NaHCO_3$ extraction-colorimetric method, and available potassium content was determined by the ammonium acetate extraction-flame photometric method.

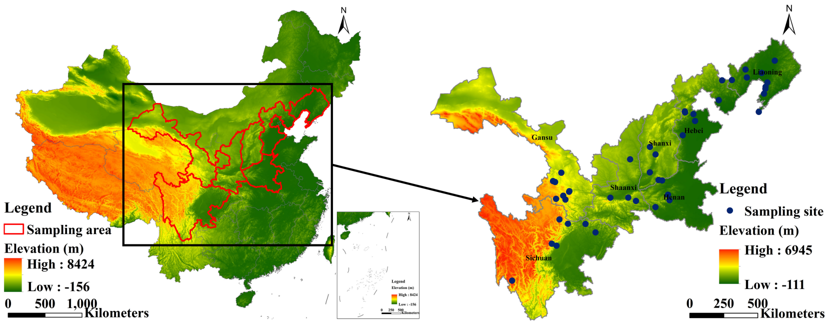


FIGURE 2
Study area and distribution of sampling sites.

TABLE 1 Sampling sites information of *R. cordifolia*.

Sampling site	Province	Long (°E)	Lat (°N)	Ele (m)	pH	OM (g kg^{-1})	AN (mg kg^{-1})	AP (mg kg^{-1})	AK (mg kg^{-1})	TEM (°C)	SSD (h)	PRE (mm)
S1	Gansu	104.568	34.615	2314	8.22	26.87	99.68	9.20	99.28	550.06	2063.25	7.60
S2	Gansu	104.591	34.650	1773	8.53	22.42	98.01	10.25	262.23	470.75	2043.05	9.46
S3	Gansu	103.909	35.893	2420	7.91	65.60	284.41	40.70	686.08	542.15	2459.26	4.59
S4	Gansu	103.909	35.893	2420	7.93	156.35	557.04	13.17	263.13	542.15	2459.26	4.59
S5	Gansu	104.089	34.342	2453	8.22	46.06	192.74	17.30	504.34	658.68	2066.46	6.47
S6	Gansu	104.275	34.085	2107	8.19	54.17	224.38	43.73	891.55	637.46	1936.29	8.70
S7	Gansu	103.411	35.256	2559	8.07	53.25	223.32	13.40	301.15	622.20	2324.24	5.06
S8	Gansu	103.238	35.317	2457	7.99	100.23	445.02	13.47	281.47	637.76	2341.98	4.81
S9	Gansu	103.517	34.124	3045	6.92	241.37	1121.45	14.30	357.81	742.23	2184.48	4.25
S10	Hebei	115.511	39.342	54	8.48	14.56	72.96	16.73	122.67	504.51	2216.87	12.58
S11	Hebei	117.909	40.019	69	7.26	24.99	132.61	14.00	223.80	624.89	2439.15	12.16
S12	Hebei	114.354	38.000	112	8.30	15.83	63.26	8.00	100.02	505.47	2162.80	14.02
S13	Hebei	118.504	41.324	1063	7.02	88.13	379.07	18.13	746.99	565.10	2690.20	5.11
S14	Hebei	114.799	39.528	1067	8.42	18.36	84.64	7.17	188.46	543.41	2509.13	7.46
S15	Hebei	114.788	39.606	1543	8.13	53.55	201.25	12.27	122.92	620.36	2578.69	4.96
S16	Hebei	115.579	38.862	14	8.33	29.97	140.56	100.23	382.16	490.43	2221.99	12.95
S17	Henan	112.601	34.112	377	7.33	28.02	120.52	15.76	205.64	657.62	1842.75	14.95
S18	Henan	111.505	33.356	324	7.80	26.18	109.05	30.87	71.76	795.80	1813.56	15.67
S19	Henan	112.674	33.729	193	6.64	60.28	266.27	33.25	300.64	749.14	1758.83	15.47
S20	Henan	112.233	35.103	496	8.13	28.47	96.22	24.63	252.01	657.78	2038.04	13.65
S21	Liaoning	122.490	40.569	18	7.99	22.16	96.19	21.10	161.43	677.13	2550.11	9.90
S22	Liaoning	122.186	40.181	73	6.45	25.89	118.76	15.00	179.02	667.86	2590.11	9.75
S23	Liaoning	121.972	39.873	69	7.34	22.31	129.30	36.17	518.81	655.27	2560.75	9.99
S24	Liaoning	122.239	40.310	80	6.63	38.95	205.31	73.18	223.42	678.32	2588.21	9.67
S25	Liaoning	123.589	41.908	110	6.13	29.24	150.83	18.60	160.69	690.18	2404.79	8.32
S26	Liaoning	122.169	41.299	5	7.71	10.02	41.18	23.97	136.25	629.54	2566.43	9.61

(Continued)

TABLE 1 Continued

Sampling site	Province	Long (°E)	Lat (°N)	Ele (m)	pH	OM (g kg^{-1})	AN (mg kg^{-1})	AP (mg kg^{-1})	AK (mg kg^{-1})	TEM (°C)	SSD (h)	PRE (mm)
S27	Liaoning	119.391	41.221	436	8.27	34.91	139.81	37.50	337.35	484.38	2677.92	8.65
S28	Liaoning	121.192	38.734	165	6.31	67.86	304.80	12.77	496.08	643.76	2515.16	10.84
S29	Liaoning	120.770	41.199	93	7.73	92.56	164.82	23.33	517.95	505.58	2596.92	9.76
S30	Liaoning	120.786	41.756	144	7.88	87.06	340.89	53.10	306.84	446.45	2653.15	9.30
S31	Liaoning	120.786	41.756	144	8.26	19.56	88.24	12.23	337.17	446.45	2653.15	9.30
S32	Shanxi	111.942	35.153	481	8.36	24.31	86.57	31.08	360.97	640.74	2036.07	13.79
S33	Shanxi	111.456	37.438	1347	8.45	19.99	78.79	34.87	193.37	605.00	2412.83	8.13
S34	Shanxi	111.275	35.721	448	8.41	37.01	78.16	36.77	312.10	483.95	2091.55	14.21
S35	Shanxi	111.881	36.896	916	8.28	33.58	122.30	33.00	277.46	542.56	2197.33	10.79
S36	Shaanxi	109.705	36.695	947	8.49	12.19	41.38	14.15	102.70	558.48	2365.28	10.36
S37	Shaanxi	109.979	33.865	771	7.10	25.46	122.09	5.87	137.81	704.92	1918.56	13.24
S38	Shaanxi	107.919	34.163	544	7.88	44.24	160.73	21.37	189.28	687.90	1796.05	13.79
S39	Shaanxi	109.374	34.114	686	7.47	26.16	84.93	13.08	213.20	732.79	1933.74	12.87
S40	Sichuan	103.815	32.748	3395	7.09	334.23	1168.00	36.90	402.24	939.84	1815.98	3.12
S41	Sichuan	104.497	32.463	1012	8.19	54.43	211.21	11.80	175.53	873.64	1414.93	13.76
S42	Sichuan	105.880	32.447	540	8.54	16.05	60.32	6.10	112.61	1114.55	1315.61	16.10
S43	Sichuan	103.616	30.989	721	8.17	46.68	179.02	15.67	271.32	1067.28	1037.23	16.40
S44	Sichuan	100.349	28.566	3012	7.86	100.78	400.23	91.53	1038.73	505.66	2312.12	9.76
S45	Sichuan	106.660	31.870	436	8.38	26.07	170.23	7.23	96.28	1206.15	1323.29	17.31
S46	Sichuan	103.250	31.118	1901	8.15	31.71	108.10	11.83	75.15	1084.87	1304.95	8.38

Long, Longitude; Lat, Latitude; Ele, Elevation; OM, Organic matter; AN, Available nitrogen; AP, Available phosphorus; AK, Available potassium; TEM, Annual mean temperature; SSD, Annual sunshine hours; PRE, Annual precipitation.

2.4 Access to climate and geographic factors

The annual mean temperature (TEM), annual precipitation (PRE), and annual sunshine hours (SSD) for 2010–2020 were obtained from the Resource and Environment Science and Data Center (https://www.resdc.cn/) (Xu, 2017). Elevation data were obtained from WorldClim (https://www.worldclim.org/). Finally, the climatic and geographic factors of each sampling site were extracted using the extraction tool in ArcGIS.

2.5 Statistical analysis

The data were pre-processed using Microsoft Excel 2019. The Spearman correlation analysis was conducted between environmental variables and secondary metabolites using the corrplot package in R 4.2.2. The stepwise regression was analyzed using the car package in R 4.4.2. Corresponding plots were constructed using ggplot2 package in R 4.2.2 and ArcGIS 10.7.

3 Results

3.1 Soil, climatic and geographic characteristics of sampling sites

Sampling sites were covered from the east to the west and from the north to the south of China (Figure 2). The longitude spanned 23.24°, and the latitude spanned 13.34°. The elevation in the study area varied from 5 to 3395 m, rising from east to west (Table 1).

Soil properties varied considerably between the sampling sites (Table 1). The average soil pH varied between 6.13 and 8.54, the average organic matter varied between 10.02 and 334.23 g kg^{-1}, the average available nitrogen content ranged from 41.18 to 1,168.00 mg kg^{-1}, the average available phosphorus content ranged from 5.87 to 100.23 mg kg^{-1}, and the average available potassium ranged from 71.76 to 1,038.73 mg kg^{-1} (Table 1).

The study area spans multiple climate zones in China, including temperate, warm temperate, subtropical, and alpine. The annual mean temperature of each sampling site varied between 3.12 and 17.31°C, annual precipitation varied between 446.45 and 1,206.15 mm, and annual sunshine hours varied between 1,037.23 and 2,690.20 h (Table 1).

3.2 Purpurin and mollugin content of *R. cordifolia* roots in different locations

The contents of purpurin and mollugin at each sampling point were calculated according to the UHPLC chromatogram (Table 2). The chromatograms of standard and sample solutions are shown in Figure 3. The purpurin and mollugin contents in *R. cordifolia* roots varied greatly among sampling sites in the study area. Purpurin content ranged from 0.00 to 3.03 mg g^{-1}, the minimum and maximum values were found at sampling points in Sichuan. The content of mollugin varied from 0.03 to 10.09 mg g^{-1}. The minimum value was found in Sichuan sampling site, while the maximum value was found in Liaoning sampling site. The contents of the two secondary metabolites in *R. cordifolia* roots varied greatly in different provinces. For example, the content of purpurin in more than half of the samples in Gansu was less than 1.00 mg g^{-1}, and the

TABLE 2 The content of purpurin and mollugin in the roots of *R. cordifolia* at each sampling point.

Sampling site	Province	Longitude	Latitude	Purpurin content (mg g^{-1})	Mollugin content (mg g^{-1})
S1	Gansu	104.568	34.615	1.06	5.28
S2	Gansu	104.591	34.650	0.69	5.61
S3	Gansu	103.909	35.893	0.74	4.61
S4	Gansu	103.909	35.893	0.99	2.01
S5	Gansu	104.089	34.342	0.99	5.56
S6	Gansu	104.275	34.085	1.11	6.60
S7	Gansu	103.411	35.256	0.78	6.47
S8	Gansu	103.238	35.317	1.06	3.70
S9	Gansu	103.517	34.124	1.34	3.17
S10	Hebei	115.511	39.342	2.05	4.23
S11	Hebei	117.909	40.019	1.85	1.66
S12	Hebei	114.354	38.000	1.41	2.04
S13	Hebei	118.504	41.324	1.16	2.81
S14	Hebei	114.799	39.528	1.21	2.93

(Continued)

TABLE 2 Continued

Sampling site	Province	Longitude	Latitude	Purpurin content (mg g^{-1})	Mollugin content (mg g^{-1})
S15	Hebei	114.788	39.606	1.15	2.20
S16	Hebei	115.579	38.862	1.26	5.10
S17	Henan	112.601	34.112	1.65	5.67
S18	Henan	111.505	33.356	1.64	4.10
S19	Henan	112.674	33.729	0.87	3.57
S20	Henan	112.233	35.103	1.03	4.59
S21	Liaoning	122.490	40.569	1.34	6.16
S22	Liaoning	122.186	40.181	1.08	4.85
S23	Liaoning	121.972	39.873	1.50	10.09
S24	Liaoning	122.239	40.310	1.27	4.63
S25	Liaoning	123.589	41.908	1.48	2.56
S26	Liaoning	122.169	41.299	1.40	6.61
S27	Liaoning	119.391	41.221	1.05	2.93
S28	Liaoning	121.192	38.734	1.02	7.83
S29	Liaoning	120.770	41.199	1.14	2.89
S30	Liaoning	120.786	41.756	0.84	7.73
S31	Liaoning	120.786	41.756	0.60	7.59
S36	Shaanxi	109.705	36.695	0.51	6.43
S37	Shaanxi	109.979	33.865	1.68	8.45
S38	Shaanxi	107.919	34.163	1.21	8.68
S39	Shaanxi	109.374	34.114	1.80	3.04
S32	Shanxi	111.942	35.153	1.56	6.06
S33	Shanxi	111.456	37.438	0.74	4.49
S34	Shanxi	111.275	35.721	1.16	4.62
S35	Shanxi	111.881	36.896	0.77	6.03
S40	Sichuan	103.815	32.748	0.00	5.89
S41	Sichuan	104.497	32.463	1.80	0.03
S42	Sichuan	105.880	32.447	1.37	3.25
S43	Sichuan	103.616	30.989	2.87	0.37
S44	Sichuan	100.349	28.566	0.65	0.94
S45	Sichuan	106.660	31.870	3.03	0.56
S46	Sichuan	103.250	31.118	0.96	0.12

highest was 1.34 mg g^{-1}. The content of purpurin in in the samples in Hebei was higher than 1.00 mg g^{-1}, and the highest content was 2.05 mg g^{-1}. However, the content of mollugin in the samples in Hebei was less than 4.00 mg g^{-1}, and the highest content was 5.10 mg g^{-1}. The content of purpurin in the roots of *R. cordifolia* in nearly half of the sampling sites in Sichuan was less than 1.00 mg g^{-1}, and the highest was 3.03 mg g^{-1}. The content of mollugin in the roots in almost all sampling sites was not up to standard, indicating that the quality of *R. cordifolia* in each sampling site in Sichuan was uneven. The content of secondary metabolites in the roots of *R. cordifolia* in Shanxi, Shaanxi and Henan provinces can basically reach the standard. In addition, the content of secondary metabolites in the roots of *R. cordifolia* in a small number of sampling points in Liaoning was not up to standard, mainly distributed in the western part of Liaoning, while the quality of *R. cordifolia* in the southeastern part of Liaoning is higher (Figure 4).

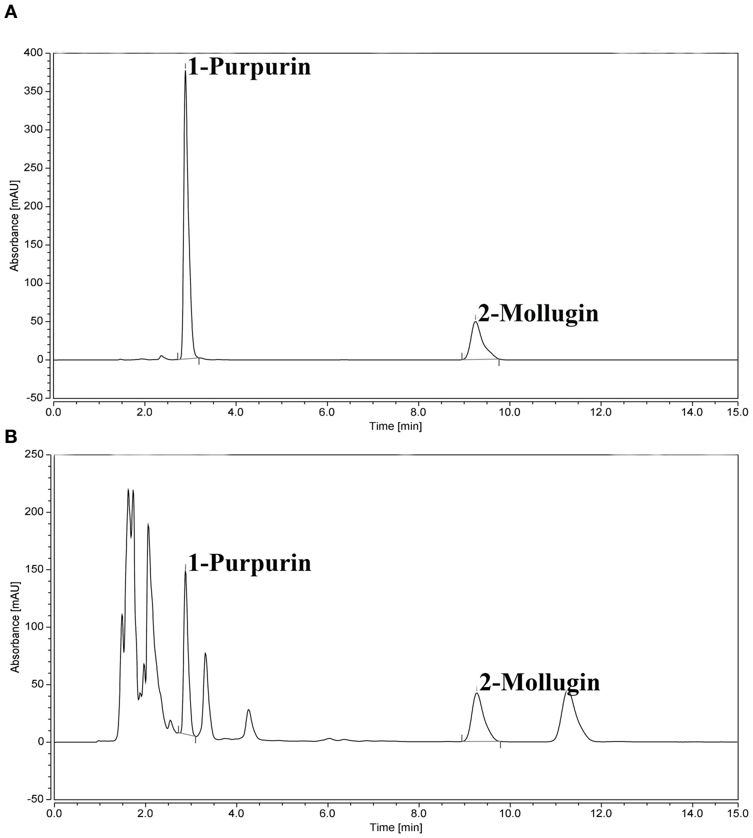


FIGURE 3
Chromatographic diagram of standard and sample solution. **(A)** Standard solution; **(B)** Sample solution.

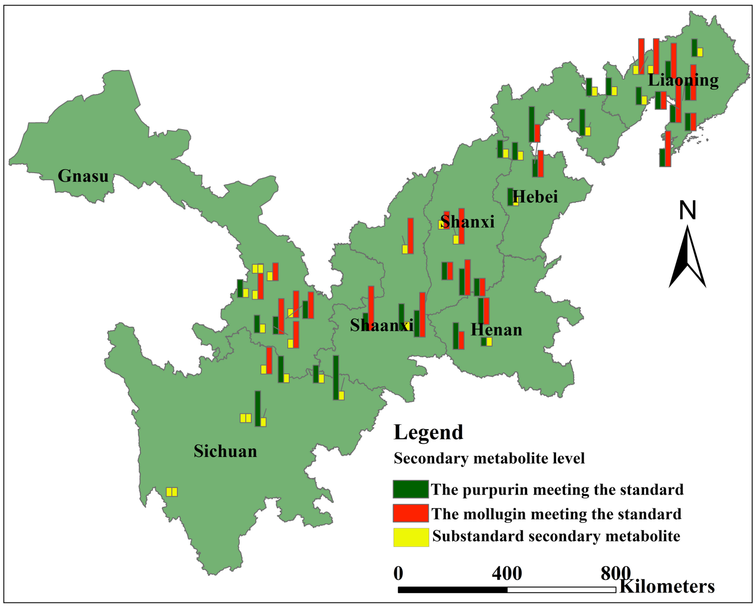


FIGURE 4
Quantitative comparison of metabolic compounds at each sampling site.

3.3 Correlation between purpurin, mollugin content and environmental variables

Purpurin and mollugin contents showed different relationships with soil, climate, and topographic factors (Figure 5). Longitude showed a significant positive correlation with purpurin content ($p < 0.05$) (Figure 5A), while elevation ($p < 0.001$) showed a significant negative correlation (Figure 5C). Annual mean temperature ($p < 0.001$) and annual precipitation ($p < 0.001$) showed highly significant positive correlations with purpurin content (Figures 5D, E), whereas annual sunshine hours showed a significant negative correlation with purpurin content ($p < 0.05$) (Figure 5F). Organic matter content ($p < 0.05$), available phosphorus content ($p < 0.05$), and available potassium content ($p < 0.01$) were significantly negatively correlated with purpurin content (Figures 5H, J, K). For mollugin, longitude ($p < 0.05$) and latitude ($p < 0.05$) were significantly positively correlated with mollugin content (Figures 5A, B). Organic matter content ($p < 0.05$) showed a significant and negative correlation with mollugin content ($p < 0.05$) (Figure 5H). There was no significant correlation between mollugin content and climatic factors (Figures 5D–F), however, there was a negative correlation between annual precipitation and mollugin content ($p = 0.052$) (Figure 5E).

3.4 Regression of purpurin, mollugin content with environmental variables

In order to further explore the relationship between the two secondary metabolites and environmental factors, a stepwise regression model between purpurin, mollugin and environmental factors was established (Table 3). Model results showed that purpurin and mollugin were affected by different factors. Purpurin in *R. cordifolia* roots was affected by latitude, annual mean temperature and annual precipitation, and the three factors had a positive impact on purpurin. Otherwise, the model results showed that root mollugin content was only negatively affected by annual precipitation.

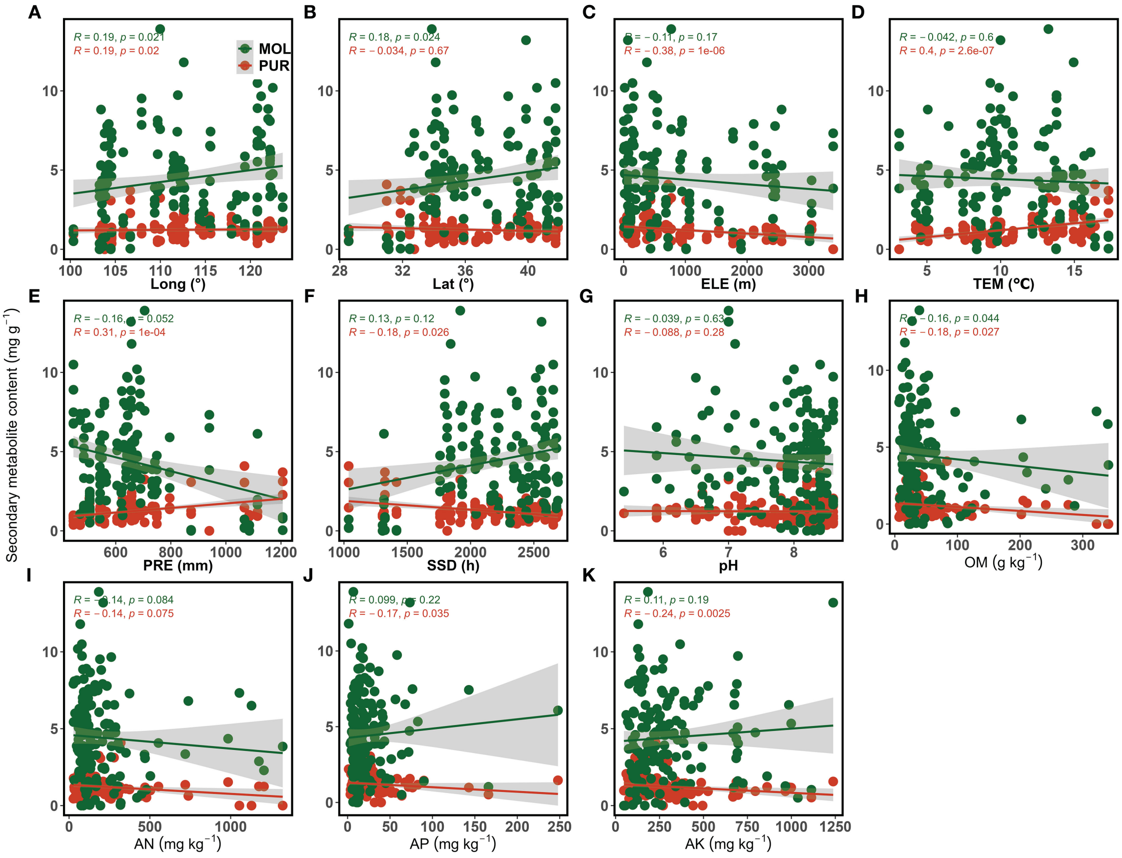


FIGURE 5
Correlation between purpurin, mollugin content, and environmental variables. **(A)** The relationship between secondary metabolite content and longitude; **(B)** The relationship between secondary metabolite content and latitude; **(C)** The relationship between secondary metabolite content and elevation; **(D)** The relationship between secondary metabolite content and annual mean temperature; **(E)** The relationship between secondary metabolite content and annual precipitation; **(F)** The relationship between secondary metabolite content and annual sunshine hours; **(G)** The relationship between secondary metabolite content and soil pH; **(H)** The relationship between secondary metabolite content and organic matter content; **(I)** The relationship between secondary metabolite content and available nitrogen content; **(J)** The relationship between secondary metabolite content and available phosphorus content; **(K)** The relationship between secondary metabolite content and available potassium content. PUR, purpurin; MOL, mollugin; Long, Longitude; Lat, Latitude; ELE, Elevation; TEM, Annual mean temperature; PRE, Annual precipitation; SSD, Annual sunshine hours; OM, Organic matter; AN, Available nitrogen; AP, Available phosphorus; AK, Available potassium.

TABLE 3 Stepwise regression model between secondary metabolites and environmental factors.

Variable	Estimate	Standard Error	Z value	*P* (>|z|)	R^2
Purpurin					
Intercept	-5.821	1.114	-5.224	< 0.001	0.281
Lat	0.091	0.024	3.796	< 0.001	
TEM	0.119	0.020	5.905	< 0.001	
PRE	0.002	0.001	3.888	< 0.001	
Mollugin					
Intercept	1.258	0.296	4.245	< 0.001	0.107
PRE	-0.002	0.000	-4.386	< 0.001	

Lat, Latitude; TEM, Annual mean temperature; PRE, Annual precipitation. Data about purpurin and mollugin were transformed using case ranking method before use in the regression model.

4 Discussion

4.1 The difference of purpurin and mollugin content in *R. cordifolia* roots in different locations

R. cordifolia is widely distributed in China; therefore, there are several places of origin recorded in the literature, including Shanxi, Sichuan, Henan, Shaanxi, Jiangsu, Hubei, Shandong, Anhui, and Hebei (Xue et al., 2009; Yan, 2021). Several studies have investigated the medicinal components in the roots of *R. cordifolia* from different origins and found that *R. cordifolia* from Shanxi, Shaanxi, Henan, Hebei, Jiangsu, Shandong, and Anhui met the Chinese Pharmacopoeia standards (Xue et al., 2009; Yu et al., 2017). Henan and Shaanxi are recognized *R. cordifolia* producing areas (Wen et al., 2022). The purpurin and mollugin contents of *R. cordifolia* roots from eastern and central regions in this study investigated met the requirements of the Chinese Pharmacopoeia, especially in central regions. Therefore, the study confirmed that *R. cordifolia* produced in Henan, Shaanxi, and Shanxi contained higher levels of purpurin and mollugin. However, *R. cordifolia* resources in Northeastern China have not yet been investigated. In this study, *R. cordifolia* from Liaoning contained higher purpurin and mollugin contents, which met the requirements of the Chinese Pharmacopoeia, suggesting that Liaoning could be a choice for *R. cordifolia* cultivation.

4.2 Effects of geographic, soil, and climate factors on the content of purpurin and mollugin

Environmental factors, including soil, climate, and topography, showed different degrees of correlation with the two secondary metabolites in the roots. The effects of geographic factors on plant secondary metabolite content have been extensively reported. The content of proanthocyanidins in *Ribes nigrum* L. was positively correlated with latitude (Yang et al., 2019b), whereas harpagoside content in *S. ningpoensis* was negatively correlated with latitude (Yang et al., 2011). Different essential oil contents in *H. italicum* showed different correlations with elevation (Melito et al., 2016). In this study, purpurin and mollugin contents were significantly and positively correlated with longitude and latitude (Figure 5; Table 2), suggesting that when the environment was closer to northern and eastern China, the purpurin and mollugin contents in *R. cordifolia* roots were higher, which is similar to the findings of Weston et al. (2013). The two secondary metabolites can cause oxidative stress in cells and therefore act as natural biocides (Robertson et al., 2009; Hook et al., 2018). The higher purpurin and mollugin content found in the eastern and central regions of this study might be due to the fact that the environmental conditions in these regions cause some kind of biotic stress, which in turn leads to the synthesis of more purpurin and mollugin. Previous studies have shown that anthraquinone content in medicinal plants *Rheum officinale* Baill. (Wang et al., 2013; Yan et al., 2017; Ge et al., 2018) and *Rumex nepalensis* Spreng (Farooq et al., 2013) increased with elevation, but in the present study elevation negatively affected purpurin content in *R. cordifolia* roots, and a similar result was found in *Rheum australe* D. Don (Pandith et al., 2014). The negative correlation between secondary metabolite content and elevation may be related to factors such as ultraviolet radiation or temperature (Spitaler et al., 2006). The positive correlation between purpurin content and annual mean temperature and the negative correlation between elevation and temperature just proved this point.

Soil nutrients play a crucial role in supporting plant growth. In the case of medicinal plants, the presence of appropriate soil nutrients not only enhances their yield but also improves their overall quality. Changes in soil nutrition will inevitably lead to changes in secondary metabolites, understanding the relationship between the content of secondary metabolites and soil nutrients can guide the cultivation of medicinal plants to improve the effectiveness of medicinal plants. However, specific soil factors that cause an increase in secondary metabolite levels in medicinal plants are not the same. Soil pH may be an important factor determining the secondary metabolite content of *Echinacea purpurea* (L.) Moench (Xu et al., 2022). The available phosphorus in the soil is thought to be a key factor influencing the secondary metabolite content in *Salvia miltiorrhiza* Bunge (Liang et al., 2021). For *Rhodiola sachalinensis* A. Bor, available phosphorus and available potassium are the main factors limiting the synthesis of salidroside (Yan et al., 2004). In this study, the findings found that higher soil nutrients were not necessarily beneficial to the synthesis of purpurin and mollugin, which is similar to the study by Shen et al. (2017), where the content of anthraquinone compounds in the roots of *Rheum tanguticum* Maxim. ex Balf was higher in nutrient-poor soil. This phenomenon might be due to the fact that under nutrient-sufficient conditions *R. cordifolia* mainly carries out primary metabolism for growth, while under poor soil conditions the plants are stressed and then carry out secondary metabolism and produce secondary metabolites to alleviate the stress.

The climate is crucial for the growth and quality of medicinal plants. Studies have shown that different medicinal plants have different requirements for climatic conditions, and secondary metabolite production can be affected by climatic conditions (Liu et al., 2020). Temperature has multiple effects on the synthesis of secondary metabolites in medicinal plants. On the one hand, temperature stress may cause medicinal plants to produce secondary metabolites for defense (Dong et al., 2011); on the other hand, high or low temperatures may increase the expression of genes involved in the synthesis of secondary metabolites (Tao et al., 2017; Yang et al., 2019b). The harpagoside content in *Scrophularia ningpoensis* Hemsl (Yang et al., 2011), and the nerolidol content in *Helichrysum italicum* (Roth) G. Don (Melito et al., 2016) are correlated positively with temperature, while there is a negative correlation between high temperature and total proanthocyanidins in *R. nigrum* fruits (Yang et al., 2019b). Temperature promoted the synthesis of purpurin in the roots of *R. cordifolia*, which may be due to high temperature stress or increased expression of related genes. Precipitation also has an impact on the synthesis of secondary metabolites in plants. Typically, drought conditions stimulate the synthesis and accumulation of secondary metabolites in plants (Humbal and Pathak, 2023). However, chlorogenic acid, choline, 3,5-o-dicaffeoylquinic acid, and coumaric acid in *E. purpurea* are positively correlated with annual precipitation (Xu et al., 2022). Two secondary metabolites in the roots of *R. cordifolia* were oppositely affected by precipitation, which promoted the synthesis of purpurin but limited the synthesis of mollugin. The impact of precipitation on the secondary metabolism of medicinal plants is complex. This complexity arises due to the potential changes in soil properties, microbial fungi, and root activity that can occur as a result of precipitation (Wang et al., 2021b; Xu et al., 2022). Sunshine hours also affect the production of plant secondary metabolites. For instance, a longer duration of sunshine can increase the content of geniposidic acid in *Eucommia ulmoides* Oliv. (Dong et al., 2011) and tanshinones in *S. miltiorrhiza* (Zhang et al., 2018). However, in the present study, correlation analysis suggested that sunshine duration might have an inhibitory effect on purpurin synthesis in *R. cordifolia* root.

4.3 Suggestions on *R. cordifolia* planting

The cultivation of medicinal plants needs to ensure both yield and quality, so Daodi medicinal materials are recognized as high-quality Chinese medicinal products. However, as wild medicinal plants have been over-exploited, they can no longer meet market demand. Therefore, field cultivation has become the main way to ensure the market supply of medicinal plants (Kling, 2016). In the case of field planting, many factors need to be considered comprehensively, including geographical location, soil, climate, etc. First of all, it is necessary to select the appropriate planting area, which needs to consider geographical and climatic factors. Therefore, eastern and northern China can be used as a choice when planting *R. cordifolia*. Higher temperatures facilitate the synthesis of purpurin, so the temperature in the suitable area needs to be considered. Precipitation has opposite effects on the two secondary metabolites in the roots of *R. cordifolia*. When planting *R. cordifolia*, the precipitation situation needs to be weighed, and the amount of watering in the field must also be considered. In order to ensure the yield of medicinal plants, fertilization is a necessary measure. However, the synthesis of two major secondary metabolites in *R. cordifolia* roots is inhibited by soil nutrients, therefore necessary fertilization experiments need to be performed to determine the appropriate fertilization ratio.

4.4 Future research prospects

The two main secondary metabolites in the root of *R. cordifolia* belong to quinone compounds, among which purpurin belongs to anthraquinone compounds (Singh et al., 2021), and mollugin belongs to naphthoquinone compounds (Morita et al., 2015). The Chinese Pharmacopoeia requires that the contents of the two secondary metabolites in the roots of *R. cordifolia* should be up to standard before they can be used as medicines (Chinese Pharmacopeia Commission, 2020), so the products of *R. cordifolia* on the market are uneven. How to improve the quality when planting *R. cordifolia* has become a problem. At present, some researchers have carried out transcriptome studies on *R. cordifolia*, and found that the synthesis of anthraquinones in the roots involves the synthesis of A, B, and C rings. Among them, A and B rings are completed by the shikimic acid pathway, while the synthesis of C ring comes from isopentenyl diphosphate (Peng et al., 2018). Previous studies have confirmed that isopentenyl diphosphate is synthesized through the 2-C-methyl-D-erythritol 4-phosphate pathway in the *Rubiaceae*. Isochorismate synthase gene, o-succinylbenzoic acid gene, 1-deoxy-D-xylulose 5-phosphate reductoisomerase gene, and 1-deoxy-D-xylulose 5-phosphate synthase gene are all key genes for the synthesis of anthraquinones (Han et al., 2001). However, the synthetic pathway of naphthoquinone compounds in the roots of *R. cordifolia* has not been reported. Therefore, in the future, on the one hand, the biosynthesis pathway of anthraquinones will be explored by molecular biology methods, and key genes will be screened. On the other hand, the relationship between the key genes in the synthesis pathway of anthraquinones and naphthoquinones and the ecological environment factors should be analyzed to find the key factors, and then measures can be taken to promote the synthesis of two secondary metabolites.

5 Conclusions

This study found that the quality of *R. cordifolia* varied considerably depending on the region; however, in general, the quality of *R. cordifolia* from southeastern Liaoning and the border of Shanxi, Shaanxi and Henan met the Chinese Pharmacopoeia standards, whereas Liaoning proved to be a potential source of *R. cordifolia* in northeastern China. Environmental factors have different effects on the two secondary metabolites. The purpurin and mollugin contents of *R. cordifolia* roots were significantly and positively correlated with longitude and latitude, while were negatively correlated with soil nutrients. Temperature contributed to the synthesis of purpurin, but sunshine duration inhibited the synthesis. Precipitation is a key factor limiting the quality of *R. cordifolia* because it has opposite effects on the synthesis of purpurin and mollugin. To meet the requirements of the Chinese Pharmacopoeia, the limiting factors affecting the two secondary metabolites should be considered when planting *R. cordifolia* in field. In the future, it is necessary to combine the key genes in the synthesis pathway of the two secondary metabolites, further analyze the key factors affecting the synthesis of the two compounds, and then promote the synthesis of the two secondary metabolites by controlling the environmental factors to improve the quality of *R. cordifolia*.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

YW: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. HL: Formal analysis, Resources, Writing – review & editing. SY: Data curation, Writing – review & editing. YZ: Resources, Writing – review & editing. YH: Resources, Writing – review & editing. XH: Conceptualization, Writing – review & editing. WC: Conceptualization, Investigation, Methodology, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Key R&D Program of China (No. 2022YFF1300500) and the Youth Innovation Promotion Association CAS (2022195).

Conflict of interest

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

Publisher's note

References

Chinese Pharmacopeia Commission (2020). *Pharmacopoeia of the People's Republic of China: Pharmacopoeia I* (Beijing, China: Chinese Medical Science Press).

Dong, J. E., Ma, X. H., Wei, Q., Peng, S. B., and Zhang, S. C. (2011). Effects of growing location on the contents of secondary metabolites in the leaves of four selected superior clones of *Eucommia ulmoides*. *Ind. Crops Prod.* 34, 1607–1614. doi: 10.1016/j.indcrop.2011.06.007

Du, H. Z., Hou, X. Y., Miao, Y. H., Huang, B. S., and Liu, D. H. (2020). Traditional Chinese Medicine: an effective treatment for 2019 novel coronavirus pneumonia (NCP). *Chin. J. Nat. Med.* 18, 206–210. doi: 10.1016/S1875-5364(20)30022-4

Farooq, U., Pandith, S. A., Singh Saggoo, M. I., and Lattoo, S. K. (2013). Altitudinal variability in anthraquinone constituents from novel cytotypes of *Rumex Nepalensis* Spreng—a high value medicinal herb of North Western Himalayas. *Ind. Crops Prod.* 50, 112–117. doi: 10.1016/j.indcrop.2013.06.044

Ge, Y. H., Sun, M. M., Salome-Abarca, L. F., Wang, M., and Choi, Y. H. (2018). Investigation of species and environmental effects on rhubarb roots metabolome using H-1 NMR combined with high performance thin layer chromatography. *Metabolomics* 14, 137. doi: 10.1007/s11306-018-1421-1

Han, Y. S., van der Heijden, R., and Verpoorte, R. (2001). Biosynthesis of anthraquinones in cell cultures of the *Rubiaceae*. *Plant Cell Tissue Organ Cult.* 67, 201–220. doi: 10.1023/A:1012758922713

Hook, I., Sheridan, H., and Reid, C. (2018). Trichomes and naphthoquinones protect *Streptocarpus dunnii* Hook.f. against environmental stresses. *South Afr. J. Bot.* 119, 193–202. doi: 10.1016/j.sajb.2018.09.016

Hu, H. R., and Wang, Y. X. (2020). *Experimental tutorial of soil science. 2nd Edition* (Beijing, China: China Forestry Publishing House).

Humbal, A., and Pathak, B. (2023). Influence of exogenous elicitors on the production of secondary metabolite in plants: A review ("VSI: secondary metabolites"). *Plant Stress.* 8, 100166. doi: 10.1016/j.stress.2023.100166

Kim, H., Choi, H. K., Jeong, T. C., Jahng, Y., Kim, D. H., Lee, S. H., et al. (2013). Selective inhibitory effects of mollugin on CYP1A2 in human liver microsomes. *Food Chem. Toxicol.* 51, 33–37. doi: 10.1016/j.fct.2012.09.013

Kling, J. (2016). Protecting medicine's wild pharmacy. *Nat. Plants.* 2, 16064. doi: 10.1038/nplants.2016.64

Li, J., Zhang, J. L., Gong, X. P., Xiao, M., Song, Y. Y., Pi, H. F., et al. (2020). Anti-inflammatory activity of mollugin on DSS-induced colitis in mice. *Curr. Med. Sci.* 40, 910–916. doi: 10.1007/s11596-020-2262-5

Liang, H. Y., Kong, Y. H., Chen, W., Wang, X. G., Jia, Z. F., Dai, Y. H., et al. (2021). The quality of wild *Salvia miltiorrhiza* from Dao Di area in China and its correlation with soil parameters and climate factors. *Phytochem. Anal.* 32, 318–325. doi: 10.1002/pca.2978

Liu, X. D., Zhang, Y., Wu, M. H., Ma, Z. G., Huang, Z. H., Tian, F., et al. (2020). The scientific elucidation of daodi medicinal materials. *Chin. Med.* 15, 86. doi: 10.1186/s13020-020-00367-1

Loreto, F., Dicke, M., Schnitzler, J. P., and Turlings, T. C. J. (2014). Plant volatiles and the environment. *Plant Cell Environ.* 37, 1905–1908. doi: 10.1111/pce.12369

Lyu, M., Fan, G. W., Xiao, G. X., Wang, T. Y., Xu, D., Gao, J., et al. (2021). Traditional chinese medicine in COVID-19. *Acta Pharm. Sin. B.* 11, 3337–3363. doi: 10.1016/j.apsb.2021.09.008

Ma, Y.g., Zhao, Y. J., Ma, X. J., Guo, J., and Huang, L. Q. (2022). Biosynthetic pathway of active components in Traditional Chinese Medicine and its application. *Scientia Sin. Vitae.* 52, 894–907. doi: 10.1360/SSV-2021-0401

Melito, S., Petretto, G. L., Podani, J., Foddai, M., Maldini, M., Chessa, M., et al. (2016). Altitude and climate influence *Helichrysum italicum* subsp microphyllum essential oils composition. *Ind. Crops Prod.* 80, 242–250. doi: 10.1016/j.indcrop.2015.11.014

Morita, H., Nishino, H., Nakajima, Y., Kakubari, Y., Nakata, A., Deguchi, J., et al. (2015). Oxomollugin, a potential inhibitor of lipopolysaccharide-induced nitric oxide production including nuclear factor kappa B signals. *J. Nat. Med.* 69, 608–611. doi: 10.1007/s11418-015-0927-3

Pandey, A., Agrawal, M., and Agrawal, S. B. (2023). Individual and combined effects of chromium and ultraviolet-B radiation on defense system, ultrastructural changes, and production of secondary metabolite psoralen in a medicinal plant *Psoralea corylifolia* L. *Environ. Sci. pollut. Res.* 30, 4372–4385. doi: 10.1007/s11356-022-22480-4

Pandith, S. A., Hussain, A., Bhat, W. W., Dhar, N., Qazi, A. K., Rana, S., et al. (2014). Evaluation of anthraquinones from Himalayan rhubarb (*Rheum emodi* Wall. ex Meissn.) as antiproliferative agents. *South Afr. J. Bot.* 95, 1–8. doi: 10.1016/j.sajb.2014.07.012

Peng, L., Zhang, G., Yan, Y. G., Sun, T., Wang, M., Li, J., et al. (2018). Analysis of transcriptomes and exploring of anthraquinones biosynthetic pathway genes in *Rubia cordifolia* L. *Chin. Tradition. Herb. Drugs* 49, 1890–1896. doi: 10.7501/j.issn.0253–2670.2018.08.024

Qin, K. M., Cai, H., Li, W. D., Lu, T. L., Jin, J. J., Liu, X., et al. (2018). Research on construction of quality control system for high quality Chinese herbal medicine and industrialization application. *Modern. Tradition. Chin. Med. Materia Medica-World Sci. Technol.* 20, 383–389.

Robertson, P. K. J., Black, K. D., Adams, M., Willis, K., Buchan, F., Orr, H., et al. (2009). A new generation of biocides for control of crustacea in fish farms. *J. Photochem. Photobiol. B: Biol.* 95, 58–63. doi: 10.1016/j.jphotobiol.2008.12.009

Selmar, D., and Kleinwächter, M. (2013). Influencing the product quality by deliberately applying drought stress during the cultivation of medicinal plants. *Ind. Crops Prod.* 42, 558–566. doi: 10.1016/j.indcrop.2012.06.020

Shan, M. Q., Yu, S., Yan, H., Chen, P. D., Zhang, L., and Ding, A. W. (2016). A review of the botany, phytochemistry, pharmacology and toxicology of *Rubiae Radix* et Rhizoma. *Molecules* 21, 1747. doi: 10.3390/molecules21121747

Shen, N., Cui, Y. L., Xu, W. H., Zhao, X. H., and Yang, L. C. (2017). Impact of phosphorus and potassium fertilizers on growth and anthraquinone content in *Rheum tanguticum* Maxim. *ex Balf. Ind. Crops Prod.* 107, 312–319. doi: 10.1016/j.indcrop.2017.05.044

Singh, J., Hussain, Y., Luqman, S., and Meena, A. (2021). Purpurin: A natural anthraquinone with multifaceted pharmacological activities. *Phytother. Res.* 35, 2418–2428. doi: 10.1002/ptr.6965

Spitaler, R., Schlorhaufer, P. D., Ellmerer, E. P., Merfort, I., Bortenschlager, S., Stuppner, H., et al. (2006). Altitudinal variation of secondary metabolite profiles in flowering heads of *Arnica montana* cv. *ARBO. Phytochem.* 67, 409–417. doi: 10.1016/j.phytochem.2005.11.018

Tao, X., Wang, M. X., Dai, Y., Wang, Y., Fan, Y. F., Mao, P., et al. (2017). Identification and expression profile of CYPome in perennial ryegrass and tall fescue in response to temperature stress. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01519

Tu, Y. Y. (1999). The development of new antimalarial drugs: Qinghaosu and dihydro-qinghaosu. *Chin. Med. J.* 112, 976–977.

Wang, W. Y., Zhou, H., Wang, Y. F., Sang, B. S., and Liu, L. (2021a). Current policies and measures on the development of Traditional Chinese Medicine in China. *Pharmacol. Res.* 163, 105187. doi: 10.1016/j.phrs.2020.105187

Wang, Y. T., Xie, Y. Z., Rapson, G., Ma, H. B., Jing, L., Zhang, Y., et al. (2021b). Increased precipitation enhances soil respiration in a semi-arid grassland on the Loess Plateau, China. *PeerJ* 9, e10729. doi: 10.7717/peerj.10729

Wang, Z., Li, M. Y., Mi, C. L., Wang, K. S., Ma, J., and Jin, X. J. (2017). Mollugin has an anti-cancer therapeutic effect by inhibiting TNF-alpha-Induced NF-kappa B activation. *Int. J. Mol. Sci.* 18, 1619. doi: 10.3390/ijms18081619

Wang, Z., Ma, P., Xu, L. J., He, C. N., Peng, Y., and Xiao, P. G. (2013). Evaluation of the content variation of anthraquinone glycosides in rhubarb by UPLC-PDA. *Chem. Cent. J.* 7, 170. doi: 10.1186/1752–153X-7–170

Wen, M., Chen, Q., Chen, W., Yang, J., Zhou, X. G., Zhang, C. X., et al. (2022). A comprehensive review of *Rubia cordifolia* L.: Traditional uses, phytochemistry, pharmacological activities, and clinical applications. *Front. Pharmacol.* 13. doi: 10.3389/fphar.2022.965390

Weston, P. A., Weston, L. A., and Hildebrand, S. (2013). Metabolic profiling in *Echium plantagineum*: presence of bioactive pyrrolizidine alkaloids and napthoquinones from accessions across southeastern Australia. *Phytochem. Rev.* 12, 831–837. doi: 10.1007/s11101-013-9306-4

Xu, X. L. (2017). "Spatial interpolation dataset of average condition of meteorological elements in China," (System, Resource and Environmental Science Data Registration and Publishing System, Beijing, China).

Xu, W. Q., Cheng, Y. L., Guo, Y. H., Yao, W. R., and Qian, H. (2022). Effects of geographical location and environmental factors on metabolite content and immune activity of *EChinacea purpurea* in China based on metabolomics analysis. *Ind. Crops Prod.* 189, 115782. doi: 10.1016/j.indcrop.2022.115782

Xue, L., Chen, S. Z., Suo, F. Y., and Yang, B. (2009). HPLC determination of rubimaillin in Chinese herbal pieces of *Radix* et *Rhizoma Rubiae*. *Chin. J. Pharm. Anal.* 29, 363–366.

Xue, R. C., Fang, Z., Zhang, M. X., Yi, Z. H., Wen, C. P., and Shi, T. L. (2012). TCMID: Traditional Chinese Medicine integrative database for herb molecular mechanism analysis. *Nucleic Acids Res.* 41, D1089–D1095. doi: 10.1093/nar/gks1100

Yan, J. F. (2021). *Study on market investigation and quality evaluation of* Rubia cordifolia (Beijing: Beijing University of Chinese Medicine).

Yan, X. F., Wu, S. X., Wang, Y., Shang, X. H., and Dai, S. J. (2004). Soil nutrient factors related to salidroside production of *Rhodiola sachalinensis* distributed in Chang Bai Mountain. *Environ. Exp. Bot.* 52, 267–276. doi: 10.1016/j.envexpbot.2004.02.005

Yan, Y. G., Wang, Y. H., Deng, C., Zhang, G., Chen, Y., Shen, X., et al. (2017). Effects of growth years, altitude, and light factors on contents of eight components in Rheum officinale. *Chin. Tradition. Herb. Drugs* 48, 2285–2291.

Yang, L. L., Zhao, Y. J., Zhang, Q., Cheng, L., Han, M., Ren, Y. Y., et al. (2019a). Effects of drought–re-watering–drought on the photosynthesis physiology and secondary metabolite production of *Bupleurum chinense* DC. *Plant Cell Rep.* 38, 1181–1197. doi: 10.1007/s00299-019-02436-8

Yang, S. T., Li, J. H., Zhao, Y. P., Chen, B. L., and Fu, C. X. (2011). Harpagoside variation is positively correlated with temperature in *Scrophularia ningpoensis* Hemsl. *J. Agric. Food Chem.* 59, 1612–1621. doi: 10.1021/jf104702u

Yang, W., Ma, X. Y., Laaksonen, O., He, W. J., Kallio, H., and Yang, B. R. (2019b). Effects of latitude and weather conditions on proanthocyanidins in blackcurrant *(Ribes nigrum*) of Finnish commercial cultivars. *J. Agric. Food Chem.* 67, 14038–14047. doi: 10.1021/acs.jafc.9b06031

Yu, R., Gao, M. J., Cui, B. B., Zhao, Y. S., Jin, M. Y., Meng, X. Y., et al. (2017). Determination of alizarin, purpurin and rubimaillin in *Rubia cordifolia* L. by HPLC-DAD. *J. Harbin Med. University.* 51, 195–199.

Zhan, P., Wang, F. Y., Xia, P. G., Zhao, G. H., Wei, M. T., Wei, F. G., et al. (2022). Assessment of suitable cultivation region for *Panax notoginseng* under different climatic conditions using MaxEnt model and high-performance liquid chromatography in China. *Ind. Crops Prod.* 176, 114416. doi: 10.1016/j.indcrop.2021.114416

Zhang, X. D., Yu, Y. G., Yang, D. F., Qi, Z. C., Liu, R. Z., Deng, F. T., et al. (2018). Chemotaxonomic variation in secondary metabolites contents and their correlation between environmental factors in *Salvia miltiorrhiza* Bunge from natural habitat of China. *Ind. Crops Prod.* 113, 335–347. doi: 10.1016/j.indcrop.2018.01.043

Zhao, Z. Y., Li, Y. D., Zhou, L. Y., Zhou, X. T., Xie, B. W., Zhang, W. J., et al. (2021). Prevention and treatment of COVID-19 using Traditional Chinese Medicine: A review. *Phytomedicine* 85, 153308. doi: 10.1016/j.phymed.2020.153308

Zheng, T., Sun, J. Q., Shi, X. J., Liu, D. L., Sun, B. Y., Deng, Y. J., et al. (2022). Evaluation of climate factors affecting the quality of red huajiao (*Zanthoxylum bungeanum* maxim.) based on UPLC-MS/MS and MaxEnt model. *Food Chem.: X.* 16, 100522. doi: 10.1016/j.fochx.2022.100522

Frontiers in Plant Science
TYPE Original Research
PUBLISHED 08 July 2024
DOI 10.3389/fpls.2024.1420231

EDITED BY
Boyang Ji,
BioInnovation Institute (BII), Denmark

REVIEWED BY
Jacinta Collado González,
Instituto Murciano de Investigación y Desarrollo Agrario y Alimentario (IMIDA), Spain
Xiujun Zhang,
Chinese Academy of Sciences (CAS), China
Arif Rashid,
Jiangsu University, China

*CORRESPONDENCE
Huachen Wang
✉ wanghc@gdou.edu.cn
Jie Chen
✉ 812242515@qq.com

†These authors have contributed equally to this work

RECEIVED 19 April 2024
ACCEPTED 14 June 2024
PUBLISHED 08 July 2024

CITATION
Yang C, Chen J, Zhang Y, Huang J, Wang H and Chen J (2024) Non-targeted metabolomics reveals the taste variations during *Baccaurea ramiflora* Lour. fruit maturation.
Front. Plant Sci. 15:1420231.
doi: 10.3389/fpls.2024.1420231

Non-targeted metabolomics reveals the taste variations during *Baccaurea ramiflora* Lour. fruit maturation

Chongcheng Yang[1†], Jiaqi Chen[1†], Yang Zhang[2], Jianjian Huang[3], Huachen Wang[1*] and Jie Chen[1*]

[1]College of Coastal Agricultural Sciences, Guangdong Ocean University, Zhanjiang, China, [2]Organic Biology Group, Jiangxi Ganzhou Eco-environmental Monitoring Center, Ganzhou, China, [3]School of Food Engineering and Biotechnology, Hanshan Normal University, Chaozhou, China

Baccaurea ramiflora Lour. is a new kind of underutilized wild fruit tree; the metabolic reasons for its fruit flavor changes are not yet clear. In this study, the pink flesh of this excellent tasting fruit (BR) was used to reveal the metabolic causes of taste variations through five developmental stages. We identified 154 common differential metabolites of different developmental stages based on non-targeted metabolomics analysis. The accumulation of sugar and fatty acids increased significantly after 73 days, while citric acid decreased significantly. Flesh color accumulation mainly occurred 53 days ago, and vitamin accumulation occurred after 93 days. Interestingly, L-sorbose and 5-hydroxyindole-3-acetic acid were positively correlated with the sugar–acid ratio but negatively correlated with titratable acids. It indicated that L-sorbose and 5-hydroxyindole-3-acetic acid may be taste biomarkers of BR *B. ramiflora*. The results provided new metabolic lines of evidence for the taste variation during the ripening process of *B. ramiflora*.

KEYWORDS
***Baccaurea ramiflora*, fruit flavor, taste biomarker, metabolites profiling, non-targeted metabonomics**

1 Introduction

Baccaurea ramiflora Lour. (Phyllanthaceae family) is a wild fruit tree native to tropical and subtropical regions, widely distributed in South China (Hainan, Lianjiang Guangdong and South Guangxi) and Southern Yunnan, as well as in South and Southeast Asian countries such as India, Nepal, Myanmar, the Andaman Islands, Thailand, and Peninsular Malaysia (Ferrer et al., 2008; Huang et al., 2021). *B. ramiflora* is a new type of underutilized wild fruit tree with good edible quality, which is suitable for fresh consumption or processed into fruit juice, dried fruit, jam, and wine (Nesa et al., 2018). Generally, *B. ramiflora* has a high economic value potential in which the yield of an adult plant reaches 100–150 kg (Pandey et al., 2018). However, *B. ramiflora* has not yet been commercially planted and utilized, although local residents usually collect these fruits from the forests and sell them in the local market (Banerjee et al., 2022).

B. ramiflora has a highly edible flesh with a rich dietary fiber; it approximately contains 35.6% water, 51.9% carbohydrate, 5.58% protein, and 20.4% fiber (Ramjan and Raghavan, 2018). The fruit is rich in magnesium (504 mg/100 g), potassium (730 mg/100 g), phosphorus (132 mg/100 g), and iron (100 mg/100 g) (Ramjan and Raghavan, 2018). In addition, *B. ramiflora* has good health and medicinal—anti-inflammatory, antioxidant (Usha et al., 2014), hypoglycemic, and hypolipidemic (Ullah et al., 2012)—properties. Despite being a wild fruit, *B. ramiflora* has gradually attracted widespread attention because of its unique advantages, namely, good taste, high nutrition value, and therapeutic applications (Huang et al., 2024). Although the various local strains of *B. ramiflora* have not been well developed and utilized, BR (white pericarp and pink flesh) in Fangchenggang, Guangxi is a more popular fruit with excellent taste.

The fruit taste of *B. ramiflora* is similar to jujuba (Guo et al., 2015), apple (Ma et al., 2015), and wampee (Yin et al., 2022), ranging from sour, sweet, to astringent mainly depending on the growth stage. Primary metabolites, particularly carbohydrates and organic acids, are closely related to the sweet and sour taste of the fruit (Jiang et al., 2023). Metabolomics was increasingly used to study fruit flavors (Liu et al., 2024; Su et al., 2024). Deng et al. studied the key substances determining the flavor of rambutan fruits using widely targeted metabolomics analysis and examined the variations of the major metabolites, such as sugars, organic acids, and amino acids, throughout the maturation process of rambutan fruits (Deng et al., 2023). Chen et al. studied the key substances determining the flavor of the *B. ramiflora* fruit by non-targeted metabolomics and explored the quantities and main types of major metabolites such as sugar, organic acid, and amino acid (Chen et al., 2023). It is very important to understand the metabolic pathway and key metabolites of fruit ripening in different stages to improve the quality of *B. ramiflora*. However, the important taste-related metabolites and the underlying mechanisms, which change during maturation, are not known and have not been characterized. Limited data constrain the systematic analysis of primary and secondary metabolites that affect fruit taste.

In this study, to better understand the flavor changes of *B. ramiflora* pulp, various metabolites, such as carbohydrates, organic acids, amino acids, fatty acids, and flavonoids at different mature stages of *B. ramiflora*, were analyzed by liquid chromatography with tandem mass spectrometry (LC-MS/MS; a fast and reliable method for the detection of various plant metabolites) (Yue et al., 2018). Through the non-targeted metabolomics analysis of five typical growth stages of the BR *B. ramiflora* fruit, all metabolites were screened out. Finally, pathway enrichment analysis identified its characteristic metabolites and potential flavor biomarkers and explained their possible mechanisms.

2 Materials and methods

2.1 Plant materials

BR (white pericarp and pink flesh) *B. ramiflora* fruits were collected from healthy fruit trees that were more than 10 years old and grown in the vicinity of Nasuo middle school in Fangchenggang, Guangxi, China (N 21°42′33″, E 108°6′29″, alt 20 m). The samples were collected at 30, 52, 73, 93, and 112 days after the flowering stage, respectively (Figure 1).

2.2 Sample preparation and metabolite extraction

We selected three trees with uniform growth and collected six fruits with the same maturity from each tree in five growing stages. Four pulps from the same period were randomly selected and mixed into one biological sample, and there were three biological replicates at each developmental stage. The samples were immediately stored in liquid nitrogen, brought back to the laboratory, and stored at −80°C, until the metabolites were extracted. Fifty milligrams of the sample was placed in a 1.5-mL centrifuge tube and added to 800 μL of extraction solution (water:methanol = 7:3, V:V, pre-cooled to −20°C), and 20 μL of the internal standard was added sequentially, ground in a tissue grinder (50 Hz, 5 min). The mixture was subjected to an ultrasonic water bath (4°C, 30 min), refrigerated (−20°C, 1 h), and centrifuged at 14,000 rpm for 15 min at 4°C using a low-temperature high-speed centrifuge (Centrifuge 5430, Eppendorf). The supernatant (600 μL) was taken and passed through a 0.22-μm filter membrane, and the filtrate sample was collected and placed in a sample bottle for ultrahigh-performance LC-MS/MS (UPLC-MS/MS) analysis (Chen et al., 2023).

2.3 UPLC-MS/MS analysis

Non-targeted metabonomics analysis identified the prepared extracts using a UPLC-HRMS/HRMS system (UPLC, Waters 2D UPLC, USA; HRMS, Thermo Fisher Scientific, USA), and data were collected in both positive and negative ion modes to improve the metabolite coverage (Di Guida et al., 2016). Five microliters of sample solutions was injected into a Hypersil GOLD aQ chromatographic column (100 mm * 2.1 mm, 1.9 μm, Thermo Fisher Scientific, USA). The UPLC mobile phase was composed of 0.1% formic acid aqueous solution (denoted as solvent A) and acetonitrile containing 0.1% formic acid (denoted as solvent B). The following gradient was used for elution: 0 min, 5% B; 2 min, 95% B; 22 min, 95% B; 27 min, 5% B solution. The overall process flow maintained a rate of 0.3 mL/min at a temperature of 40°C. The Q Exactive mass spectrometer was used for collecting primary and secondary mass spectrometry data. Parameters included the scanned mass-to-charge ratios (*m/z*) within the 150–1,500 range, a primary resolution of 70,000, an AGC of 1×10^6, and a maximum injection time (MIT) set to 100 ms. The top three *m/z* peaks were selected based on the precursor ion intensity for fragmentation

FIGURE 1
Five fruit profiles of the maturation process of BR *B. ramiflora*. From left to right, BR1 (30 days), BR2 (52 days), BR3 (73 days), BR4 (93 days), and BR5 (112 days) after full flowering, respectively. Scale bar = 2 cm.

and secondary information collection. The secondary resolution was 35,000, with an AGC of 2×10^5 and an MIT of 50 ms; the fragmentation energy (stepped NCE) was set to 20, 40, and 60 eV. The sheath gas and Aux gas flow rates of the ion source (ESI) were set to 40 mL/min and 10 mL/min, respectively. Spray voltage (|kV|) was set to 3.80 in the positive ion mode and 3.20 in the negative ion mode. The ion transfer tube temperature (capillary temperature) and the auxiliary gas heater temperature (aux gas heater temperature) were 320°C and 350°C, respectively.

2.4 Quality control measurement

The quality control (QC) sample was prepared by mixing equal volumes (20 µL) of each sample from five developmental stages, and the QC sample was measured using UPLC-MS/MS with the same parameter settings as the rest of the samples. The repeatability and stability of the UPLC-MS/MS analysis process were evaluated through the number of peaks, peak response intensity, chromatogram overlap, and principal component analysis (PCA) of QC.

2.5 MS data analysis

Data preprocessing was performed by importing the raw data collected from LC-MS/MS into Compound discoverer 3.1 (Thermo Fisher Scientific, USA). Data processing includes peak extraction, background peak labeling, intra-group and inter-group retention time correction, missing value filling, adduct ion merging, and metabolite identification, and then data on compound molecular weight, peak area, retention time, and identification results were exported in the end. The identification of metabolites was comprehensively referenced from multiple databases including BERRY Library, mzCloud, and ChemSpider [the Human Metabolome Database (HMDB), the Lipid Metabolites and Pathways Strategy, LipidMaps, and Kyoto Encyclopedia of Genes and Genomes (KEGG)].

The results obtained from Compound discoverer 3.1 were imported into metaX for data preprocessing, relative peak area was obtained using Probabilistic Quotient Normalization (PQN) (Di Guida et al., 2016), batch effect correction of the actual sample was performed via local polynomial regression fitting signal correction by QC-based robust LOESS signal correction (QC-RLSC) (Dunn et al., 2011) based on QC sample information, and the compounds of QC samples with a coefficient of variation (CV) greater than 30% were deleted.

2.6 MS statistical analysis

Statistical analyses of metabolites including PCA, hierarchical clustering analysis (HCA), fold change (FC), variable importance projection (VIP), metabolite classification, and functional annotations were performed using the metabonomic R software package metaX and the metabonomic information analysis process developed by Berry Hekang Gene Company (Wen et al., 2017). The Partial Least Squares Method Discriminant Analysis (PLS-DA) model was used to calculate the VIP of two principal components, which can measure the intensity of impact and explanatory power of different metabolite expression patterns on the classification discrimination of each sample group and assist in screening metabolic markers (Li et al., 2015). Log_2 logarithmic conversion on data was performed, and the PLS-DA model between the comparative analysis group (two groups of samples) was established, with the scaling method of Par. Sevenfold cross-validation of the established PLS-DA model was performed, and 200 times response permutation testing was carried out to judge the model quality. Differential metabolites (DMs) were screened based on the VIP values of the first two principal components obtained from the PLS-DA model, combined with the results of FC and Student's *t*-test obtained from univariate analysis. PCA and FC were converted by Log_2, and the screening criteria were *p*-value <0.05, VIP ≥1, and FC ≥1.2 or ≤0.83. The increase or decrease of DMs was based on the five growing stages of BR *B. ramiflora*.

2.7 KEGG pathway enrichment analysis of metabolites

KEGG is a well-known and reliable database that explains molecular-level details of chemicals in organisms (Kanehisa et al., 2012) and was used to interpret the identified BR *B. ramiflora* metabolites. Annotated metabolites were mapped to the KEGG pathway database, and key pathways were identified by pathway enrichment analysis based on enrichment factors and the number of metabolites.

2.8 Statistical analysis

SPSS (22.0, IBM Corp., Armonk, NY, USA) software and OriginLab (2019, OriginLab Inc., Northampton, MA, USA) software were used for data statistical analysis and graphing, and expressed as the mean ± standard deviation (SD). Data were evaluated by one-way analysis of variance (ANOVA) using Tukey's honestly significant difference (HSD) test ($p < 0.05$) (Feng et al., 2024).

3 Results and analysis

3.1 Statistical analysis of BR pulp metabolites

PCA showed a clear separation between samples of five developmental stages and QC, with significant differences ($p < 0.05$) as depicted in Figure 2. Among 18 samples, BR1 prominently separated from the rest, BR2 and BR3 were separated and proximal, while BR4 and BR5 were almost indistinguishable. The peak area of each metabolite was transformed by Log_2, and subsequent HCA was performed to eliminate the influence of quantity on pattern recognition. HCA revealed five distinct groups related to BR1, BR2, BR3, BR4, and BR5, respectively (Figure 3). Overall, non-targeted metabolite analysis revealed the unique metabolite profiles of the five developmental stages of BR pulp, indicating the distinctive metabolic characteristics of the *B. ramiflora* fruit in different maturation stages.

3.2 Identification and classification of metabolites

A total of 536 metabolites in BR *B. ramiflora* were identified by non-targeted metabolite analysis utilizing LC-MS/MS, of which 361 were identified in positive ion mode and 201 were identified in negative ion mode, including 26 duplicates. The main substances affecting fruit flavor were 73 primary metabolites (12 carbohydrates, 3 organic acids, 7 amino acids and 8 derivatives, 2 vitamins, and 41 fatty acids) and 166 secondary metabolites (42 flavonoids, 8 phenols, 6 phenolic acids, 26 phenylpropanoids, 4 steroids and 5 derivatives, and 75 terpenoids).

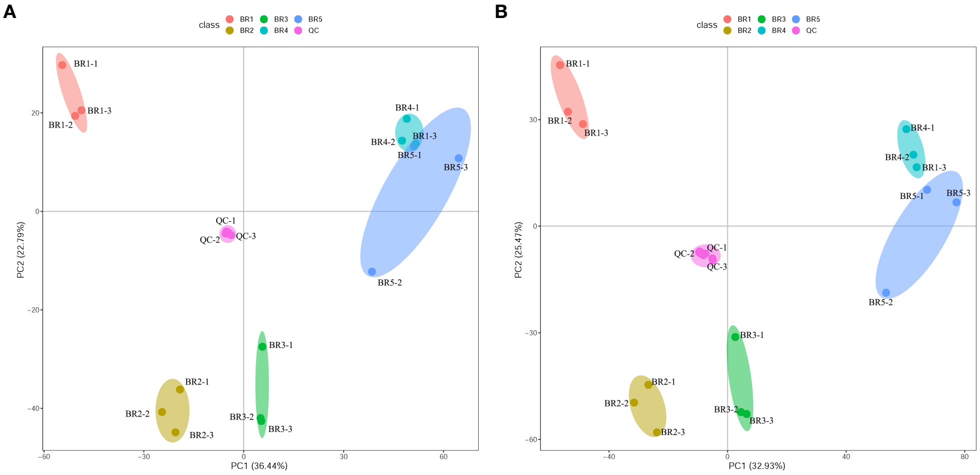


FIGURE 2
PCA of metabolites in the five developmental stages of BR *B. ramiflora*. QC represents a mixed equal amount of BR pulp samples. **(A)** Positive ion mode; **(B)** negative ion mode.

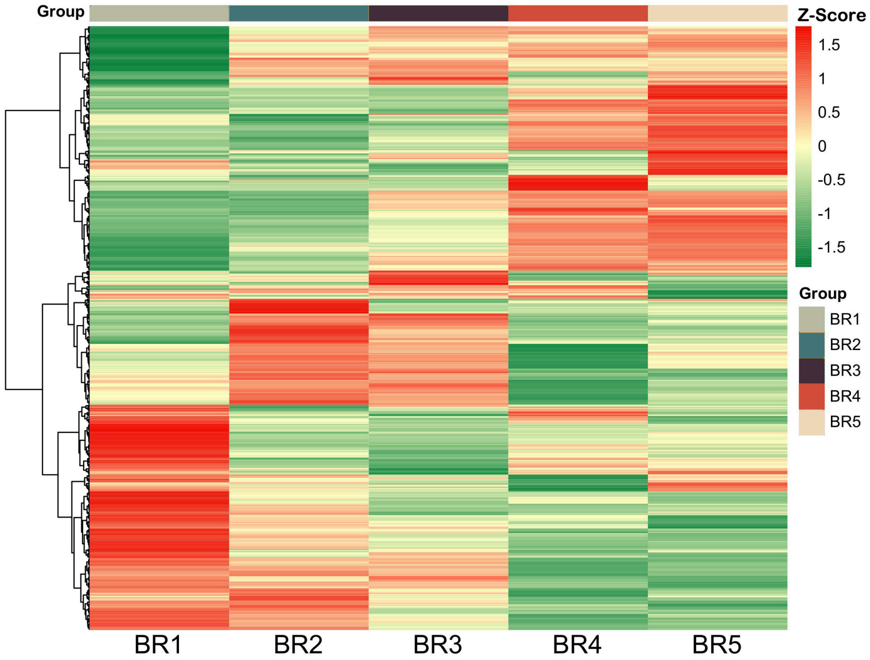


FIGURE 3
Hierarchical cluster analysis of metabolites identified from BR *B. ramiflora*. The color from green (low) to red (high) indicates the level of each metabolite. The *Z*-score represents the deviation from the mean by standard deviation units.

3.3 Analysis of the dynamic changes in sugars and organic acids

Fruit taste mainly includes sweet, sour, and astringent, which is an important factor that consumers consider when buying fruits. These tastes, primarily determined by soluble sugars and organic acids, are intrinsically linked to fruit composition (Feng et al., 2024). Generally, sugar in fruits accumulates progressively from the unripe stage to full ripeness (Al-Maiman and Ahmad, 2002). BR fruits contain 12 kinds of carbohydrates, with the majority being L-sorbose, D-(+)-glucose, bis(methylbenzylidene) sorbitol, and sucrose (Supplementary Figure 1). Fructose was not detected, which may due to its low content. L-sorbose, which was the most abundant carbohydrate, showed a significant non-accumulation in the first three stages (30, 52, and 73 days), followed by a significant increase at 93 days (p = 0.033, FC = 2.327, VIP = 1.301), and reached the peak at 112 days. D-(+)-glucose was the second most abundant carbohydrate, and the accumulation trend was similar to L-sorbose. The contents of L-sorbose and D-(+)-glucose increased significantly at 93 days, indicating that the BR fruit entered the mature stage at 73 days and began to accumulate sugar rapidly. The content of bis(methylbenzylidene) sorbitol remained relatively steady, suggesting its negligible impact on the fruit's texture or taste. The sucrose content was the smallest, showing an upward trend similar to L-sorbose and D-(+)-glucose in the first four developmental stages, further confirming that BR began to mature at 73 days. In addition, compared with BR4, the sucrose content of BR5 was decreased but was not significant, suggesting that sucrose was converted into glucose during the maturation period from 93 days to 112 days.

Citric acid was the main organic acid in the pulp of BR *B. ramiflora*, and malic and oxalic acids were undetectable. Citric acid initially increased and then decreased, which was consistent with the taste change of BR *B. ramiflora* pulp from astringent to sour in the early stage and from sour to sweet in the later stage. Citric acid decreased significantly in BR3 to BR4 (p = 0.007, FC = −2.974, VIP = 1.708), suggesting that BR began to mature at 73 days, which was consistent with the soluble sugars. Zhen (2015) considered that although the organic acid content could not directly determine the sweetness of fruits, the sugar–acid ratio significantly affects the flavor of fruits. As shown in Table 1, the soluble sugar [sucrose, L-sorbose, and D-(+)-glucose] content continuously increased from 61.67% in BR1 to 73.23% in BR5. Organic acids, mainly citric acid, increased slightly from BR1 to BR3, then significantly decreased from 21.897% in BR3 to 18.88% in BR4, and finally showing no significant increase in BR4 and BR5. With the change in the content of soluble sugar and organic acid, the sugar–acid ratio continually increased from 2.91 in BR1 to 3.87 in BR4, subsequently sliding to 3.78 in BR5. These results indicated that the change of sugar–acid ratio affected the taste of the BR *B. ramiflora* fruit, and the main reason was the increase of soluble sugar.

3.4 Analysis of the dynamic changes in amino acids

Amino acids, sugars, and organic acids are the primary metabolic products of fruits that determine fruit quality (Zhang et al., 2010). The composition and content of amino acids are important indexes to evaluate nutritional quality and taste. In this study, L-phenylalanine

TABLE 1 Sugars, acids, and basic taste qualities of BR pulp at five mature stages.

Growing stage	Sugars				Acids		Sugar–acid ratio
	Soluble sugar (%)	Sucrose	L-sorbose	D-(+)-glucose	Titratable acid (%)	Citric acid	
BR1	61.66 ± 0.12c	18.65 ± 0.51c	22.58 ± 0.18c	20.42 ± 0.49d	21.17 ± 0.25a	21.17 ± 0.25a	2.91 ± 0.04b
BR2	63.90 ± 1.52bc	20.28 ± 0.77bc	22.56 ± 0.29c	21.06 ± 0.56cd	21.80 ± 0.14a	21.80 ± 0.14a	2.93 ± 0.09b
BR3	65.70 ± 1.93b	21.48 ± 0.57ab	22.61 ± 0.95c	21.61 ± 0.48c	21.90 ± 0.03a	21.90 ± 0.03a	3.00 ± 0.09b
BR4	72.94 ± 1.20a	24.15 ± 0.70a	25.08 ± 0.36b	23.71 ± 0.15b	18.88 ± 0.44b	18.88 ± 0.44b	3.86 ± 0.05a
BR5	73.23 ± 3.70a	22.31 ± 3.02ab	26.22 ± 0.62a	24.69 ± 0.33a	19.38 ± 1.31b	19.38 ± 1.31b	3.78 ± 0.18a

Sugar–acid ratio is soluble sugar divided by titratable acid. Different letters on the number meant significant differences between growing stages ($p < 0.05$).

was the main amino acid affecting flavor, followed by L-tyrosine and DL-arginine in lower proportions (Supplementary Figure 2). L-phenylalanine decreased slightly in the first four developmental stages and increased significantly only in the last developmental stage, which may improve the taste of fruit; DL-arginine did not change significantly in the first three development stages, and then increased significantly, which indicated that DL-arginine had a certain effect on pulp quality. There was no significant difference in L-tyrosine among the five developmental stages, but it increased at 112 days. The contents of all three amino acids increased at the final developmental stage, indicating that it may be a reliable indicator of maturation. The results showed that the accumulation of amino acids increased during the mature stage, which played an important role in the fruit taste.

3.5 Analysis of the dynamic changes in fatty acids

Fatty acids are essential components that contribute to fruit quality (Zishun et al., 2016). A total of 38 fatty acids were identified in the BR pulp, mainly oleamide, A-eleostearic acid, and corchorifatty acid F, and their content variations are shown in Supplementary Figure 3. The oleamide content remained stable in the first three growth stages and gradually increased from 73 days to 112 days. A-eleostearic acid did not show significant fluctuations in five developmental stages. Corchorifatty acid F showed a downward trend, with significant reductions in BR2 vs. BR1 ($p = 0.0007$, FC = −2.887, VIP = 1.347) and BR5 vs. BR4 ($p = 0.1293$, FC = 1.407, VIP = 1.7769). The increase of oleamide in the late stage, as well as the significant decrease of corchorifatty acid F, played an important role in improving the taste of BR *B. ramiflora*.

3.6 Analysis of the dynamic changes in flavonoids

Flesh color is usually determined by carotenoids, anthocyanins, and flavonoids, which are also pigments of various plants (Ferrer et al., 2008). The synthesis of anthocyanin is related to flavonoids and phenylpropanoids, which are precursors of anthocyanin synthesis. Flavonoids provide protection for plants under diverse stress conditions and provide beneficial health effects in humans (Jaakola, 2013; Zhang et al., 2014). The biosynthetic pathway of flavonoids plays a pivotal role in the production and regulation of anthocyanins, proanthocyanidins, and flavonols (Wang et al., 2019).

A total of 42 flavonoids were identified in BR flesh, with the absence of carotenoids. Rhusflavanone, procyanidin B1, and catechin constituted the primary flavonoids. Rhusflavanone, the most abundant one in content, progressively increased during the whole fruit maturation. Rhusflavanone underwent a significant increase across different developmental stages, BR3 vs. BR2 ($p = 0.0002$, FC = 1.926, VIP = 1.234) and BR4 vs. BR3 ($p = 0.0004$, FC = 2.778, VIP = 1.230), indicating that rhusflavanone was evident from 73 days to 112 days and culminated in elevated quantities in the mature stage. Procyanidin B1 showed a decreasing trend, suggesting that the reduction of bitterness in the pulp was related to its gradual reduction. Catechin, a preliminary substance for procyanidin B1 synthesis, displayed a decreasing trend, which corresponded to the reduced synthesis of procyanidin B1. The gradual decrease in procyanidin B1 and catechins may be related to the reduction of anthocyanidin synthesis or upstream substrate synthesis.

Ten metabolites related to the biosynthetic processes of flavonoids, including anthocyanins, were found: naringenin chalcone, naringenin, eriodictyol, dihydroquercetin, kaempferol, quercetin, (+)-catechin hydrate, (−)-catechin gallate, procyanidin B1, and rhusflavanone. Ten metabolites showed an overall downward trend during development (Supplementary Figure 4), with the accumulation of the contents of each substance mainly in the first two periods, indicating that anthocyanins were mainly synthesized in the earlier period. The content of dihydroquercetin, an essential precursor for anthocyanin synthesis, was significantly decreased in BR5 vs. BR4 ($p = 0.020$, FC = −1.338, VIP = 1.209), speculating that the increase of its content was an important reason for the pink flesh of the BR fruit. Kaempferol experienced a significant decrease in BR5 vs. BR4 ($p = 0.021$, FC = −1.339, VIP = 1.256), indicating a diminished synthesis of anthocyanins. Thus, anthocyanin synthesis gradually decreased in the late stage.

3.7 Analysis of the dynamic changes in polyphenolics

Polyphenolics can affect the color and biological characteristics of fruits, which is an important index to evaluate fruit quality.

Generally, unripe fruits are unpalatable due to the presence of polyphenolics, resulting in a bitter and astringent taste. The reduction of polyphenolic content during fruit ripening will lessen these bad flavors and lead to a better fruit flavor (Aldhanhani et al., 2022). Therefore, the maturity and harvest time of fruits can be judged by the content of polyphenolics (Cuthbertson et al., 2012; Prakash et al., 2020). Seven high-content polyphenolic metabolites were identified in the BR pulp, 2-hydroxycinnamic acid, caffeic acid, phloretin, hexylcinnamaldehyde, dihydromethysticin, demethoxyyangonin, and apocynin. Polyphenolic content gradually decreased as the fruit ripened (Supplementary Figure 5). A significant decrease appeared at the stage from BR4 to BR3, and a higher polyphenolic content from BR1 to BR3 than from BR4 and BR5 indicated that the synthesis and accumulation of polyphenolic substances mainly happen in the early stages of BR *B. ramiflora* development.

3.8 Analysis of the dynamic changes in vitamins

Two vitamin metabolites, pantothenic acid and DL-thioctic acid, were identified (Supplementary Figure 6). Pantothenic acid is a precursor of coenzyme A synthesis and widely distributed in plants (Serrano-García et al., 2022). The pantothenic acid content remained relatively stable in 30, 52, 73, and 93 days, and remarkably increased in the last stage ($p = 1.8453$, FC = 1.043, VIP = 1.557). In DL-thioctic acid, there was no significant difference in 30, 52, and 73 days, but there was a significant decrease in BR4 vs. BR3 ($p = 2.3727$, FC = −3.359, VIP = 1.644). The results suggested that pantothenic acid and DL-thioctic acid may affect the taste of the BR pulp in the later maturation stages.

3.9 Differential metabolites at five different maturity stages

To identify DMs in five growing stages of BR *B. ramiflora*, we selected metabolites based on p-value < 0.05, VIP ≥1, $\text{Log}_2\text{FC} \geq 1.2$, and $\text{Log}_2\text{FC} \leq 0.83$. In total, 247 DMs were identified, which could be divided into 15 classes (Figure 4A). The metabolites of 10 groups were compared and analyzed, and major DMs were mainly fatty acyls (FA), terpenoids, flavonoids, phenylpropanoids, and others (Figure 4B).

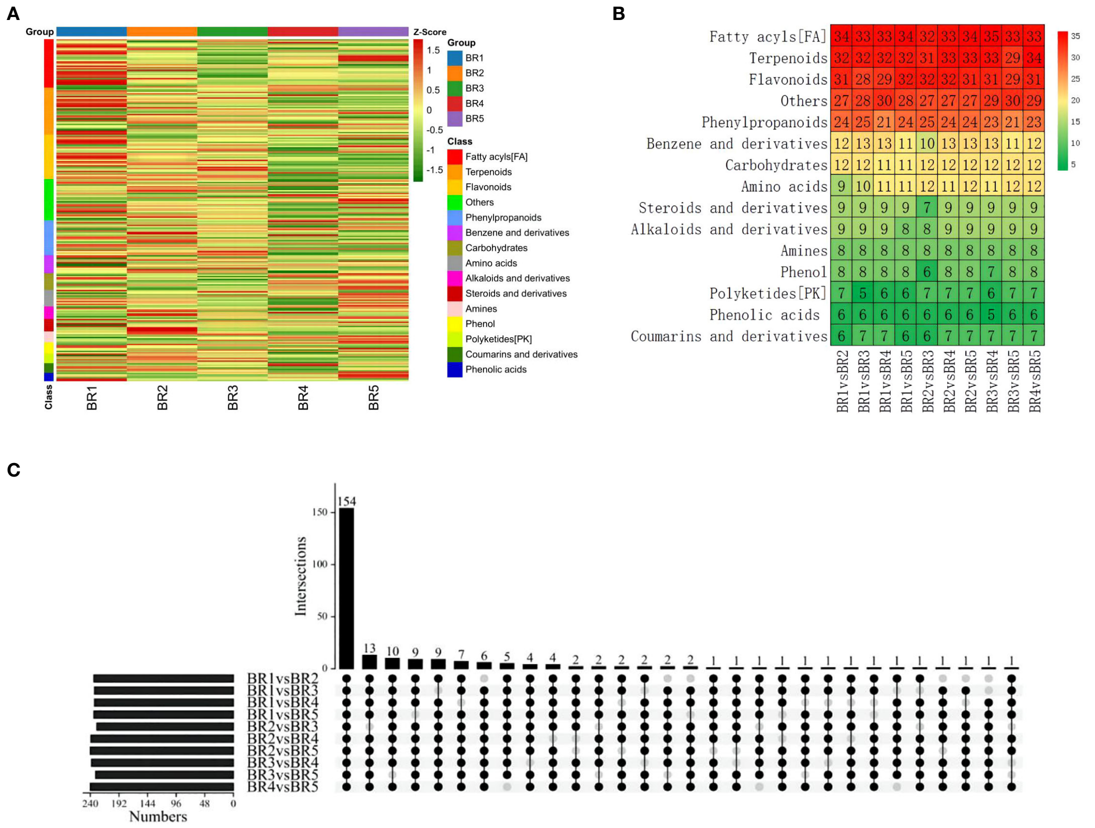
A total of 154 common DMs were identified from 10 groups in five stages by the UpSet plot analysis (Figure 4C), namely, 21 FAs, 21 flavonoids, 20 terpenoids, 14 phenylpropanoids, 10 carbohydrates, 8 amines, 7 alkaloids and derivatives, 7 benzene and derivatives, 7 steroids and derivatives, 6 amino acids, 5 phenols, 4 phenolic acids, 4 coumarins and derivatives, 3 polyketides, and 17 others. The comparison analysis of metabolites at different maturity stages showed 234, 233, 233, and 234 DMs in BR1 vs. BR2, BR3, BR4, BR5; 229, 239, and 240 DMs in BR2 vs. BR3, BR4, and BR5; 238 and 231 DMs in BR3 vs. BR4 and BR5; and 240 DMs in BR4 vs. BR5.

Flavonoids and terpenoids affect fruit coloration, and FAs played an important role in fruit development and affected the structure and function of plant cell membrane (Zheng et al., 2022). The significant variation of flavonoids, FAs, and terpenoids may explain fruit coloring and the morphological alterations in fruit cross-sections and provide important support for fruit development. Sugars, organic acids, and amino acids were the major compounds contributing to sensory variation. Five carbohydrates exhibited a continuous increase, namely, D-(+)-glucose, lactose, L-sorbose, lusitanicoside, and uridine 5’-diphosphogalactose, indicating that carbohydrate synthesis was active during the maturation process of BR *B. ramiflora*. Additionally, five amino acids were actively synthesized (zeatin-7-n-glucoside biocytin, L-alanyl-l-proline, DL-arginine, L-phenylalanine, and L-tyrosine) and one amino acid was degraded (methionine sulfoxide) during fruit ripening.

3.10 KEGG enrichment analysis of DMs

The comprehensive analysis of metabolites in five mature stages of BR *B. ramiflora* was performed using the KEGG database, and the main metabolic pathways were elucidated (Supplementary Figure 7). There were 37 metabolic pathways, mainly sugar and acid and amino acid metabolism. Sugar and acid metabolism included glycolysis/gluconeogenesis, fructose and mannose metabolism, pyruvate metabolism, and citrate cycle, and the metabolites involved in the glycometabolism pathway were mainly L-sorbose, citric acid, D-(−)-salicin, and 2-isopropylmalic acid. Twelve pathways were established in amino acid metabolism, namely, 2-oxocarboxylic acid metabolism; alanine, aspartate, and glutamate metabolism; beta-alanine metabolism; biosynthesis of amino acids; cyanoamino acid metabolism; glycine, serine, and threonine metabolism; phenylalanine metabolism; phenylalanine, tyrosine, and tryptophan biosynthesis; tryptophan metabolism; tyrosine metabolism; valine, leucine, and isoleucine biosynthesis; and aminoacyl-tRNA biosynthesis. Twelve metabolic pathways of amino acid synthesis, such as amino acid biosynthesis and L-phenylalanine, tyrosine, and tryptophan biosynthesis, share common metabolic pathways. A total of nine metabolites were identified, and L-tyrosine and L-tryptophan were the most active metabolites involved in six amino acid metabolic pathways. KEGGs found that DMs were mainly enriched in plant hormone signal transduction, linoleic acid metabolism, oxidative phosphorylation, phenylalanine, tyrosine and tryptophan biosynthesis, and citrate cycle. Metabolite analysis showed that many metabolites were involved in the synthesis of amino acids, which may significantly improve the nutritional value and taste of the BR fruit during ripening.

It provided a comprehensive analysis of the KEGG pathway based on carbohydrate metabolism and amino acid biosynthesis pathway, and explained the mechanism of flavor change of the BR fruit (Figure 5). L-sorbose, a significant sugar in BR, was active in the fructose and mannose metabolism pathway with a significant accumulation from BR1 to BR5, showing that L-sorbose played a



	BR1vsBR2	BR1vsBR3	BR1vsBR4	BR1vsBR5	BR2vsBR3	BR2vsBR4	BR2vsBR5	BR3vsBR4	BR3vsBR5	BR4vsBR5
Fatty acyls[FA]	34	33	33	34	32	33	34	35	33	33
Terpenoids	32	32	32	32	31	33	33	33	29	34
Flavonoids	31	28	29	32	32	32	31	31	29	31
Others	27	28	30	28	27	27	27	29	30	29
Phenylpropanoids	24	25	21	24	25	24	24	23	21	23
Benzene and derivatives	12	13	13	11	10	13	13	13	11	12
Carbohydrates	12	12	11	11	12	12	12	12	12	12
Amino acids	9	10	11	11	12	11	12	11	12	12
Steroids and derivatives	9	9	9	9	7	9	9	9	9	9
Alkaloids and derivatives	9	9	9	8	8	9	9	9	9	9
Amines	8	8	8	8	8	8	8	8	8	8
Phenol	8	8	8	8	6	8	8	7	8	8
Polyketides[PK]	7	5	6	6	7	7	7	6	7	7
Phenolic acids	6	6	6	6	6	6	6	5	6	6
Coumarins and derivatives	6	7	7	6	6	7	7	7	7	7

FIGURE 4

Identification and classification of metabolic discrepancy. **(A)** Hierarchical cluster analysis of metabolites identified from BR *B. ramiflora*. The color from green (low) to red (high) indicates the level of each metabolite. The *Z*-score represents the deviation from the mean by standard deviation units. **(B)** Number of differential metabolites. **(C)** A total of 154 common differential accumulated metabolites by the UpSet plot.

key role in the ripening and flavor change of the BR fruit. Conversely, D-(−)-salicin, involved in the glycolysis pathway, continuously decreases from BR1 to BR5, suggesting an ongoing breakdown during the whole maturation of the BR fruit. 2-Isopropylmalic acid entered the valine, leucine, and isoleucine biosynthesis pathway via phosphoenolpyruvate, and undergoes a notable change throughout the maturation, showing a trend of first increasing, then decreasing, and finally increasing, representing high synthesis and decomposition activity. Two dynamic amino acids, L-tryptophan and L-tyrosine, were simultaneously found in the amino acid biosynthesis pathway. L-tyrosine remained consistent throughout the five maturation stages, while L-tryptophan showed a significant decrease from BR1 to BR4 and a significant increase from BR4 to BR5, suggesting that L-tryptophan continuously decomposes and synthesizes during BR fruit maturation. 5-Hydroxyindole-3-acetic acid, a metabolite found in the tryptophan metabolism pathway, continuously and significantly decreased from BR1 to BR5, suggesting that 5-hydroxyindole-3-acetic acid potentially played a major role in the flavor changes of BR fruit ripening. 2-Hydroxycinnamic acid was identified, entering the phenylalanine metabolism pathway via L-tyrosine, with a significant reduction in content from BR1 to BR4 and an increase from BR4 to BR5. Salidroside was found to enter the tryptophan metabolism pathway via L-tyrosine, decreasing significantly from BR1 to BR2 and decreasing continuously from BR2 to BR5. The changes in salidroside and 2-hydroxyquinamic acid indicated their dynamic role in synthesis and decomposition during the ripening process of BR fruits. In addition, citric acid and pantothenic acid, which respectively entered the citrate cycle pathway through acetyl-CoA and the beta-alanine metabolism pathway, were involved in the synthesis and decomposition of metabolites during the ripening of BR *B. ramiflora*.

3.11 Potential taste biomarker

Of the 154 common metabolites, only L-sorbose consistently increased with increasing BR maturity, resulting in a persistent increase in the sugar–acid ratio, presumably suggesting that L-sorbose may be a taste biomarker for BR fruits. Interestingly, KEGG enrichment analysis found that 5-hydroxyindole-3-acetic acid in the tryptophan metabolism pathway consistently decreased with increasing BR maturity. Further analysis showed that L-sorbose was positively correlated with sugar–acid ratio (R^2 = 0.9126) and negatively correlated with titratable acid (R^2 = 0.8382), and 5-hydroxyindole-3-acetic acid was positively correlated with sugar–acid ratio (R^2 = 0.8897) and negatively correlated with titratable acid (R^2 = 0.745) (Figure 6).

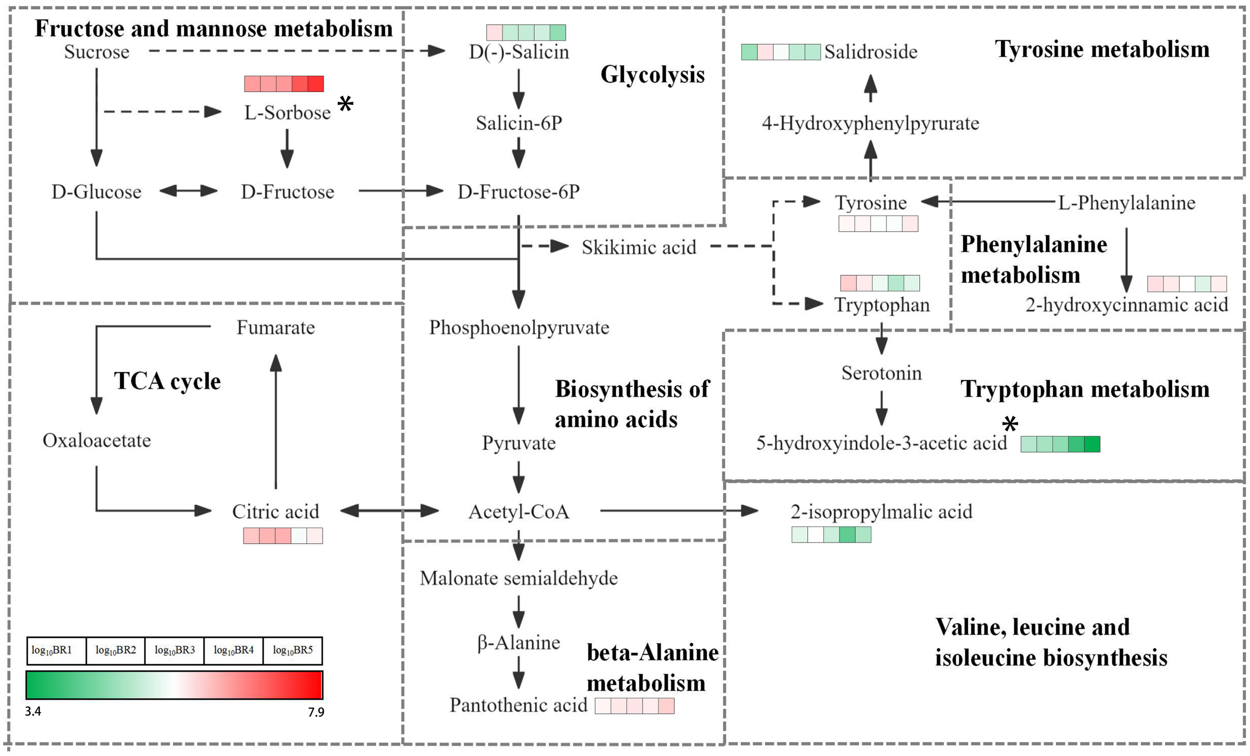


FIGURE 5
KEGG map of key DMs in BR *B. ramiflora*. This map was constructed based on the KEGG pathway and literary references. Colored boxes in front of each metabolite indicate $\log_{10}$BR1, $\log_{10}$BR2, $\log_{10}$BR3, $\log_{10}$BR4, and $\log_{10}$BR5 values according to the color scale. * represents the taste biomarker.

Carbohydrate is the main energy source for plant growth and development, and L-sorbose was the main carbohydrate of the BR fruit (Dhungana and Braun, 2021). 5-Hydroxyindole-3-acetic acid is a derivative of tryptophan and is a precursor to many natural products, such as pigments, alkaloids, hormones, and cell wall components in plants (Maeda and Dudareva, 2012). In conclusion, these results suggested that L-sorbose and 5-hydroxyindole-3-acetic acid may be key taste biomarkers of the BR fruit.

4 Discussion

B. ramiflora, mostly found in tropical forests with wild distribution, has not yet been commercially planted and utilized (Ramjan and Raghavan, 2018). There were significant differences between the flowering and fruiting stages from south to north. Generally, *B. ramiflora* bloom and bear fruit approximately 45 days earlier in Southeast Asian regions than those in mainland China (Huang et al., 2024). The fruits ripen between late May and early August, similar to litchi and longan. It has the potential to alleviate fruit shortages of this season, given the different breeding and fruiting periods of different provinces (Chen et al., 2023). Understanding the phenological period of *B. ramiflora* is not only valuable to their harvest time, but also beneficial to their commercial cultivation (Fyfe et al., 2023). Through the non-targeted metabolomics analysis of five typical growth stages of BR *B. ramiflora*, we have identified key metabolic changes that characterize each phase of development. The results showed that L-sorbose and D-(+)-glucose increased significantly at 93 days, indicating that the BR fruit entered the mature stage at 73 days and began to accumulate sugar rapidly. Citric acid displayed a trend of initial increase and then decrease, which was consistent with the taste change of the BR *B. ramiflora* pulp from astringency to sour in the early stage and from sour to sweet in the later stage. By analyzing the contents of sugars, organic acids, amino acids, fatty acids, flavonoids, polyphenolics, and vitamins in different fruit stages, we found that BR3 to BR4 may be the key period of respiratory climacteric. This is consistent with previous studies; maturity III (98 days after flowering) has good quality and high storage resistance, which can better meet the fresh marketing and storage of the *B. ramiflora* fruit after harvest (Kong et al., 2024).

Harvest maturity is an important factor affecting fruit quality and storability (Khalil et al., 2023). Understanding the mechanism of fruit ripening is critical for fruit storage and quality improvement (Zhang et al., 2021; Wang et al., 2022). In this study, non-target metabolome analysis was used to clarify the metabolic change of the BR *B. ramiflora* fruit (peel white, pulp pink) during five ripening stages and to reveal the key substances of taste change. A total of 536 metabolites were identified. The main primary metabolites were L-sorbose, D-(+)-glucose, citric acid, L-phenylalanine, oleamide, and α-eleostearic acid; the secondary metabolites were mainly rhusflavanone, procyanidin B1, caffeic acid, and pantothenic acid. The discovery of key biomarkers is important for the early prediction, diagnosis, and classification of fruit quality. Deng et al.

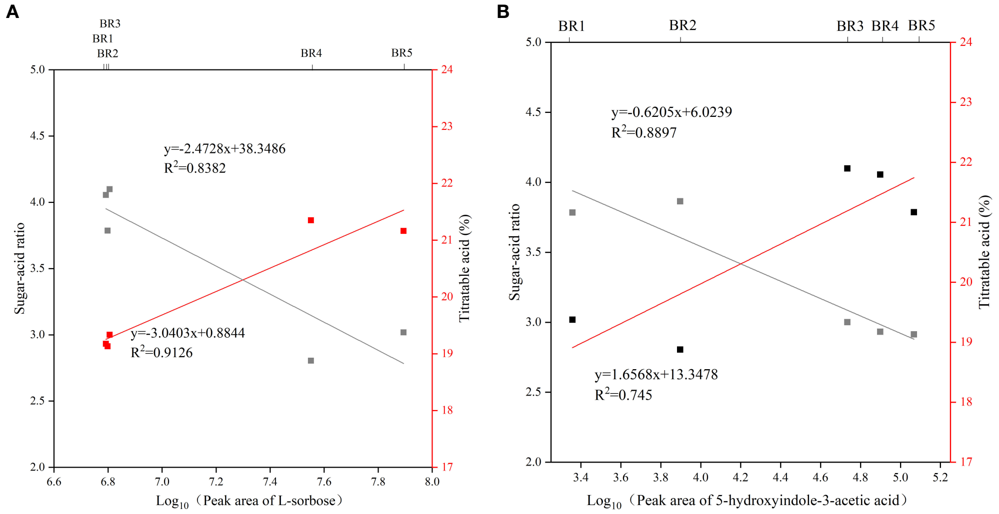


FIGURE 6
Correlation between potential taste biomarker, sugar–acid ratio, and titratable acids in BR *B. ramiflora*. **(A)** L-sorbose; **(B)** 5-hydroxyindole-3-acetic acid.

(2023) explained the main reasons for fruit taste changes by mining key biomarkers of rambutan. L-sorbose and 5-hydroxyindole-3-acetic acid were closely associated with sugar–acid ratio and titratable acids in fruit development, demonstrating that they may be key taste biomarkers of the BR fruit. Our study provides not only a new insight into the metabolic changes during *B. ramiflora* ripening but also an important reference for the improvement of *B. ramiflora* quality.

B. ramiflora has high medicinal and health value (Inta et al., 2013). However, the use of natural plant products for medicinal purposes is still challenging because of their low content and complex structure and issues in mass production and synthesis (Sirirungruang et al., 2022). Fortunately, advances in technology related to synthetic chemistry, pharmaceutical chemistry, pharmacology, and molecular biology have made it much easier to discover new natural plant products (Li et al., 2023). In this study, 536 metabolites in BR *B. ramiflora* were identified by non-targeted metabolite analysis utilizing LC-MS/MS. According to different medicinal needs, we can determine the optimal fruit picking time by non-targeted metabolomes (Fang et al., 2023). The discovery of metabolites will provide a useful reference to further develop the medicinal value of *B. ramiflora*.

5 Conclusions

In conclusion, different from the previous physiological and biochemical studies on various stages of fruit ripening, this study provided a new perspective for understanding the critical metabolites during the various stages of *B. ramiflora* ripening. This study showed that the changes of metabolites at different maturity stages were closely associated with *B. ramiflora* fruit quality. As a result, through the non-targeted metabolomics analysis of five typical growth stages of the BR *B. ramiflora* fruit, 536 metabolites and 154 common metabolites were screened out. On the basis of pathway enrichment analysis, its characteristic metabolites and potential flavor biomarkers were identified, and its possible mechanisms were explained. The results provided new metabolic lines of evidence for taste variation during the ripening process of *B. ramiflora*. Future research can be combined with genomics and transcriptomics to better understand the relationship between gene expression, protein synthesis, and metabolite changes, so as to further explore the molecular basis of *B. ramiflora* fruit quality formation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

CY: Data curation, Formal analysis, Methodology, Project administration, Resources, Writing – original draft. JQC: Data curation, Formal analysis, Methodology, Project administration, Resources, Writing – original draft. YZ: Supervision, Validation, Writing – review & editing. JH: Investigation, Resources, Validation, Writing – original draft. HW: Data curation, Formal analysis, Software, Validation, Writing – review & editing. JC: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the program for scientific research start-up funds of Guangdong Ocean University (060302052305), Forestry Science and Technology Innovation Project of Guangdong province (2018KJCX023).

Acknowledgments

We would like to thank the editor and reviewers who revised this manuscript and for providing insightful comments.

Conflict of interest

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1420231/full#supplementary-material

References

Aldhanhani, A. R. H., Ahmed, Z. F. R., Tzortzakis, N., and Singh, Z. (2022). Maturity stage at harvest influences antioxidant phytochemicals and antibacterial activity of jujube fruit (*Ziziphus mauritiana* Lamk. and *Ziziphus spina-christi* L.). *Ann. Agric. Sci.* 67, 196–203. doi: 10.1016/j.aoas.2022.12.003

Al-Maiman, S. A., and Ahmad, D. (2002). Changes in physical and chemical properties during pomegranate (*Punica granatum* L.) fruit maturation. *Food Chem.* 76, 437–441. doi: 10.1016/S0308-8146(01)00301-6

Banerjee, S., Islam, S., Chattopadhyay, A., Sen, A., and Kar, P. (2022). Synthesis of silver nanoparticles using underutilized fruit *Baccaurea ramiflora* (Latka) juice and its biological and cytotoxic efficacy against MCF-7 and MDA-MB 231 cancer cell lines. *South Afr. J. Bot.* 145, 228–235. doi: 10.1016/j.sajb.2021.09.016

Chen, J., Wu, F., Wang, H., Guo, C., Zhang, W., Luo, P., et al. (2023). Identification of key taste components in *Baccaurea ramiflora* Lour. fruit using non-targeted metabolomics. *Food Sci. Hum. Wellness* 12, 94–101. doi: 10.1016/j.fshw.2022.07.027

Cuthbertson, D., Andrews, P. K., Reganold, J. P., Davies, N. M., and Lange, B. M. (2012). Utility of metabolomics toward assessing the metabolic basis of quality traits in apple fruit with an emphasis on antioxidants. *J. Agric. Food Chem.* 60, 8552–8560. doi: 10.1021/jf3031088

Deng, H., Wu, G., Zhang, R., Yin, Q., Xu, B., Zhou, L., et al. (2023). Comparative nutritional and metabolic analysis reveals the taste variations during yellow rambutan fruit maturation. *Food Chemistry: X* 17, 100580. doi: 10.1016/j.fochx.2023.100580

Dhungana, S. R., and Braun, D. M. (2021). Sugar transporters in grasses: Function and modulation in source and storage tissues. *J. Plant Physiol.* 266, 153541. doi: 10.1016/j.jplph.2021.153541

Di Guida, R., Engel, J., Allwood, J. W., Weber, R. J. M., Jones, M. R., Sommer, U., et al. (2016). Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics* 12 (5), 1–14. doi: 10.1007/S11306-016-1030-9/TABLES/5

Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-Mcintyre, S., Anderson, N., et al. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 6, 1060–1083. doi: 10.1038/nprot.2011.335

Fang, Q., Zheng, H., Fu, G., Yin, M., Jiang, L., Zhao, Y., et al. (2023). Integrated untargeted metabolome, full-length sequencing, and transcriptome analyses reveal insights into the fruit quality at different harvest times of *Chaenomeles speciosa*. *Food Res. Int.* 164, 112314. doi: 10.1016/j.foodres.2022.112314

Feng, B. S., Liu, L. X., Sun, J., Leng, P., Wang, L., Guo, Y., et al. (2024). Combined metabolome and transcriptome analyses of quality components and related molecular regulatory mechanisms during the ripening of Huangjin Peach. *Scientia Hortic.* 327, 112787. doi: 10.1016/j.scienta.2023.112787

Ferrer, J. L., Austin, M. B., Stewart, C., and Noel, J. P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol. Biochem.* 46, 356–370. doi: 10.1016/j.plaphy.2007.12.009

Fyfe, S., Smyth, H. E., Joachim, H., Rychlik, M., Sultanbawa, Y., and Kurniawan, N. D. (2023). Postharvest Biology and Technology Physical properties and magnetic resonance imaging of the Australian green plum (*Buchanania obovata*) through maturity and ripening and across locations. *Postharvest Biol. Technol.* 205, 112494. doi: 10.1016/j.postharvbio.2023.112494

Guo, S., Duan, J. A., Qian, D., Tang, Y., Wu, D., Su, S., et al (2015). Content variations of triterpenic acid, nucleoside, nucleobase, and sugar in jujube (Ziziphus jujuba) fruit during ripening. *Food Chemistry* 167, 468–474. doi: 10.1016/J.FOODCHEM.2014.07.013

Huang, J., Chen, J., Shi, M., Zheng, J., Chen, M., Wu, L., et al. (2024). Genome assembly provides insights into the genome evolution of *Baccaurea ramiflora* Lour. *Sci. Rep.* 14, 1–11. doi: 10.1038/s41598-024-55498-4

Huang, H., Zou, S., and Cheng, C. (2021). Domestication and breeding strategy of wild fruit trees on track of plant introduction and domestication history. *J. Plant Genet. Resour.* 22, 1463–1473. doi: 10.13430/J.CNKI.JPGR.20210902001

Inta, A., Trisonthi, P., and Trisonthi, C. (2013). Analysis of traditional knowledge in medicinal plants used by Yuan in Thailand. *J. Ethnopharmacology* 149, 344–351. doi: 10.1016/j.jep.2013.06.047

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003

Jiang, X., Liu, K., Peng, H., Fang, J., Zhang, A., Han, Y., et al. (2023). Comparative network analysis reveals the dynamics of organic acid diversity during fruit ripening in peach (Prunus persica L. Batsch). *BMC Plant Biol.* 23, 1–14. doi: 10.1186/s12870-023-04037-w

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research* 40 (Database issue), 109–114. doi: 10.1093/NAR/GKR988

Khalil, U., Rajwana, I. A., Razzaq, K., Farooq, U., Saleem, B. A., and Brecht, J. K. (2023). Quality attributes and biochemical changes in white and colored table grapes as influenced by harvest maturity and ambient postharvest storage. *South Afr. J. Bot.* 154, 273–281. doi: 10.1016/j.sajb.2023.01.044

Kong, F., Yan, Z., Zhou, Z., Zhou, C., Lu, M., Zhou, F., et al. (2024). Quality and storability of *Baccaurea ramiflora* Lour. with different maturity. *Food Fermentation Industries*. doi: 10.13995/j.cnki.11-1802/ts.037979

Li, Z., Wang, Y., Xu, M., Liu, H., Li, L., and Xu, D. (2023). Molecular mechanism overview of metabolite biosynthesis in medicinal plants. *Plant Physiol. Biochem.* 204, 108125. doi: 10.1016/j.plaphy.2023.108125

Li, C. F., Yao, M. Z., Ma, C. L., Ma, J. Q., Jin, J. Q., and Chen, L. (2015). Differential metabolic profiles during the albescent stages of "Anji Baicha" (*Camellia sinensis*). *PloS One* 10, 1–18. doi: 10.1371/journal.pone.0139996

Liu, Z., Wang, H., Zhang, J., Chen, Q., He, W., Zhang, Y., et al. (2024). Comparative metabolomics profiling highlights unique color variation and bitter taste formation of Chinese cherry fruits. *Food Chem.* 439, 138072. doi: 10.1016/j.foodchem.2023.138072

Ma, B., Chen, J., Zheng, H., Fang, T., Ogutu, C., Li, S., et al. (2015). Comparative assessment of sugar and malic acid composition in cultivated and wild apples. *Food Chemistry* 172, 86–91. doi: 10.1016/J.FOODCHEM.2014.09.032

Maeda, H., and Dudareva, N. (2012). The shikimate pathway and aromatic amino Acid biosynthesis in plants. *Annu. Rev. Plant Biol.* 63, 73–105. doi: 10.1146/annurev-arplant-042811-105439

Nesa, M. L., Karim, S. M. S., Api, K., Sarker, M. M. R., Islam, M. M., Kabir, A., et al. (2018). Screening of *Baccaurea ramiflora* (Lour.) extracts for cytotoxic, analgesic, anti-inflammatory, neuropharmacological and antidiarrheal activities. *BMC Complementary Altern. Med.* 18, 1–10. doi: 10.1186/s12906-018-2100-5

Pandey, Y., Upadhyay, S., Bhatt, S. S., Sharma, L., Manivannan, S., and Chanbisana, C. (2018). Nutritional Compositions of *Baccaurea sapida* and *Eleaocarpus sikkimnesis* of Sikkim Himalaya. *Int. J. Curr. Microbiol. Appl. Sci.* 7, 2101–2106. doi: 10.20546/ijcmas.2018.702.250

Prakash, O., Supriya, A., and Kudachikar, V. B. (2020). Physicochemical changes, phenolic profile and antioxidant capacities of colored and white grape (*Vitis Vinifera* L.) varieties during berry development and maturity. *Int. J. Fruit Sci.* 20, S1773–S1783. doi: 10.1080/15538362.2020.1833809

Ramjan, M., and Raghavan, M. (2018). Burmese grape (*Baccaurea ramiflora* Lour.): A promising fruit crop for future generations. *J. Medicinal Plants Stud.* 6, 50–52.

Serrano-García, I., Hurtado-Fernández, E., Gonzalez-Fernandez, J. J., Hormaza, J. I., Pedreschi, R., Reboredo-Rodríguez, P., et al. (2022). Prolonged on-tree maturation vs. cold storage of Hass avocado fruit: Changes in metabolites of bioactive interest at edible ripeness. *Food Chem.* 394, 133447. doi: 10.1016/j.foodchem.2022.133447

Sirirungruang, S., Markel, K., and Shih, P. M. (2022). Plant-based engineering for production of high-valued natural products. *Natural Product Rep.* 39, 1492–1509. doi: 10.1039/D2NP00017B

Su, C., Yang, M., Chen, S., Fu, C., Zhang, L., Liu, S., et al. (2024). Multiple metabolite profiles uncover remarkable bioactive compounds and metabolic characteristics of noni fruit (*Morinda citrifolia* L.) at various stages of ripeness. *Food Chem.* 450, 139357. doi: 10.1016/j.foodchem.2024.139357

Ullah, M. O., Urmi, K. F., Howlader, M. A., Hossain, M. K., Ahmed, M., and Hamid, K. (2012). Hypoglycemic, hypolippidemic and antioxidant effects of leaves methanolic extract of *Baccaurea ramiflora*. *Int. J. Pharm. Pharm. Sci.* 4, 266–269.

Usha, T., Middha, S. K., Bhattacharya, M., Lokesh, P., and Goyal, A. K. (2014). Rosmarinic acid, a new polyphenol from *baccaurea ramiflora* lour. Leaf: A probable compound for its anti-inflammatory activity. *Antioxidants* 3, 830–842. doi: 10.3390/antiox3040830

Wang, T., Peng, H., Cao, Y., Xu, J., Xiong, Y., Liu, K., et al. (2022). Dynamic Network Biomarker Analysis Reveals the Critical Phase Transition of Fruit Ripening in Grapevine. *Genes (Basel).* 13 (10), 1851. doi: 10.3390/genes13101851

Wang, H., Zhang, H., Yang, Y., Li, M., Zhang, Y., Liu, J., et al. (2019). The control of red colour by a family of MYB transcription factors in octoploid strawberry (Fragaria × ananassa) fruits. *Plant Biotechnol. J.* 18, 1169–1184. doi: 10.1111/PBI.13282

Wen, B., Mei, Z., Zeng, C., and Liu, S. (2017). metaX: A flexible and comprehensive software for processing metabolomics data. *BMC Bioinf.* 18, 1–14. doi: 10.1186/s12859-017-1579-y

Yin, Q.-C., Ji, J.-B., Zhang, R.-H., Duan, Z.-W., Xie, H., Chen, Z., et al. (2022). Identification and verification of key taste components in wampee using widely targeted metabolomics. *Food Chemistry: X* 13, 100261. doi: 10.1016/j.fochx.2022.100261

Yue, W., Sun, W., Rao, R. S. P., Ye, N., Yang, Z., and Chen, M. (2018). Non-targeted metabolomics reveals distinct chemical compositions among different grades of Bai Mudan white tea. *Food Chem.* 277, 289–297. doi: 10.1016/j.foodchem.2018.10.113

Zhang, Y., Butelli, E., and Martin, C. (2014). Engineering anthocyanin biosynthesis in plants. *Curr. Opin. Plant Biol.* 19, 81–90. doi: 10.1016/j.pbi.2014.05.011

Zhang, Y., Li, P., and Cheng, L. (2010). Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh. *Food Chem.* 123, 1013–1018. doi: 10.1016/j.foodchem.2010.05.053

Zhang, A., Zhou, H., Jiang, X., Han, Y., and Zhang, X. (2021). The draft genome of a flat peach (Prunus persica l. cv. '124 pan') provides insights into its good fruit flavor traits. *Plants* 10, 1–17. doi: 10.3390/plants10030538

Zhen, Z. L. J. (2015). Advances in research on sugars, organic acids and their effects on taste of fruits. *J. Fruit Sci.* 32, 304–312. doi: 10.13925/j.cnki.gsxb.20140271

Zheng, B., Zhao, Q., Wu, H., Ma, X., Xu, W., Li, L., et al. (2022). Metabolomics and transcriptomics analyses reveal the potential molecular mechanisms of flavonoids and carotenoids in guava pulp with different colors. *Scientia Horticulturae* 305, 111384. doi: 10.1016/j.scienta.2022.111384

Zishun, H., Fangfang, L., Shaoling, Z., Ru, B., and Huping, Z. (2016). Changes in free fatty acid and free amino acid contents in fruit of Pyrus sinkiangensis'Kuerle Xiangli'by bagging. *J. Fruit Sci.* 33, 804–813. doi: 10.13925/j.cnki.gsxb.20150510

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 19 July 2024
DOI 10.3389/fpls.2024.1433575

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Ying Lyu,
Hainan University, China
Hansheng Gong,
Ludong University, China

*CORRESPONDENCE
Huajun Gao
gaohuajun_81@163.com
Zhaoliang Geng
zhaolianggeng1104@126.com

†These authors have contributed equally to this work

RECEIVED 16 May 2024
ACCEPTED 03 July 2024
PUBLISHED 19 July 2024

CITATION
Yan T, Cai B, Li F, Guo D, Xia C, Lv H, Lin B, Gao H and Geng Z (2024) Proteomic and metabolomic revealed the effect of shading treatment on cigar tobacco. *Front. Plant Sci.* 15:1433575. doi: 10.3389/fpls.2024.1433575

Proteomic and metabolomic revealed the effect of shading treatment on cigar tobacco

Tongjing Yan†, Bin Cai†, Fangyou Li, Dong Guo, Changjian Xia, Hongkun Lv, Beisen Lin, Huajun Gao* and Zhaoliang Geng*

Haikou cigar Research Institute, Hainan Provincial Branch of China National Tobacco Corporation, Haikou, China

Shading or low light conditions are essential cultivation techniques for cigar wrapper tobacco leaves production, yet their impact on protein and metabolic regulatory networks is not well understood. In this study, we integrated proteomic and metabolomic analyses to uncover the potential molecular mechanisms affecting cigar tobacco leaves under shading treatment. Our findings include: (1) Identification of 780 significantly differentially expressed proteins (DEPs) in the cigar wrapper tobacco leaves, comprising 560 up-regulated and 220 down-regulated proteins, predominantly located in the chloroplast, cytoplasm, and nucleus, collectively accounting for 50.01%. (2) Discovery of 254 significantly differentially expressed metabolites (DEMs), including 148 up-regulated and 106 down-regulated metabolites. (3) KEGG pathway enrichment analysis revealed that the mevalonate (MVA) pathway within 'Terpenoid backbone biosynthesis' was inhibited, leading to a down-regulation of 'Sesquiterpenoid and triterpenoid biosynthesis'. Conversely, the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway was enhanced, resulting in an up-regulation of 'Monoterpenoid biosynthesis', 'Diterpenoid biosynthesis', and 'Carotenoid biosynthesis', thereby promoting the synthesis of terpenoids such as carotenoids and chlorophylls. Simultaneously, the Calvin cycle in 'Carbon fixation in photosynthetic organisms' was amplified, increasing photosynthetic efficiency. These results suggest that under low light conditions, cigar tobacco optimizes photosynthetic efficiency by reconfiguring its energy metabolism and terpenoid biosynthesis. This study contributes valuable insights into protein and metabolic analyses, paving the way for future functional studies on plant responses to low light.

KEYWORDS

proteomic, metabolomic, shading treatment, cigar tobacco, biosynthesis of terpenoids, Calvin cycle

1 Introduction

Light plays a crucial role in the lifecycle of plants, acting as a key ingredient in their growth and development processes. This was emphasized by De Wit and his colleagues in 2016 (De Wit et al., 2016). Light is not only pivotal for photosynthesis, but also for the plant's ability to produce food, demanding adaptability to the ever-changing sunlight exposure, as noted by Bukhov in 2004 (Bukhov, 2004). The ability of a plant to effectively adapt to varying light conditions is essential for its growth, development, and survival. Plants derive their energy from light; photosynthesis converts light energy into chemical energy via light-dependent reactions, a process detailed by Liang and others in 2016 (Liang and Lindblad, 2016). Insufficient light, therefore, poses a major challenge to plant growth (Lu et al., 2019), impacting their physiological traits, including morphology, chlorophyll content, photosynthetic efficiency, and the expression of various genes and proteins. This was explored in studies by Liang, Feng, Gao, and Yang between 2016 and 2020 (Liang and Lindblad, 2016; Feng et al., 2018; Gao et al., 2020; Yang et al., 2020). The costly and elusive cigar wrapper leaves, essential for the production of premium cigars, are sourced from plants cultivated in shaded conditions. These cigar wrapper leaves, renowned for their distinctive qualities, serve as the outer covering during the final stages of crafting exquisite cigars. Borges and his team in 2012 found that shading enhances the flavor quality of the leaves and prevents them from becoming too thick or nicotine-rich (Borges et al., 2012). Some scholars also proposed that low light could change the photosynthesis process of cigars by regulating chlorophyll content, chlorophyll fluorescence and gene expression (Wu et al., 2021). These shading treatments make tobacco leaves thinner and better looking, which could be used as cigar wrapper and increase the value of tobacco leaves. Despite these practices, the self-regulatory mechanisms of cigar tobacco under shaded conditions are not fully understood.

The rise of high-throughput sequencing and mass spectrometry has dramatically advanced omics technologies, encompassing genomics, transcriptomics, proteomics, and metabolomics. These methods offer a comprehensive view of gene transcription, signal transduction, protein expression, and metabolic reactions, as cited in sources (Lei et al., 2014; Hafidh et al., 2016; Li et al., 2017). Proteomics, in particular, bridges the gap between gene transcription and metabolic responses, quantifying protein levels. Integrating proteomics with transcriptomics enhances understanding of regulatory networks. Combining it with metabolomics sheds light on post-transcriptional regulation in metabolic changes, as found in studies by Liu, Haider, and Liang between 2013 and 2016 (Haider and Pal, 2013; Liang et al., 2016; Liu et al., 2019a). Metabolomics, a critical component of systems biology, identifies a wide range of endogenous metabolites and is widely used in food, tea, and tobacco research (Daglia et al., 2014; Liu et al., 2019b, 2020; Tsaballa et al., 2020; Zheng et al., 2021). Previous research, including transcriptomic and proteomic analyses, revealed complex gene and protein expression patterns in plants under stress, as Venkataramanan et al. high-lighted in 2015 (Venkataramanan et al., 2015). Studies have explored transcriptomic responses to low light in crops like rice and cherry (Sekhar et al., 2019; Tian et al., 2020), and proteomic responses in maize and rice (Liu et al., 2014; Gao et al., 2020). However, research on cigar tobacco's proteomic and metabolomic responses to low light is limited, with a few studies addressing high light conditions and changes in cigar tobacco leaves under varying light (Schuster et al., 2020; Ma et al., 2023).

This study investigates the adaptive mechanisms and regulatory responses of cigar tobacco to low light. Using proteomic and metabolomic analyses, we aim to uncover the changes in proteins and metabolites in cigar tobacco leaves under these conditions. Our approach seeks to provide new insights into the plant's response and physiological adaptations to light stress, enhancing our understanding of plant resilience in suboptimal light environments.

2 Materials and methods

2.1 Plant materials and growth conditions

For our study, we selected the Haiyan 103 variety of *Nicotiana tabacum L.*, grown in Danzhou City, Hainan Province. We used the same variety for both shaded (ST) and non-shaded treatments (NST) to ensure consistency. In line with local guidelines for quality cigar tobacco cultivation, each treatment was given 180 Kg of pure nitrogen, 270 Kg of P_2O_5, and 360 Kg of K_2O per hectare. Planting commenced on January 8, 2023, and 66 square meters were planted for ST and NST groups, respectively. Shading started on February 8, 2023. The shading net material is white polyethylene, and the transparency of the shading net is 75%. Sampling from both ST and NST was conducted on March 15, 2023. At the time of sampling, eight plants were randomly selected from each group as a biological duplicate sample. The 7th leaf on cigar tobacco plants (counting from the base to the tip, intact leaves with an area of around 48 × 26 cm) was selected for sampling, and the main vein of leaf was removed. The leaves were pooled, wrapped in foil, and flash-frozen in liquid nitrogen. Three independent biological replicates were used for proteomic analysis, while six independent biological replicates were used for metabolomic analysis. For each independent biological replicate, we conducted sufficient freezing and grinding with liquid nitrogen during sampling to ensure the uniformity of the sample.

2.2 Proteomic analysis

2.2.1 Protein isolation and digestion process

Following Wu et al.'s methodology (Wu et al., 2020), protein isolation began with grinding 0.5 g of tobacco leaf tissue into powder in liquid nitrogen. The powder was then processed with a lysis buffer (50 mM Tris-HCl, pH 8, 8 M urea, 0.2% Sodium dodecyl sulfate). To break disulfide bonds, we added 10 mM dithiothreitol to the mixture and incubated it at 56°C for one hour. Afterward, we added iodoacetic acid and let the mixture sit for another hour in a dark environment at room temperature. The samples were then mixed with four times their volume of cold acetone, vortexed, and left at −20°C overnight. Following centrifugation (12000 × g, 4°C for 5 min), we washed the sample twice with cold acetone (4 × volume), dissolved them in a solution of 0.1 M Triethylammonium bicarbonate (TEAB, pH 8.5) and 8 M urea, and measured the protein concentration using the Bradford method.

Protein digestion was carried out as described by Zhang et al (Zhang et al., 2016). We diluted each protein sample to 100 µL using DB lysis buffer (8 M urea, 100 mM TEAB, pH 8.5), then added trypsin and 100 mM TEAB, incubating it at 37°C for four hours. Further trypsin and $CaCl_2$ were added for overnight digestion. Formic acid was mixed with digested sample, adjusted pH to 3, and centrifuged the sample at 12000 × g for 5 min at room temperature. The supernatant was slowly loaded onto the C_{18} desalting column, washed 3 times with washing buffer (0.1% formic acid, 3% acetonitrile), and then eluted with the elution buffer (0.1% formic acid, 70% acetonitrile). The eluates were then lyophilized for analysis.

2.2.2 Proteomic analysis via UHPLC-MS/MS

For proteomic analysis, we used an Easy-nLC^{TM} 1200 UHPLC and a Q Exploris^{TM} HF-X mass spectrometer. We loaded 4 µg of each sample, mixed with iRT reagent, onto a C_{18} Nano-Trap column. The gradient profile ranged from 5% to 95% acetonitrile in formic acid over 92 min, with a flow rate of 600 nL/min. The peptides underwent analysis by a Q Exactive^{TM} HF-X mass spectrometer, with specific settings for full scan range, resolution, AGC target, ion injection time, and fragmentation settings. The top 40 precursors were selected for MS/MS analysis, and DIA mode was utilized for broader peptide coverage.

2.2.3 Identification and quantification of proteins

We independently searched the spectra from each fraction against the Nicotiana tabacum protein database using Proteome Discoverer 2.2. The search parameters were finely tuned for precision, and the proteins identified needed to meet stringent criteria including FDR and amino acid coverage.

2.2.4 Data analysis and statistical approaches

For functional annotation, we utilized Gene Ontology and InterPro analyses via InterProScan, analyzing against a comprehensive protein database. We also employed COG and KEGG databases for further analysis of protein families and pathways. Differential protein expression was assessed using various analytical tools, including Volcano plot and heat map analysis. Protein-protein interactions were predicted using the STRING-db server.

2.3 Metabolomic analysis

2.3.1 Process of metabolite extraction and untargeted metabolomic analysis

We began by grinding 0.1 g of leaf tissue in liquid nitrogen to create a fine powder. This powder was then reconstituted in 80% methanol that had been chilled beforehand. After vigorous shaking of the mixture, it was left on ice for 5 min before being subjected to centrifugation at 15,000 × g at a temperature of 4°C, lasting for 20 min. We then diluted a part of the clear supernatant to achieve a concentration of 53% methanol, utilizing water of UHPLC-MS/MS grade. The diluted mixture was placed into new Eppendorf tubes and centrifuged once more under identical conditions. The final supernatant obtained from this process was then prepared for analysis via UHPLC-MS/MS.

The analysis of the samples was conducted using a Vanquish UHPLC system from ThermoFisher (Germany), in conjunction with an Orbitrap Q ExactiveTM HF mass spectrometer. Samples were injected onto a Hypersil Gold column (100 × 2.1 mm, 1.9 µm) using a 12 min linear gradient at a flow rate of 0.2 mL/min. The eluents for the positive polarity mode were eluent A (0.1% Formic acid in Water) and eluent B (Methanol). The eluents for the negative polarity mode were eluent A (5 mM Ammonium acetate, pH 9.0) and eluent B (Methanol). The solvent gradient was set as follows: 2% B, 1.5 min; 2–85% B, 3 min; 85–100% B, 10 min; 100–2% B, 10.1 min; 2% B, 12 min. The mass spectrometer settings, including spray voltage, capillary temperature, gas flow rates, and S-lens RF level, were finely adjusted. The MS scan covered a range from 90 to 900 m/z, utilizing fragmentation data acquisition methods in both polarity modes.

2.3.2 Processing and analysis of metabolomic data

For the processing of metabolomics data, we used Compound Discoverer 3.3, following the approach described by Ribbenstedt et al (Ribbenstedt et al., 2018). The metabolite identification process involved querying the acquired data against an array of databases such as mzCloud, mzVault, KEGG, and several others, maintaining a precise mass tolerance. For the purpose of annotating these metabolites, we utilized well-known databases like KEGG, HMDB, and LIPID Maps, each offering a wealth of information for accurate metabolite identification and classification.

2.4 Synthesis of proteomic and metabolomic data

The integration of the identified DEPs and DEMs into the KEGG pathway maps facilitated the visualization of alterations in crucial metabolic pathways. We focused on pathways including 'Terpenoid backbone biosynthesis', and others related to terpenoid synthesis, as well as the 'Carbon fixation in photosynthetic organisms' pathway to depict the Calvin cycle. This integration helped in constructing a detailed representation of the terpenoid biosynthesis process.

3 Results

3.1 Comparative analysis of protein profiles in ST and NST groups

3.1.1 Multivariate statistical evaluation of protein profiles

The unsupervised, multi-factor principal component analysis (PCA) method was used to analyze the protein profile data of cigar tobacco leaves treated with shading and non-shading. The results are shown in Figure 1. The two groups of cigar tobacco leaves were significantly separated along the first principal component, while the dispersion trend between samples within each treatment group was not obvious and all samples were within the 95% confidence interval, indicating that there were significant differences between the protein profile of cigar tobacco leaves in ST and NST groups.

3.1.2 Examination of DEPs in cigar tobacco leaves

The proteomic of cigar tobacco leaves in ST and NST groups allowed us to quantify a total of 12553 proteins. In this study, all proteins were screened and analyzed using a threshold of VIP > 1, FC > 1.5 or FC < 0.67, *P* < 0.05. The results showed that compared to NST group, cigar tobacco leaves in ST group included 780 DEPs, of which 560 DEPs were up-regulated and 220 DEPs were down-regulated (Figure 2A; Supplementary Table S1). The heatmap analysis was conducted on 780 DEPs of cigar tobacco leaves in ST and NST groups (Figure 2B). The red color in heatmap represents DEPs with higher than average content, while the blue color represents DEPs with lower than average content. Comparing the intensity of all colors, it was found that most DEPs in the same treatment group showed similar patterns. According to Pearson correlation coefficients, all samples can be divided into two categories: ST and NST. This is similar to the results of PCA, indicating that there are significant differences in proteins of cigar tobacco leaves between ST and NST groups. As shown in Figure 2C and Supplementary Table S2, subcellular localization analysis found that most of DEPs of cigar tobacco leaves in ST and NST groups were located in chloroplast (19.50%), cytoplasm (16.04%), and nucleus (14.47%), accounting for 50.01%.

3.2 Comparative metabolite analysis in ST and NST groups

3.2.1 Multivariate statistical evaluation of metabolite profiles

The PCA results are shown in Figure 3A. The samples from ST and NST groups were significantly separated along the first and second principal components, while the dispersion trend between samples within each treatment group was not obvious and all samples were within the 95% confidence interval, indicating significant differences between the metabolite profiles of cigar tobacco leaves in the two treatment groups. The partial least squares discriminant analysis (PLS-DA) is shown in Figure 3B. The two groups of cigar samples are clearly separated in the PLS-DA score diagram. In the PLS-DA model, the first two components R2 Y and Q2 Y dominate, with values of 99% and 78%, respectively, indicating that the model has good stability and that the selection of the first two components is sufficient to test the dataset. In addition, permutation testing was used to evaluate whether the PLS-DA model was 'overfitting'. As shown in Figure 3C, R^2 was 0.92 and Q^2 was -0.70. R^2>Q^2 and the intercept of the Q^2 regression line with the Y axis was less than 0, indicating that the PLS-DA model was not 'overfitting' and could better describe the classification of samples, which indicating that ST had a significant impact on the metabolite profile of cigar tobacco leaves.

3.2.2 Exploration of DEMs between ST and NST groups

The metabolic of cigar tobacco leaves in ST and NST groups allowed us to quantify a total of 1163 metabolites. In this study, all

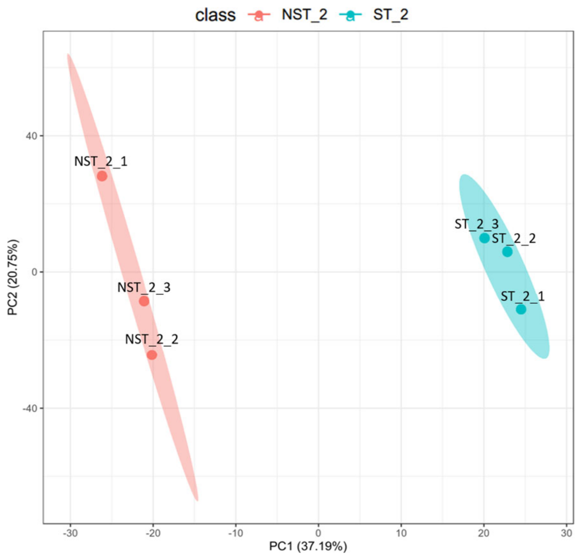


FIGURE 1
PCA diagram of cigar tobacco leaves in ST and NST groups. The abscissa PC1 and the ordinate PC2 in the figure represent the scores of the first and second principal components, respectively. The ellipse is the 95% confidence interval. The sample number of NST_2 is NST_2_1~3, and the sample number of ST_2 is ST_2_1~3.

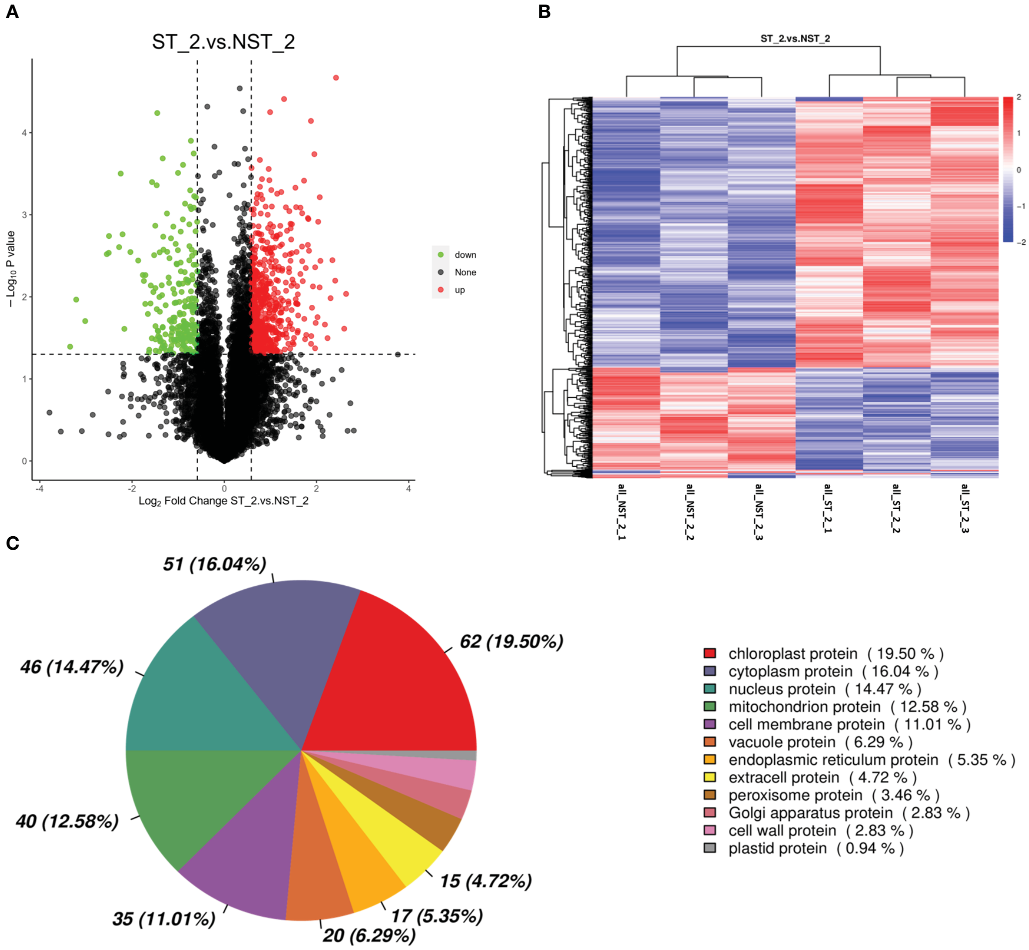


FIGURE 2
Volcano map, cluster heat map and subcellular localization of DEPs of cigar tobacco leaves in ST and NST groups. **(A)** Abscissa represents the multiple change of expression of proteins in different groups ($\log_2$ Fold Change), ordinate indicates the significant level of difference ($-\log_{10}$ P-value), each point in the volcanic map represents a protein, significantly up-regulated proteins are represented by red dots, significantly down-regulated proteins are represented by green dots. **(B)** ST_2_1–3 and NST_2_1–3 represent the six repeated samples for the analysis of proteomic of ST and NST groups, respectively. **(C)** Location of DEPs in cells.

metabolites were screened and analyzed using a threshold of VIP > 1, FC > 1.2 or FC < 0.83, $P < 0.05$. The results showed that compared to NST group, cigar tobacco leaves in ST group included 254 DEMs, of which 148 DEMs were up-regulated and 106 DEMs were down-regulated (Figure 4A; Supplementary Table S3). The heatmap analysis was conducted on 254 DEMs of cigar tobacco leaves in ST and NST groups (Figure 4B). The red color in heatmap represents DEMs with higher than average content, while the blue color represents DEMs with lower than average content. Comparing the intensity of all colors, it was found that most DEMs in the same treatment group showed similar patterns. According to Pearson correlation coefficients, all samples can be divided into two categories: ST and NST. This is similar to the results of PCA and PLS-DA, indicating that there are significant differences in metabolites of cigar tobacco leaves between ST and NST groups.

3.3 Merging proteomic and metabolomic data studies

Seven key pathways were obtained through KEGG enrichment pathway association analysis of DEPs and DEMs screened out above (Figure 5). DEPs and DEMs were mainly enriched in the pathways of 'Terpenoid backbone biosynthesis', 'Cutin, suberine and wax biosynthesis', 'Sesquiterpenoid and triterpenoid biosynthesis', 'Riboflavin metabolism', 'Purine metabolism', 'Diterpenoid biosynthesis', and 'Carbon fixation in photosynthetic organisms'. Among them, 'Terpenoid backbone biosynthesis', 'Sesquiterpenoid and triterpenoid biosynthesis' and 'Diterpenoid biosynthesis' pathways refer to the process of synthesizing terpenoids such as chlorophyll and carotenoids in cigar tobacco leaves through biochemical pathways; 'Carbon fixation in photosynthetic organisms' pathway refers to the process of converting atmospheric carbon

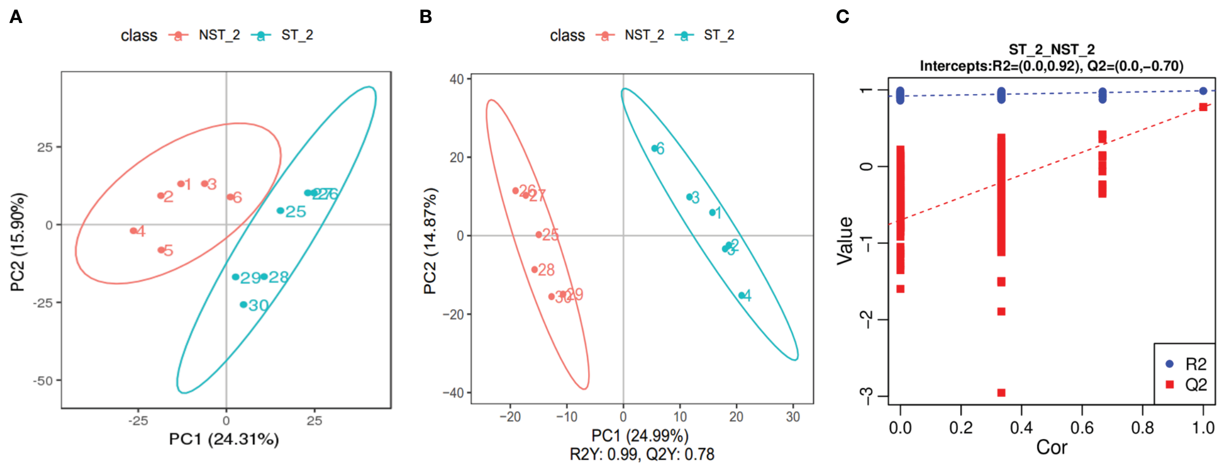


FIGURE 3
PCA diagram, PLS-DA score scatter diagram and sequencing verification diagram of cigar tobacco leaves in ST and NST groups. **(A)** The abscissa PC1 and the ordinate PC2 in the figure represent the scores of the first and second principal components, respectively. The ellipse is the 95% confidence interval. The sample number of NST_2 is 1~6, and the sample number of ST_2 is 25~30. **(B)** The abscissa is the score of the sample on the first principal component, the ordinate is the score of the sample on the second principal component. The sample number of NST_2 is 1~6, and the sample number of ST_2 is 25~30. **(C)** The abscissa represents the correlation between the Y of the random grouping and the original group Y, and the ordinate represents the scores of R^2 and Q^2.

dioxide into carbon in its own organic matter through photosynthesis in cigar tobacco leaves; 'Cutin, suberine and wax biosynthesis' pathway refers to the process of combining long-chain fatty acids and alcohol compounds through biochemical pathways to form substances such as cutin, suberin and wax that constitute the cell structure of cigar tobacco leaves; 'Riboflavin metabolism' and 'Purine metabolism' pathways refer to the process of synthesizing and decomposing riboflavin and purine in cigar tobacco leaves through biochemical pathways.

Through association analysis of the enriched metabolic pathways (Figure 6), a terpenoid biosynthesis process diagram was obtained by combining the metabolic pathways of 'Terpenoid backbone biosynthesis', 'Sesquiterpenoid and triterpenoid biosynthesis', 'Monoterpenoid biosynthesis', 'Diterpenoid biosynthesis' and 'Carotenoid biosynthesis'. Calvin cycle diagram was also drawn based on the metabolic pathway of 'Carbon fixation in photosynthetic organisms'.

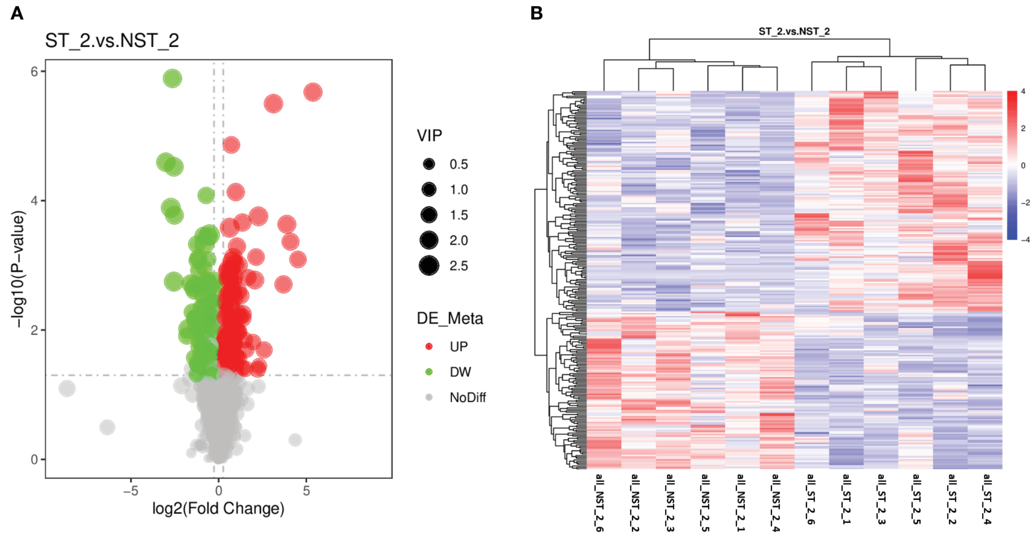


FIGURE 4
Volcano map and cluster heat map of DEMs of cigar tobacco leaves in ST and NST groups. **(A)** Abscissa represents the multiple change of expression of metabolites in different groups ($\log_2$ Fold Change), ordinate indicates the significant level of difference ($-\log_{10}$ P-value), each point in the volcanic map represents a metabolite, significantly up-regulated metabolites are represented by red dots, significantly down-regulated metabolites are represented by green dots, and the size of dots represents VIP values. **(B)** ST_2_1−6 and NST_2_1−6 represent the six repeated samples for the analysis of metabolomic of ST and NST groups, respectively.

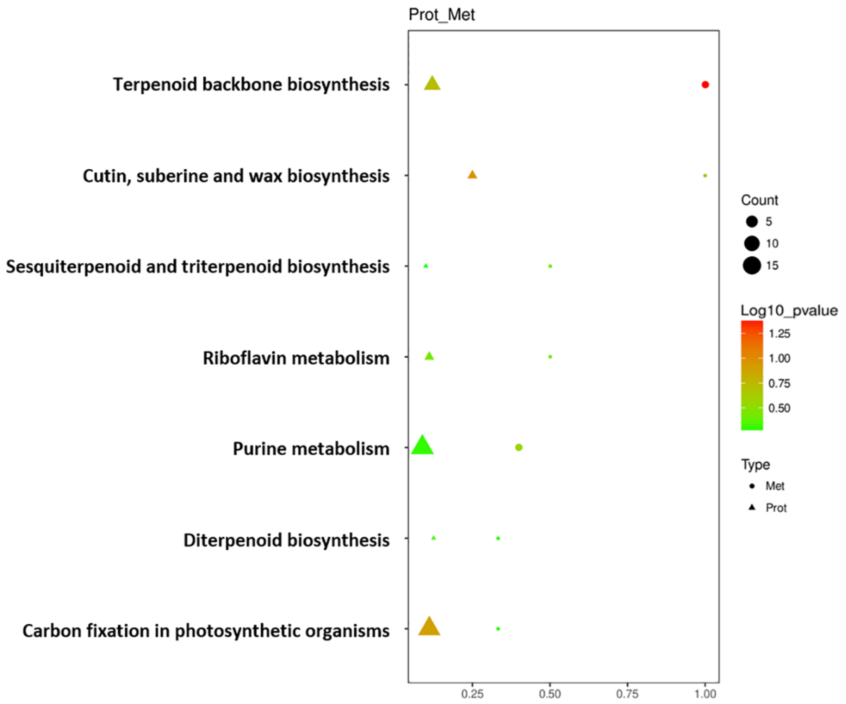


FIGURE 5
Bubble diagram of enrichment pathway of DEPs and DEMs of cigar tobacco leaves in ST and NST groups. Protein was represented by triangles, metabolite was represented by circles. Horizontal axis indicates the enrichment degree of DEPs and DEMs. Point color represents the P-value, and point size indicates the number of DEPs and DEMs in the corresponding pathway.

4 Discussion

Light's transformation into biochemical energy is vital in plant growth, playing an indispensable role in their development, as highlighted by Evans (Evans, 2013). In tobacco plants, shading impedes the availability of light, diminishing the rate of photosynthesis in leaves, leading to less material accumulation and thinner leaves, a phenomenon documented by Wu et al. (2021). Plants, however, have developed adaptive mechanisms to combat low-light stress, as detailed in the works of Gao et al (Gao et al., 2020). and Tian et al (Tian et al., 2020). Adopting a multi-faceted data integration approach, encompassing genetic, epigenetic, transcriptomic, and proteomic perspectives, can shed light on plant responses to various non-living stresses, a concept

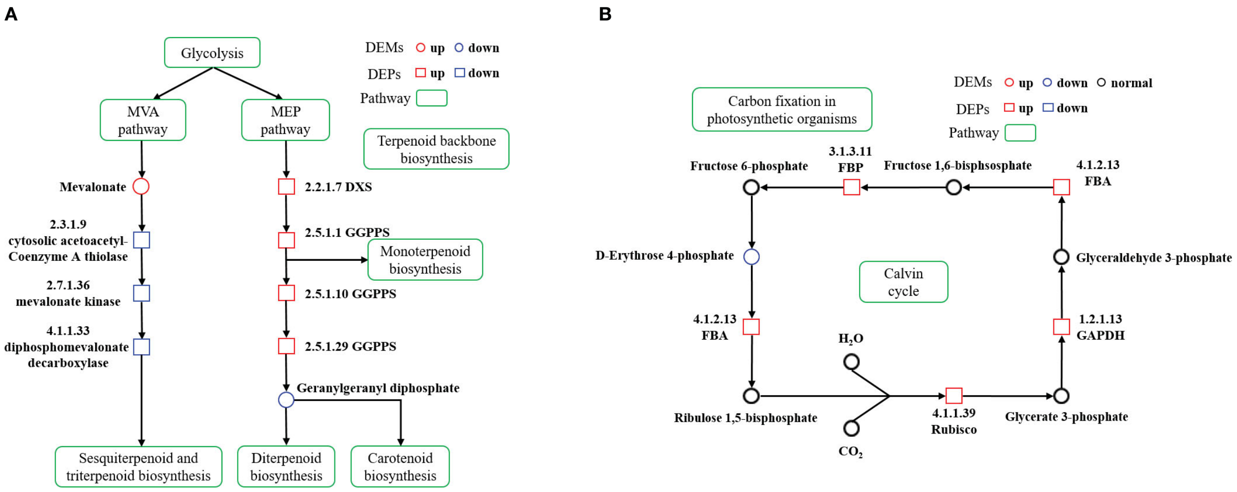


FIGURE 6
Terpenoid biosynthesis pathway diagram and differences in protein and metabolite content **(A)** Calvin cycle diagram and differences in protein and metabolite content **(B)** in cigar tobacco leaves. The proposed metabolic pathways were based on KEGG PATHWAY Database. The red represents high relative content and the blue represents low relative content. The green represents metabolic pathway.

explored by Haak et al (Haak et al., 2017). It's critical to recognize, as Liu et al. point out, that protein levels may not always be indicative of metabolite abundance (Liu et al., 2016). Sole reliance on a single methodology, like proteomics or metabolomics, risks partial or incorrect interpretations of the underlying processes. Thus, merging various omics methods is crucial for an accurate interpretation, as emphasized by Zapalska-Sozoniuk et al (Zapalska-Sozoniuk et al., 2019). In our research, shading treatment (ST) revealed significant variances in differentially expressed proteins (DEPs) and metabolites (DEMs). However, a direct correlation between the proteomic and metabolomic data was absent. This inconsistency could stem from factors like delayed metabolite reactions. Ding et al. noted that proteins react more swiftly to environmental shifts than metabolites, despite using identical plant samples and treatments (Ding et al., 2020).

4.1 The impact of shading treatment on terpenoid synthesis in cigar tobacco leaves

Terpenoids, a vast class of over 55,000 natural secondary metabolites, are vital in plant primary and secondary metabolism (Christianson, 2017; Pu et al., 2021). These compounds are synthesized through two distinct pathways: MVA and MEP pathways. Monoterpenes, diterpenes, and carotenoids typically originate from the MEP pathway, while sesquiterpenes and triterpenes are produced via the MVA pathway (Moser and Pichler, 2019; Chen et al., 2020). In the ST group, we noticed an upregulation of enzymes like geranylgeranyl diphosphate synthase (GGPPS, 2.5.1.1, 2.5.1.10, 2.5.1.29) and 1-deoxy-D-xylulose-5-phosphate synthase (DXS, 2.2.1.7) in the MEP pathway. On the contrary, cytosolic acetoacetyl-coenzyme A thiolase (2.3.1.9), mevalonate kinase (2.7.1.36) and diphosphomevalonate decarboxylase (4.1.1.33) in the MVA pathway were significantly downregulated, with an increase in mevalonic acid. These findings suggest a suppression of the MVA pathway and an enhancement of the MEP pathway, facilitating an increased conversion of glycolysis metabolites through the MEP pathway. The GGPPS in cigar tobacco promoted the further conversion of GGPP to terpenoids such as chlorophyll and carotenoids, thus enabling cigar tobacco leaves to adapt to the stress caused by shading (Figure 6A).

The MEP pathway starts with the transketolase decarboxylation of pyruvate and glyceraldehyde 3-phosphate (GA-3P), catalyzed by DXS, leading to 1-Deoxy-D-xylulose-5-phosphate (DXP). Subsequent reactions yield precursors for various isoprenoids, including centellosides. DXS, being the initial and flux-controlling enzyme, plays a critical role in isoprenoid biosynthesis (Estevez et al., 2001; Xiang et al., 2007; Zhang et al., 2018). However, DXS expression doesn't always correlate directly with terpenoid levels. For instance, Umaima showed that over-expressing cyanobacterial DXS in chloroplasts disrupted the chlorophyll-carotenoid balance in Chlamydomonas reinhardtii (Hoqani et al., 2022). Erpenoids in plants start from isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Compounds like geranyl diphosphate (GPP), farnesyl diphosphate (FPP), and geranyl-geranyl diphosphate (GGPP) evolve from these precursors. GGPP is particularly crucial for various terpenoids like diterpenes, gibberellins, a-tocotrienol, chlorophyll, and carotenoids (Zhou et al., 2017; Dong et al., 2022). GGPPS is instrumental in converting FPP to GGPP, earmarking it for specific terpenoid biosynthesis (Ruiz-Sola et al., 2016; Akhtar et al., 2017). In Arabidopsis, the enzyme AtGGPPS11, a critical isozyme, collaborates with enzymes involved in the biosynthesis of vital compounds such as carotenoids, chlorophyll, and plastoquinone. This includes interactions with enzymes like phytoene synthase (PSY), geranylgeranyl reductase (GGR), and solanesyl diphosphate synthase 2 (SPS2) (Walter and Strack, 2011; Ruiz-Sola et al., 2016). Similarly, in rice, the enzyme OsGGPPS1 plays a pivotal role in the biosynthesis of chlorophyll, leading to enhanced plant size and increased carotenoid levels (Zhou et al., 2017). In pepper, the enzyme CaGGPPS1 is noted for its specific interaction with PSY, directing the flow of GGPP towards carotenoid production (Wang et al., 2018). Tobacco plants have shown a unique adaptation, where the enzyme NtGGPPS1, functioning as a homodimer, works alongside NtPSY1 in chloroplasts to boost carotenoid biosynthesis, a pathway that differs from AtGGPPS11's broader role in terpenoid biosynthesis (Ruiz-Sola et al., 2016; Zhou et al., 2017). Advancements in this field include the design of a modified version of NtGGPPS1, as developed by Dong et al (Dong et al., 2022). This engineered enzyme has demonstrated a significant increase in carotenoid levels, photosynthetic efficiency, and biomass in tobacco plants. The improvement in biomass can be attributed to enhanced metabolic pathways linked to carotenoid biosynthesis, which includes the production of photoprotective molecules like β-carotene, zeaxanthin, and violaxanthin. These molecules contribute to improved photoprotection, leading to enhanced photosynthetic efficiency and biomass (Moreno et al., 2020; Kossler et al., 2021).

4.2 The impact of shading treatment on photosynthesis in cigar tobacco leaves

The Calvin cycle, a crucial carbon metabolism process in higher plants, operates in the chloroplast matrix and serves as the primary mechanism for CO_2 assimilation during the dark reaction stage of photosynthesis. This cycle produces carbon compounds essential for plant growth and development (Raines, 2003; Smith and Stitt, 2007; Anoman et al., 2016). In our study, we observed a significant upregulation of proteins and metabolites related to the 'Carbon fixation in photosynthetic organisms' pathway under low light conditions. Key enzymes in the Calvin cycle, such as ribulose bisphosphate carboxylase (Rubisco, 4.1.1.39), glycer-aldehyde-3-phosphate dehydrogenase (GAPDH, 1.2.1.13), and fructose-bisphosphate aldolase (FBA, 4.1.2.13) and fructose-bisphosphatase (FBP, 3.1.3.11) were found to be increased, while levels of D-erythrose 4-phosphate were decreased (Figure 6B). Rubisco, a rate-limiting enzyme in the Calvin cycle, plays a critical role in converting atmospheric CO_2 into energy storage molecules (Cheng et al., 1998; Portis and Parry, 2007). Despite its low catalytic efficiency, enhancing Rubisco's activity is seen as a key strategy to improve plant photosynthetic carbon assimilation efficiency (Wang et al., 2014; Wei et al., 2022). GAPDH, an enzyme with multiple isoforms, operates in both glycolysis and the Calvin cycle, playing a dual role in energy production and CO_2 fixation into carbohydrates (El Kadmiri et al., 2014; Anoman et al., 2016; Yu et al., 2020). FBA, a crucial enzyme in the RuBP regeneration stage of the Calvin cycle, significantly influences carbon fixation and flow (Zeng et al., 2014; Liang and Lindblad, 2016). Overexpression of FBA genes has been linked to improved plant growth and metabolism (Uematsu et al., 2012). FBP, a key enzyme in gluconeogenesis and sucrose synthesis, catalyzes the conversion of fructose 1,6-bisphosphate into fructose 6-phosphate and inorganic phosphate (Miyagawa et al., 2001). Enhancements in photosynthetic efficiency have been achieved by introducing modified versions of this enzyme into plant chloroplasts (De Porcellinis et al., 2018; Siddiqui et al., 2022). Our study found that key enzymes of the Calvin cycle were up-regulated under low light conditions, while the down-regulation of D-erythrose 4-phosphate suggests an efficient operation of this cycle. These findings indicate that cigar tobacco plants can adapt to low light conditions by reorganizing their energy metabolism, enabling them to endure stress more effectively (Thalmann and Santelia, 2017; Hwang et al., 2020).

5 Conclusion

In this study, we demonstrated that the changes in the intrinsic proteins and metabolites of cigar tobacco leaves are significantly related to the decrease in light intensity, manifested as the promotion of the biosynthesis of terpenoids such as carotenoids and chlorophyll, as well as the enhancement of the Calvin cycle in the "carbon fixation of photosynthetic organisms" pathway. This is the result of plants adjusting their energy requirements to compensate for the stress caused by shading, allowing plants to more effectively harvest light. This study offers novel insights into both proteomic and metabolomic processes as response mechanisms under different light intensities in cigar tobacco plants. In subsequent studies, we could combine transcriptomic and enzymatic activity assays to further clarify the impact of changes in key enzymes and metabolites on cigar tobacco leaves undergoing shading treatment.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

TY: Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. BC: Data curation, Methodology, Writing – review & editing. FL: Conceptualization, Writing – review & editing. DG: Data curation, Methodology, Software, Writing – original draft. CX: Methodology, Software, Writing – original draft. HL: Methodology, Writing – original draft. BL: Conceptualization, Writing – original draft. HG: Formal analysis, Writing – review & editing. ZG: Formal analysis, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

All authors were employed by the company Hainan Provincial Branch of China National Tobacco Corporation. The authors declare that this study received funding from Science and Technology Major Project of China National Tobacco Corporation (110202201038(XJ-09), 110202103017, 110202101012(XJ-04)). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1433575/full#supplementary-material

References

Akhtar, T. A., Surowiecki, P., Siekierska, H., Kania, M., Van Gelder, K., Rea, K. A., et al. (2017). Polyprenols are synthesized by a plastidial cis-prenyltransferase and influence photosynthetic performance. *Plant Cell* 29, 1709–1725. doi: 10.1105/tpc.16.00796

Anoman, A. D., Flores-Tornero, M., Rosa-Tellez, S., Munoz-Bertomeu, J., Segura, J., and Ros, R. (2016). The specific role of plastidial glycolysis in photosynthetic and heterotrophic cells under scrutiny through the study of glyceraldehyde-3-phosphate dehydrogenase. *Plant Signal Behav.* 11, e1128614. doi: 10.1080/15592324.2015.1128614

Borges, A., Morejón, R., Izquierdo, A., Ortega, E., and Rodés, R. (2012). Nitrogen fertilization for optimizing the quality and yield of shade grown Cuban cigar tobacco: required nitrogen amounts, application schedules, adequate leaf nitrogen levels, and early season diagnostic tests. *Beiträge zur Tabakforschung International/Contributions to Tobacco Res.* 25, 336–349. doi: 10.2478/cttr-2013-0913

Bukhov, N. G. (2004). Dynamic light regulation of photosynthesis (A review). *Russian J. Plant Physiol.* 51, 742–753. doi: 10.1023/B:RUPP.0000047822.66925.bf

Chen, X., Zhang, C., and Lindley, N. D. (2020). Metabolic engineering strategies for sustainable terpenoid flavor and fragrance synthesis. *J. Agric. Food Chem.* 68, 10252–10264. doi: 10.1021/acs.jafc.9b06203

Cheng, S. H., Moore, B., and Seemann, J. R. (1998). Effects of short- and long-term elevated CO2 on the expression of ribulose-1,5-bisphosphate carboxylase/oxygenase genes and carbohydrate accumulation in leaves of Arabidopsis thaliana (L.) Heynh. *Plant Physiol.* 116, 715–723. doi: 10.1104/pp.116.2.715

Christianson, D. W. (2017). Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* 117, 11570–11648. doi: 10.1021/acs.chemrev.7b00287

Daglia, M., Antiochia, R., Sobolev, A. P., and Mannina, L. (2014). Untargeted and targeted methodologies in the study of tea (Camellia sinensis L.). *Food Res. Int.* 63, 275–289. doi: 10.1016/j.foodres.2014.03.070

De Porcellinis, A. J., Norgaard, H., Brey, L. M. F., Erstad, S. M., Jones, P. R., Heazlewood, J. L., et al. (2018). Overexpression of bifunctional fructose-1,6-bisphosphatase/sedoheptulose-1,7-bisphosphatase leads to enhanced photosynthesis and global reprogramming of carbon metabolism in Synechococcus sp. PCC 7002. *Metab. Eng.* 47, 170–183. doi: 10.1016/j.ymben.2018.03.001

De Wit, M., Galvao, V. C., and Fankhauser, C. (2016). Light-mediated hormonal regulation of plant growth and development. *Annu. Rev. Plant Biol.* 67, 513–537. doi: 10.1146/annurev-arplant-043015-112252

Ding, H., Mo, S., Qian, Y., Yuan, G., Wu, X., and Ge, C. (2020). Integrated proteome and transcriptome analyses revealed key factors involved in tomato (Solanum lycopersicum) under high temperature stress. *Food Energy Secur.* 9, e239. doi: 10.1002/fes3.239

Dong, C., Qu, G., Guo, J., Wei, F., Gao, S., Sun, Z., et al. (2022). Rational design of geranylgeranyl diphosphate synthase enhances carotenoid production and improves photosynthetic efficiency in Nicotiana tabacum. *Sci. Bull. (Beijing)* 67, 315–327. doi: 10.1016/j.scib.2021.07.003

El Kadmiri, N., Slassi, I., El Moutawakil, B., Nadifi, S., Tadevosyan, A., Hachem, A., et al. (2014). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and Alzheimer's disease. *Pathol. Biol. (Paris)* 62, 333–336. doi: 10.1016/j.patbio.2014.08.002

Estevez, J. M., Cantero, A., Reindl, A., Reichler, S., and Leon, P. (2001). 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J. Biol. Chem.* 276, 22901–22909. doi: 10.1074/jbc.M100854200

Evans, J. R. (2013). Improving photosynthesis. *Plant Physiol.* 162, 1780–1793. doi: 10.1104/pp.113.219006

Feng, L., Raza, M. A., Li, Z., Chen, Y., Khalid, M. H. B., Du, J., et al. (2018). The influence of light intensity and leaf movement on photosynthesis characteristics and carbon balance of soybean. *Front. Plant Sci.* 9, 1952. doi: 10.3389/fpls.2018.01952

Gao, J., Liu, Z., Zhao, B., Liu, P., and Zhang, J. W. (2020). Physiological and comparative proteomic analysis provides new insights into the effects of shade stress in maize (Zea mays L.). *BMC Plant Biol.* 20, 60. doi: 10.1186/s12870-020-2264-2

Haak, D. C., Fukao, T., Grene, R., Hua, Z., Ivanov, R., Perrella, G., et al. (2017). Multilevel regulation of abiotic stress responses in plants. *Front. Plant Sci.* 8, 1564. doi: 10.3389/fpls.2017.01564

Hafidh, S., Potesil, D., Fila, J., Capkova, V., Zdrahal, Z., and Honys, D. (2016). Quantitative proteomics of the tobacco pollen tube secretome identifies novel pollen tube guidance proteins important for fertilization. *Genome Biol.* 17, 81. doi: 10.1186/s13059-016-0928-x

Haider, S., and Pal, R. (2013). Integrated analysis of transcriptomic and proteomic data. *Curr. Genomics* 14, 91–110. doi: 10.2174/1389202911314020003

Hoqani, U. A., Leon, R., and Purton, S. (2022). Over-expression of a cyanobacterial gene for 1-deoxy-d-xylulose-5-phosphate synthase in the chloroplast of Chlamydomonas reinhardtii perturbs chlorophyll: carotenoid ratios. *J. King Saud Univ. Sci.* 34, None. doi: 10.1016/j.jksus.2022.102141

Hwang, J. H., Yu, S. I., Lee, B. H., and Lee, D. H. (2020). Modulation of energy metabolism is important for low-oxygen stress adaptation in Brassicaceae species. *Int. J. Mol. Sci.* 21, 1787. doi: 10.3390/ijms21051787

Kossler, S., Armarego-Marriott, T., Tarkowska, D., Tureckova, V., Agrawal, S., Mi, J., et al. (2021). Lycopene beta-cyclase expression influences plant physiology, development, and metabolism in tobacco plants. *J. Exp. Bot.* 72, 2544–2569. doi: 10.1093/jxb/erab029

Lei, B., Lu, K., Ding, F., Zhang, K., Chen, Y., Zhao, H., et al. (2014). RNA sequencing analysis reveals transcriptomic variations in tobacco (Nicotiana tabacum) leaves affected by climate, soil, and tillage factors. *Int. J. Mol. Sci.* 15, 6137–6160. doi: 10.3390/ijms15046137

Li, W., Zhang, H., Li, X., Zhang, F., Liu, C., Du, Y., et al. (2017). Intergrative metabolomic and transcriptomic analyses unveil nutrient remobilization events in leaf senescence of tobacco. *Sci. Rep.* 7, 12126. doi: 10.1038/s41598-017-11615-0

Liang, C., Cheng, S., Zhang, Y., Sun, Y., Fernie, A. R., Kang, K., et al. (2016). Transcriptomic, proteomic and metabolic changes in Arabidopsis thaliana leaves after the onset of illumination. *BMC Plant Biol.* 16, 43. doi: 10.1186/s12870-016-0726-3

Liang, F., and Lindblad, P. (2016). Effects of overexpressing photosynthetic carbon flux control enzymes in the cyanobacterium Synechocystis PCC 6803. *Metab. Eng.* 38, 56–64. doi: 10.1016/j.ymben.2016.06.005

Liu, G. T., Ma, L., Duan, W., Wang, B. C., Li, J. H., Xu, H. G., et al. (2014). Differential proteomic analysis of grapevine leaves by iTRAQ reveals responses to heat stress and subsequent recovery. *BMC Plant Biol.* 14, 110. doi: 10.1186/1471-2229-14-110

Liu, P., Luo, J., Zheng, Q., Chen, Q., Zhai, N., Xu, S., et al. (2020). Integrating transcriptome and metabolome reveals molecular networks involved in genetic and environmental variation in tobacco. *DNA Res.* 27, dsaa006. doi: 10.1093/dnares/dsaa006

Liu, X., Xu, J., Zhu, J., Du, P., and Sun, A. (2019a). Combined transcriptome and proteome analysis of RpoS regulon reveals its role in spoilage potential of Pseudomonas fluorescens. *Front. Microbiol.* 10, 94. doi: 10.3389/fmicb.2019.00094

Liu, X., Yang, Q., Lu, Y., Li, Y., Li, T., Zhou, B., et al. (2019b). Effect of purslane (Portulaca oleracea L.) extract on anti-browning of fresh-cut potato slices during storage. *Food Chem.* 283, 445–453. doi: 10.1016/j.foodchem.2019.01.058

Liu, Y., Beyer, A., and Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell* 165, 535–550. doi: 10.1016/j.cell.2016.03.014

Lu, T., Yu, H., Li, Q., Chai, L., and Jiang, W. (2019). Improving plant growth and alleviating photosynthetic inhibition and oxidative stress from low-light stress with exogenous GR24 in tomato (Solanum lycopersicum L.) seedlings. *Front. Plant Sci.* 10, 490. doi: 10.3389/fpls.2019.00490

Ma, X., Ren, X., Gao, H., Wu, X., Chen, K., and Khan, R. (2023). Proteomic, transcriptomic, biochemical, and physio-anatomical analyses provide insights into energy metabolism, light usage, and photosynthesis in cigar tobacco under different light intensities. *Ind. Crops Prod.* 198, 116651. doi: 10.1016/j.indcrop.2023.116651

Miyagawa, Y., Tamoi, M., and Shigeoka, S. (2001). Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. *Nat. Biotechnol.* 19, 965–969. doi: 10.1038/nbt1001-965

Moreno, J. C., Mi, J., Agrawal, S., Kossler, S., Tureckova, V., Tarkowska, D., et al. (2020). Expression of a carotenogenic gene allows faster biomass production by redesigning plant architecture and improving photosynthetic efficiency in tobacco. *Plant J.* 103, 1967–1984. doi: 10.1111/tpj.14909

Moser, S., and Pichler, H. (2019). Identifying and engineering the ideal microbial terpenoid production host. *Appl. Microbiol. Biotechnol.* 103, 5501–5516. doi: 10.1007/s00253-019-09892-y

Portis, A. R. Jr., and Parry, M. A. (2007). Discoveries in Rubisco (Ribulose 1,5-bisphosphate carboxylase/oxygenase): a historical perspective. *Photosynth. Res.* 94, 121–143. doi: 10.1007/s11120-007-9225-6

Pu, X., Dong, X., Li, Q., Chen, Z., and Liu, L. (2021). An update on the function and regulation of methylerythritol phosphate and mevalonate pathways and their evolutionary dynamics. *J. Integr. Plant Biol.* 63, 1211–1226. doi: 10.1111/jipb.13076

Raines, C. A. (2003). The Calvin cycle revisited. *Photosynth. Res.* 75, 1–10. doi: 10.1023/A:1022421515027

Ribbenstedt, A., Ziarrusta, H., and Benskin, J. P. (2018). Development, characterization and comparisons of targeted and non-targeted metabolomics methods. *PloS One* 13, e0207082. doi: 10.1371/journal.pone.0207082

Ruiz-Sola, M. A., Coman, D., Beck, G., Barja, M. V., Colinas, M., Graf, A., et al. (2016). Arabidopsis GERANYLGERANYL DIPHOSPHATE SYNTHASE 11 is a hub isozyme required for the production of most photosynthesis-related isoprenoids. *New Phytol.* 209, 252–264. doi: 10.1111/nph.13580

Schuster, M., Gao, Y., Schottler, M. A., Bock, R., and Zoschke, R. (2020). Limited responsiveness of chloroplast gene expression during acclimation to high light in tobacco. *Plant Physiol.* 182, 424–435. doi: 10.1104/pp.19.00953

Sekhar, S., Panda, D., Kumar, J., Mohanty, N., Biswal, M., Baig, M. J., et al. (2019). Comparative transcriptome profiling of low light tolerant and sensitive rice varieties induced by low light stress at active tillering stage. *Sci. Rep.* 9, 5753. doi: 10.1038/s41598-019-42170-5

Siddiqui, M. H., Mukherjee, S., Kumar, R., Alansi, S., Shah, A. A., Kalaji, H. M., et al. (2022). Potassium and melatonin-mediated regulation of fructose-1,6-bisphosphatase (FBPase) and sedoheptulose-1,7- bisphosphatase (SBPase) activity improve photosynthetic efficiency, carbon assimilation and modulate glyoxalase system accompanying tolerance to cadmium stress in tomato seedlings. *Plant Physiol. Biochem.* 171, 49–65. doi: 10.1016/j.plaphy.2021.12.018

Smith, A. M., and Stitt, M. (2007). Coordination of carbon supply and plant growth. *Plant Cell Environ.* 30, 1126–1149. doi: 10.1111/j.1365-3040.2007.01708.x

Thalmann, M., and Santelia, D. (2017). Starch as a determinant of plant fitness under abiotic stress. *New Phytol.* 214, 943–951. doi: 10.1111/nph.14491

Tian, T., Qiao, G., Wen, Z., Deng, B., Qiu, Z., Hong, Y., et al. (2020). Comparative transcriptome analysis reveals the molecular regulation underlying the adaptive mechanism of cherry (Cerasus pseudocerasus Lindl.) to shelter covering. *BMC Plant Biol.* 20, 27. doi: 10.1186/s12870-019-2224-x

Tsaballa, A., Sarrou, E., Xanthopoulou, A., Tsaliki, E., Kissoudis, C., Karagiannis, E., et al. (2020). Comprehensive approaches reveal key transcripts and metabolites highlighting metabolic diversity among three oriental tobacco varieties. *Ind. Crops Prod.* 143, 111933. doi: 10.1016/j.indcrop.2019.111933

Uematsu, K., Suzuki, N., Iwamae, T., Inui, M., and Yukawa, H. (2012). Increased fructose 1,6-bisphosphate aldolase in plastids enhances growth and photosynthesis of tobacco plants. *J. Exp. Bot.* 63, 3001–3009. doi: 10.1093/jxb/ers004

Venkataramanan, K. P., Min, L., Hou, S., Jones, S. W., Ralston, M. T., Lee, K. H., et al. (2015). Complex and extensive post-transcriptional regulation revealed by integrative proteomic and transcriptomic analysis of metabolite stress response in Clostridium acetobutylicum. *Biotechnol. Biofuels* 8, 81. doi: 10.1186/s13068-015-0260-9

Walter, M. H., and Strack, D. (2011). Carotenoids and their cleavage products: biosynthesis and functions. *Nat. Prod. Rep.* 28, 663–692. doi: 10.1039/c0np00036a

Wang, L., Czedik-Eysenberg, A., Mertz, R. A., Si, Y., Tohge, T., Nunes-Nesi, A., et al. (2014). Comparative analyses of C(4) and C(3) photosynthesis in developing leaves of maize and rice. *Nat. Biotechnol.* 32, 1158–1165. doi: 10.1038/nbt.3019

Wang, Q., Huang, X. Q., Cao, T. J., Zhuang, Z., Wang, R., and Lu, S. (2018). Heteromeric Geranylgeranyl Diphosphate Synthase Contributes to Carotenoid Biosynthesis in Ripening Fruits of Red Pepper (Capsicum annuum var. conoides). *J. Agric. Food Chem.* 66, 11691–11700. doi: 10.1021/acs.jafc.8b04052

Wei, S., Li, X., Lu, Z., Zhang, H., Ye, X., Zhou, Y., et al. (2022). A transcriptional regulator that boosts grain yields and shortens the growth duration of rice. *Science* 377, eabi8455. doi: 10.1126/science.abi8455

Wu, S., Cao, G., Adil, M. F., Tu, Y., Wang, W., Cai, B., et al. (2020). Changes in water loss and cell wall metabolism during postharvest withering of tobacco (*Nicotiana tabacum* L.) leaves using tandem mass tag-based quantitative proteomics approach. *Plant Physiol. Biochem.* 150, 121–132. doi: 10.1016/j.plaphy.2020.02.040

Wu, X., Khan, R., Gao, H., Liu, H., Zhang, J., and Ma, X. (2021). Low light alters the photosynthesis process in cigar tobacco via modulation of the chlorophyll content, chlorophyll fluorescence, and gene expression. *Agriculture* 11, 755. doi: 10.3390/agriculture11080755

Xiang, S., Usunow, G., Lange, G., Busch, M., and Tong, L. (2007). Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *J. Biol. Chem.* 282, 2676–2682. doi: 10.1074/jbc.M610235200

Yang, H., Dong, B., Wang, Y., Qiao, Y., Shi, C., Jin, L., et al. (2020). Photosynthetic base of reduced grain yield by shading stress during the early reproductive stage of two wheat cultivars. *Sci. Rep.* 10, 14353. doi: 10.1038/s41598-020-71268-4

Yu, A., Xie, Y., Pan, X., Zhang, H., Cao, P., Su, X., et al. (2020). Photosynthetic phosphoribulokinase structures: enzymatic mechanisms and the redox regulation of the Calvin-Benson-Bassham cycle. *Plant Cell* 32, 1556–1573. doi: 10.1105/tpc.19.00642

Zapalska-Sozoniuk, M., Chrobak, L., Kowalczyk, K., and Kankofer, M. (2019). Is it useful to use several "omics" for obtaining valuable results? *Mol. Biol. Rep.* 46, 3597–3606. doi: 10.1007/s11033-019-04793-9

Zeng, Y., Tan, X., Zhang, L., Jiang, N., and Cao, H. (2014). Identification and expression of fructose-1,6-bisphosphate aldolase genes and their relations to oil content in developing seeds of tea oil tree (Camellia oleifera). *PloS One* 9, e107422. doi: 10.1371/journal.pone.0107422

Zhang, H., Liu, T., Zhang, Z., Payne, S. H., Zhang, B., Mcdermott, J. E., et al. (2016). Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell* 166, 755–765. doi: 10.1016/j.cell.2016.05.069

Zhang, T., Sun, M., Guo, Y., Shi, X., Yang, Y., Chen, J., et al. (2018). Overexpression of LiDXS and LiDXR from lily (Lilium 'Siberia') enhances the terpenoid content in tobacco flowers. *Front. Plant Sci.* 9, 909. doi: 10.3389/fpls.2018.00909

Zheng, P. C., Qin, C. Y., Liu, P. P., Feng, L., Ling, T. J., Ning, J. M., et al. (2021). Untargeted metabolomics combined with bioassay reveals the change in critical bioactive compounds during the processing of Qingzhuan tea. *Molecules* 26, 6718. doi: 10.3390/molecules26216718

Zhou, F., Wang, C. Y., Gutensohn, M., Jiang, L., Zhang, P., Zhang, D., et al. (2017). A recruiting protein of geranylgeranyl diphosphate synthase controls metabolic flux toward chlorophyll biosynthesis in rice. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6866–6871. doi: 10.1073/pnas.1705689114

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 25 July 2024
DOI 10.3389/fpls.2024.1425759

Diterpenoids target SARS-CoV-2 RdRp from the roots of *Euphorbia fischeriana* Steud

Ting Ruan[1,2,3,4†], Zheng-Rui Xiang[1,3,4†], Yun-Wu Zhang[1,3,5], Shi-Rui Fan[1,3,4], Juan Ren[1,5], Qian Zhao[1,4], Xiao-Long Sun[1,6], Shi-Li Wu[1,4], Li-Li Xu[1,4], Miao Qiao[1,4], Chen-Xu Jing[7], Xiao-Jiang Hao[1,2*] and Duo-Zhi Chen[1,2*]

[1]State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China, [2]Yunnan Characteristic Plant Extraction Laboratory, Kunming, China, [3]Research Unit of Chemical Biology of Natural Anti-Virus Products, Chinese Academy of Medical Sciences, Beijing, China, [4]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China, [5]Department of Chemical Science and Engineering, Yunnan University, Kunming, China, [6]Institute of International Rivers and Eco-Security, Yunnan University, Kunming, China, [7]Research Center of Traditional Chinese Medicine, The Affiliated Hospital to Changchun University of Chinese Medicine, Changchun, China

OPEN ACCESS

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Pawan Kumar,
Agricultural Research Organization (ARO), Israel
Xian-Wen Yang,
State Oceanic Administration, China
Li-Dong Shao,
Yunnan University of Chinese Medicine, China

*CORRESPONDENCE
Duo-Zhi Chen
✉ chenduozhi@mail.kib.ac.cn
Xiao-Jiang Hao
✉ haoxj@mail.kib.ac.cn

[†]These authors have contributed equally to this work

RECEIVED 30 April 2024
ACCEPTED 01 July 2024
PUBLISHED 25 July 2024

CITATION
Ruan T, Xiang Z-R, Zhang Y-W, Fan S-R, Ren J, Zhao Q, Sun X-L, Wu S-L, Xu L-L, Qiao M, Jing C-X, Hao X-J and Chen D-Z (2024) Diterpenoids target SARS-CoV-2 RdRp from the roots of *Euphorbia fischeriana* Steud. *Front. Plant Sci.* 15:1425759.
doi: 10.3389/fpls.2024.1425759

Introduction: Currently, the development of new antiviral drugs against COVID-19 remains of significant importance. In traditional Chinese medicine, the herb *Euphorbia fischeriana* Steud is often used for antiviral treatment, yet its therapeutic effect against the COVID-19 has been scarcely studied. Therefore, this study focuses on the roots of *E. fischeriana* Steud, exploring its chemical composition, antiviral activity against COVID-19, and the underlying basis of its antiviral activity.

Methods: Isolation and purification of phytochemicals from *E. fischeriana* Steud. The elucidation of their configurations was achieved through a comprehensive suite of 1D and 2D NMR spectroscopic analyses as well as X-ray diffraction. Performed cytopathic effect assays of SARS-CoV-2 using Vero E6 cells. Used molecular docking to screen for small molecule ligands with binding to SARS-CoV-2 RdRp. Microscale thermophoresis (MST) was used to determine the dissociation constant Kd.

Results: Ultimately, nine new ent-atisane-type diterpenoid compounds were isolated from *E. fischeriana* Steud, named Eupfisenoids A-I (compounds 1-9). The compound of 1 was established as a C-19-degraded ent-atisane-type diterpenoid. During the evaluation of these compounds for their antiviral activity against COVID-19, compound 1 exhibited significant antiviral activity. Furthermore, with the aid of computer virtual screening and microscale thermophoresis (MST) technology, it was found that this compound could directly bind to the RNA-dependent RNA polymerase (RdRp, NSP12) of the COVID-19, a key enzyme in virus replication. This suggests that the compound inhibits virus replication by targeting RdRp.

Discussion: Through this research, not only has our understanding of the antiviral components and material basis of *E. fischeriana* Steud been enriched, but also the potential of atisane-type diterpenoid compounds as antiviral agents against COVID-19 has been discovered. The findings mentioned above will provide valuable insights for the development of drugs against COVID-19.

KEYWORDS

Euphorbia fischeriana **Steud, diterpenoids, antiviral, SARS-CoV-2 RdRp, microscale thermophoresis**

1 Introduction

In late 2019, viral infections spurred by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) profoundly impacted global human existence and productivity (Hu et al., 2021). As of early March 2024 (World Health Organization, 2024), the World Health Organization (WHO) documented 775 million confirmed infections and 7.04 million deaths worldwide. Despite accessible medications and ongoing vaccination drives, viral mutations and waning immune responses have hampered vaccine efficacy, leading to re-infections (Sievers et al., 2022). SARS-CoV-2, an optimistic single-stranded RNA virus (Gorbalenya et al., 2020), employs RNA-dependent RNA polymerase (RdRp) for genome replication and gene transcription (Snijder et al., 2016). RdRp, as a pivotal enzyme in the viral life cycle, stands as a recognized target for antiviral medications (te Velthuis, 2014). Human life and productivity remain imperiled, and the development of targeted COVID-19 medications is urgently needed (Feikin et al., 2022). Natural products have long been esteemed as pivotal sources of medications for diverse ailments. *Euphorbia* is the largest genus of the Euphorbiaceae, one of the largest families of higher plants (Li et al., 2009). *Euphorbia fischeriana* Steud, a member of the genus *Euphorbia* (Euphorbiaceae), is a perennial herb, and the entire plant is poisonous and is grown primarily in Mongolia, Eastern Siberia, and China (Flora of China Editorial Committee, 1997). *E. fischeriana* Steud serves prominently in traditional Chinese medicine, the roots of which has a long history of traditional use for treating conditions such as edema, ascites, cough, and cancer, and also exhibits prowess in combating viral infections (Pharmacopoeia Commission of PRC, 2020). Its diterpenoids possess a spectrum of pharmacological activities, including antiviral, anti-tumor, antibacterial, anti-inflammatory, and other therapeutic properties (Li et al., 2021) (Figure 1). In traditional Chinese medicine, *E. fischeriana* Steud is an important antiviral herb. During the outbreak of the COVID-19 pandemic, efforts have been made in some regions of China to utilize traditional Chinese herbal remedies, including E. fischeriana Steud, to treat COVID-19 infections, with some degree of success (Dai et al., 2020). Through literature review, it was found that the material basis of *E. fischeriana* Steud's anti-COVID-19 activity, specifically the antiviral activity of its natural compounds and related mechanisms, has not been studied. Therefore, we selected *E. fischeriana* Steud as the research subject to investigate its chemical composition. We employed the latest techniques in pharmaceutical chemistry research, such as molecular docking and microscale thermophoresis (MST), to evaluate its potential anti-SARS-CoV-2 effects and related targets. Through *in vitro* cellular experiments, we identified *ent*-atisane-type diterpenoid compounds and potential targets among the anti-COVID-19 active herbal constituents, providing new insights for the development of COVID-19 therapeutic drugs. We have successfully isolated and identified nine diterpenoids with a novel structure from the roots of *E. fischeriana* Steud (Figure 2); in addition, we evaluated the anti-SARS-CoV-2 activity of the compounds and explored the targets and mechanisms of action.

2 Results

2.1 Isolated phytochemicals from *Euphorbia fischeriana* Steud

Compound 1, white crystals, has the molecular formula $C_{22}H_{34}O_4$ as determined from a high-resolution electrospray ionization mass spectrometry (HR-ESI-MS) peak at *m/z* 363.2529 $[M+H]^+$ (calcd. $C_{23}H_{35}O_4$, 363.2530). The molecular formula indicates six indices of hydrogen deficiency. The ^{13}C NMR and DEPT (Tables 1, 2) spectra of 1 displayed 22 carbon signals and resonances attributable to four methyls, eight methylenes, five methines (one oxygenated methine), and five quaternary carbons (one carbonyl and two oxygenated carbons), corresponding to the units in its ^1H NMR data (Table 1). The ^1H-^1H COSY spectrum of this substance shows the existence of correlations of H-1/H-2 and H-4/H-5. Furthermore, analysis of the HMBC spectrum uncovered correlations of H-1/C-2, C-3 (δ_C 213.8), C-5 (δ_C 55.5), C-10, and H-4/C-10, H-5/C-3, showing the presence of an A ring with the carbonyl group at C-3. The ^1H-^1H COSY spectrum showed a

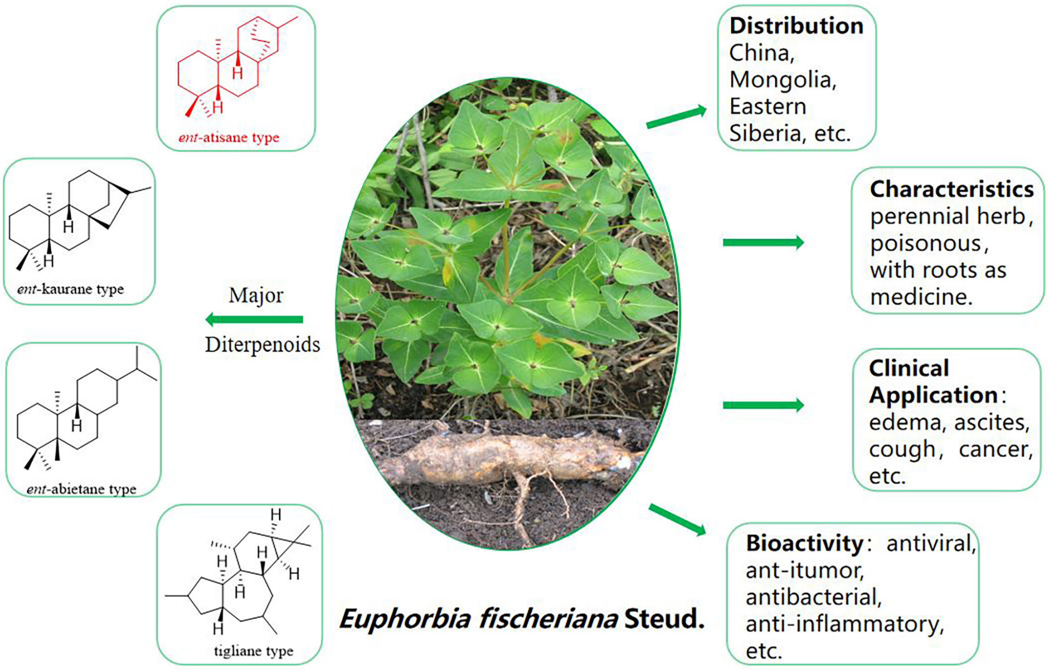


FIGURE 1
Euphorbia fischeriana Steud plant pictures and other information.

correlation between CH_3-18/H-4/H-5, and the HMBC spectrum showed a correlation between H-18/C-3, C-4, and C-5, H-1, H-5/C-20, indicating that C-18 (δ_C 11.7) and C-20 (δ_C 14.8) have a methyl group; C-20 is a horn methyl group. Compound 1 was one methyl less than *ent*-3-oxoatisan-16*α*, 17-acetonide (Yan et al., 2018), suggesting the absence of CH_3-19 in 1. The ^1H-^1H COSY (Figure 3B) spectrum showed the correlation of H-9/H-11/H-12 and H-12/H-13/H-14, and the HMBC spectrum showed the correlation of H-11/C-16, C-10, which, combined with the HSQC (Tables 1, 2) (δ_C 67.4, C-11, δ_H 4.66, H-11), indicated an additional

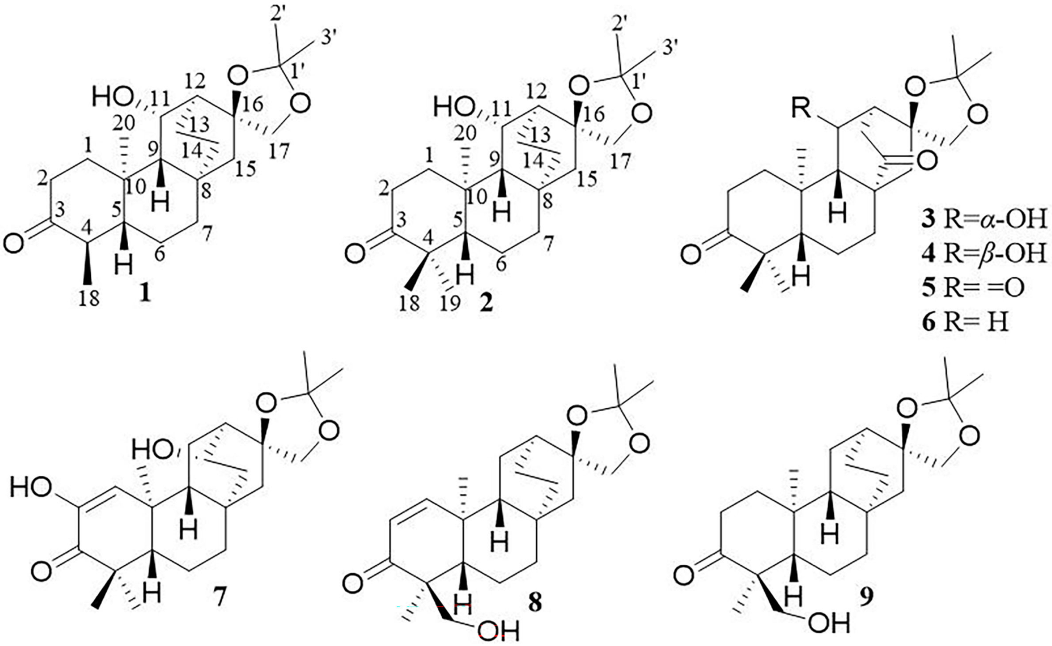


FIGURE 2
Structures of compounds 1–9.

TABLE 1 ^{1}H NMR spectroscopic data (δ) for compounds 1–4.

Position	1^a	2^b	3^a	4^a
1a	2.72, ddd (13.0,6.4, 2.3)	1.93, ddd (13.3, 6.8, 3.1)	2.67, m	2.66, m
1b	1.49, dd(13.9, 10.5)	1.41, m	1.56, m	1.65, m
2a	2.52, td(14.4, 6.4)	2.67, ddd (16.0, 12.5, 6.9)	2.32, ddt (6.6, 3.4)	2.62, m
2b	2.29, m	2.30, ddd (16.0, 5.9, 3.1)	2.19, m	2.34, m
3				
4	2.29, m			
5	1.06, m	1.33, m	1.23, dd(12.4, 2.3)	1.29, dd(12.4, 2.2)
6a	1.49, s	1.76, m	1.62, m	1.62, m
6b	1.31, m	1.30, m	1.43, ddd (11.2, 4.6, 2.4)	1.43, m
7a	1.36, s	1.46, m	2.30, m	2.27, dddd (32.7, 8.6, 5.6, 2.0)
7b	1.10, m	1.31, m	0.75, m	0.84, td(13.5, 4.9)
8				
9	1.39, d	1.76, m	1.60, m	1.92, d(9.8)
10				
11a	4.66, m	3.91, m	4.50, d(5.6, 4.4)	4.95, dd(5.8, 3.9)
11b				
12	1.68, d(2.8)	1.77, m	2.19, m	2.18, m
13a	2.07, m	1.50, m	2.72, m	2.88, dd(19.5, 2.9)
13b	1.36, m	1.24, m	2.59, m	2.01, d(19.5)
14a	2.05, m	1.80, m		
14b	2.00, m	1.32, m		
15a	1.47, s	1.45, m	1.83, d(15.0)	1.74, d(15.1)
15b		1.32, m	1.62, m	
16				
17a	3.98, d(8.4)	4.02, d(8.6)	4.08, d(8.6)	3.93, d(8.7)
17b	3.66, d(8.4)	3.61, d(8.6)	3.69, d(8.6)	3.61, d(8.7)
18	0.98, d(6.5)	1.05, s	0.99, s	1.01, s
19		1.07, s	1.04, s	1.05, s
20	1.41, s	1.12, s	1.13, s	1.14, s
1’				
2’	0.85, td(12.3, 7.2)	1.35, s	1.33, s	1.35, s
3’	1.34, s	1.35, s	1.34, s	1.38, s

[a] Chemical shifts (ppm) referenced to solvent peak (δ_H7.26 in $CDCl_3$) at 500 MHz.

[b] Chemical shifts (ppm) referenced to solvent peak (δ_H 3.31 in methanol-d_4) at 500 MHz.

hydroxyl substitution at C-11. The correlation of H-17/C-12, C-15, C-16, C-1′ in the HMBC spectrum (δ_C 109.0) demonstrated that C-16 (δ_C 81.7) and C-17 (δ_C 74.7) form an acetone dimethyl acetal. The ROESY (Figure 3C) spectrum of 1 shows the existence of correlations of H-5/H9, H-5/H-11, H-12/H$_2$-17, and H-12/H-18, indicating that H-5, H-9, and H-11 are on the same faces of rings B and C. Fortunately, compound 1 crystallizes from methanol and acetone, yielding crystals that are suitable for x-ray diffraction [Cu

TABLE 2 ^{13}CNMR spectroscopic data (δ) for compounds 1–9.

Position	1[a]	2[b]	3[a]	4[a]	5[a]	6[a]	7[a]	8[a]	9[a]
1	40.4	39.5	38.6	38.5	38.2	37.2	128.0	125.4	37.8
2	37.6	35.1	34.4	34.4	34.0	34.1	143.4	160.1	27.6
3	213.8	219.9	217.1	217.0	215.7	216.3	201.0	205.9	219.0
4	44.6	38.5	47.8	47.8	47.8	47.6	43.9	49.0	52.5
5	55.5	56.8	56.7	56.5	55.0	55.4	53.7	47.2	49.3
6	22.3	16.8	20.1	20.2	19.6	19.9	18.9	18.7	19.2
7	38.2	20.8	31.0	31.0	30.9	31.3	39.3	38.7	23.4
8	34.2	48.8	48.7	48.9	49.4	47.8	34.4	33.8	37.0
9	53.3	51.1	56.7	55.7	64.9	51.1	46.0	45.4	50.2
10	37.8	35.0	38.9	38.9	39.3	37.6	39.0	40.1	33.2
11	67.4	67.6	68.2	66.6	210.8	23.8	15.6	23.1	23.2
12	43.8	43.9	47.2	46.7	55.9	36.9	42.9	34.1	34.1
13	16.0	40.5	34.3	35.2	35.1	40.3	67.2	23.3	35.0
14	27.0	39.1	215.9	215.6	212.1	216.2	39.0	28.0	38.7
15	54.5	53.2	50.1	49.2	48.2	50.1	51.8	53.6	53.8
16	81.7	82.6	79.6	80.2	78.3	81.0	81.0	82.5	82.6
17	74.7	74.6	73.3	73.8	73.9	74.1	73.2	74.2	74.4
18	11.7	22.2	21.7	21.8	21.5	21.8	26.9	16.8	16.8
19		26.5	26.6	26.5	26.3	26.0	21.9	67.9	66.8
20	14.8	15.5	15.6	15.6	14.9	12.7	19.7	18.5	13.9
1'	109.0	109.8	109.4	109.8	110.3	109.5	109.0	109.0	108.8
2'	27.1	27.2	27.2	26.6	26.3	26.9	26.8	27.0	27.1
3'	27.3	27.8	26.8	27.3	27.3	27.2	27.6	27.6	27.5

[a] Chemical shifts (ppm) referenced to solvent peak (δ_C77.16 in $CDCl_3$) at 126 MHz.

[b] Chemical shifts (ppm) referenced to solvent peak (δ_C49.00 in methanol-d_4) at 121 MHz.

Kα radiation, Flack parameter 0.01(11)] (Figure 3A). Analysis of the structure shows that the absolute configurations at the stereogenic centers in compound 1 is 4*R*,5*S*,8*S*,9*S*,10*R*,11*S*,12*S*,16*S*. Compound 1 was named Eupfisenoid A, and the structure was assigned to be (4*R*,5*S*,8*S*,9*S*,10*R*,11*S*,12*S*,16*S*)-*ent*-11*α*-hydroxy-16*α*,17-acetonide-19-noratisan-3-one.

Compound 2, white amorphous powder, has the molecular formula $C_{23}H_{36}O_4$ as determined from an HR-ESI-MS peak at *m/z* 399.2512 $[M+Na]^+$ (calcd. $C_{23}H_{36}NaO_4$, 399.2506). The molecular formula indicates six indices of hydrogen deficiency. The ^{13}C NMR and DEPT (Tables 1, 2) spectra of 2 displayed 23 carbon signals and resonances attributable to five methyls, eight methylenes, four methines, and five quaternary carbons (one carbonyl and two oxygenated carbons). Comparison of literature data analysis revealed that compound 2 is structurally similar to the known compound *ent*-3-oxoatisan-16*α*,17-acetonide (Yan et al., 2018). The OH moiety was attached to C-11 according to the ^{1}H-^{1}H COSY (Figure 3B) correlations of H-9/H/-11/H-12/H-13/H-14 and HMBC spectrum (Figure 3B) H-11/C-8 and C-16, combined with HSQC (Tables 1, 2) (δ_C 67.6, C-11, δ_H 3.91, H-11). The relative configuration of compound 2 was determined by the ROESY spectrum, and the ROESY spectrum (Figure 3C) showed a correlation between H-12/H_2-17, H-5/H-1a, and H-1a/H-11, indicating the *β*-configuration of H-11, suggesting the *α*-orientation of 11-OH. Compound 2 differs from compound 1 in that methyl is added at C-19. Compound 2 showed similar ROESY correlation signals C-12 and C-17, indicating that they share the same relative configuration at C-17. Compound 2 was named Eupfisenoid B, and the structure was assigned to be *ent*-11*α*-hydroxyatisan-16*α*,17-acetonide-3-one (Yan et al., 2018).

Compound 3, colorless crystals, has the molecular formula $C_{23}H_{34}O_5$ as determined from an HR-ESI-MS peak at *m/z* 391.2486 $[M+H]^+$ (calcd. $C_{23}H_{35}O_5$,391.2479). The molecular formula indicates seven indices of hydrogen deficiency. The ^{13}C NMR and DEPT (Tables 1, 2) spectra of 3 displayed 23 carbon signals and resonances attributable to five methyls, seven methylenes, four methines, and seven quaternary carbons (two carbonyls and two oxygenated carbons). Analysis of the ^{1}H and ^{13}C NMR data (Tables 1, 2) indicated that the structure of 3 was similar to that of 2. The difference is that compound **3** has another

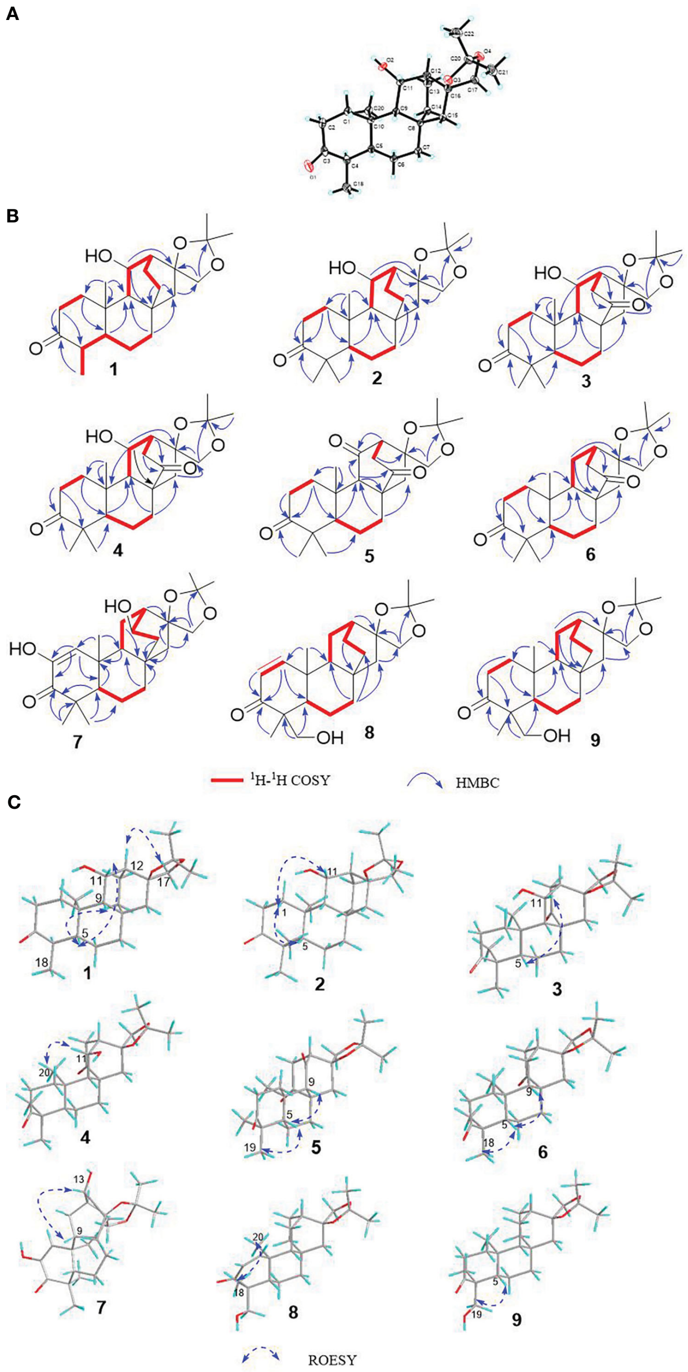


FIGURE 3
(A) Single-crystal x-ray figure structure of compound **1**; **(B, C)** key 2D NMR correlations of compounds 1–9.

carbonyl. The ^1H-^1H COSY spectrum of compound **3** (Figure 3B) H-9/H-11/H-12/H-13 and HMBC correlation between H-7, H-12, H-14, and H-15/C-14, combined with HSQC (Tables 1, 2) (δ_C 68.2, C-11, δ_H 4.50, H-11; δ_C 215.9, C-14), revealed a carbonyl group at C-14. The relative configuration of compound **3** was determined by the ROESY spectrum, showing the correlation of H-12/H_2-17 and H-5/H-11, indicating the β-configuration of H-11, and the α-orientation of 11-OH can be concluded. Compound 3 showed similar ROESY correlation signals, δ values (δ_C 73.3), and *J* values (δ_H 4.08, d, *J* = 8.6 Hz, H-17a, δ_H 3.69, d, *J* = 8.6 Hz, H-17b) for key protons to compound 1 at C-17, indicating that they share the same relative configuration at C-16 and C-17. Compound 3 was named Eupfisenoid C, and the structure was assigned to be *ent*-11α-hydroxyatisan-16α,17-acetonide-3,14-dione.

Compound 4, white crystals, has the molecular formula $C_{23}H_{34}O_5$ as determined from an HR-ESI-MS peak at *m/z* 391.2477 $[M+H]^+$ (calcd. $C_{23}H_{35}O_5$, 391.2477). Analysis of the ^1H and ^{13}C NMR data (Tables 1, 2) indicated that the structure of 4 was similar to that of 3. The ^1H-^1H COSY spectra (Figure 3B) show correlations for H-9/H-11/H-12/H-13, and the HMBC spectra show (Figure 3B) correlations for H-11/C-8, C-16, which, in combination with the HSQC (Figure 3B) (δ_C 66.6, C-11, δ_H 4.95, *J* = 5.8, 3.9 Hz, H-11; δ_C 52.3, C-9, δ_H 1.92, d, *J* = 9.8 Hz, H-9), like compound **3**, shows that compound 4 has a hydroxyl group at C-11; the only difference is the configuration of the hydroxyl group. The ROESY spectrum H-12/H_2-17, H-11/H_3-20, assigned the *α*-configuration of H-11, suggests the *β*-orientation of 11-OH. Compound 4 showed similar ROESY correlation signals, δ values (δ_C 73.8), and *J* values (δ_H 3.93, d, *J* = 8.7 Hz, H-17a, δ_H 3.61, d, *J* = 8.7 Hz, H-17b) for key protons to compound 1 at C-17, indicating that they share the same relative configuration at C-16 and C-17. The compound was named Eupfisenoid D, and the structure was assigned to be *ent*-11*β*-hydroxyatisan-16*α*,17-acetonide-3,14-dione.

Compound 5, white powder, has a molecular formula of $C_{23}H_{32}O_5$ according to an HR-ESI-MS peak at *m/z* 433.2229 $[M+COOH]^-$ (calcd. $C_{24}H_{33}O_7$, 433.2232). The ^{13}C NMR and DEPT (Tables 2, 3) spectra of 5 displayed 23 carbon signals, and analysis of the ^1H and ^{13}C NMR data (Tables 1–3) shows that compound 5 and compound 4 are structurally similar. They differ by a decrease in the hydroxyl group and an increase in the carbonyl group of compound 5. The location of the introduced carbonyl group was attached to C-2 by the correlation of H-12/H-13 in the ^1H-^1H COSY spectrum (Figure 3B) and H-9, H-12/C-11 (δ_C 210.8) in the HMBC spectrum (Figure 3B), in combination with the HSQC (Tables 2, 3) (δ_C 64.9, C-9, δ_H 1.88, H-9; δ_C 55.9, C-12, δ_H 2.80, H-12). Compound **5** showed similar ROESY (Figure 3C) correlation signals, δ values (δ_C 73.9), and *J* values (δ_H 3.93, d, *J* = 8.8 Hz, H-17a, δ_H 3.64, d, *J* = 8.8 Hz, H-17b) for key protons to compound 1 at C-17, indicating that they share the same relative configuration at C-16 and C-17. The compound was named Eupfisenoid E, and the structure was assigned to be *ent*-atisan-16*α*,17-acetonide-3,14-dione.

Compound 6, colorless crystals, has the molecular formula $C_{23}H_{34}O_4$ as determined from an HR-ESI-MS peak at *m/z* 373.2382 $[M-H]^-$ (calcd. $C_{23}H_{33}O_4$, 373.2384). The ^{13}C NMR and DEPT (Tables 2, 3) spectra of 6 displayed 23 carbon signals, and the analysis of the ^1H and ^{13}C NMR data (Tables 2, 3) indicated that the structure of 6 was similar to that of 5. In the ^1H-^1H COSY spectrum (Figure 3B), there is a correlation between H-9/H-11/H-12/H-13, and in combination with the HSQC (Tables 2, 3) (δ_C 23.8, C-11, δ_H 1.79, 1.62, H-11; δ_C 36.9, C-12, δ_H 2.20, H-12), compound 6 has one less carbonyl group at C-11 compared to compound 5. Compound 6 showed similar ROESY (Figure 3C) correlation signals, δ values (δ_C 74.1), and *J* values (δ_H 4.06, d, *J* = 8.5 Hz, H-17a, δ_H 3.75, d, *J* = 8.5 Hz, H-17b) for key protons to compound 1 at C-17, indicating that they share the same relative configuration at C-16 and C-17. Compound 6 was named Eupfisenoid F, and the structure was assigned to be *ent*-atisan-16*α*,17-acetonide-3-one.

Compound 7, colorless needle crystals, has the molecular formula $C_{22}H_{32}O_5$ as determined from an HR-ESI-MS peak at *m/z* 435.2387 $[M+COOH]^-$ (calcd. $C_{23}H_{35}O_7$ 435.2388). The ^{13}C NMR and DEPT (Tables 2, 3) spectra of **7** displayed 23 carbon signals. By comparing literature data, it was determined that compound **7** has two more hydroxyl groups compared to the *ent*-3-oxoatis-1-en-16*α*,17-acetonide (Huang, 2021). ^1H-^1H COSY spectra (Figure 3B) showed a correlation of H-9/H-11/H-12/13, which, in combination with HMBC spectra (Figure 3B), shows a correlation between H-1/C-2/C-3, C-5, and C-20, indicating a hydroxyl group at the C-2 position of the double bond. From the HSQC (Tables 2, 3) (δ_C 67.2, C-13, δ_H 3.96, H-13; δ_C 46.0, C-9, δ_H 1.60, H-9), combined with the ^1H-^1H COSY spectra, we can determine another hydroxyl group at C-13. The relative configuration of compound 7 can be inferred by analyzing the ROESY correlation signals. According to the ROESY spectrum (Figure 3C), the correlation signals of H-12/H_2-17 and H-13/H-9 indicate the *β*-configuration of H-13, suggesting the *α*-orientation of 13-OH. Compound **7** showed similar ROESY correlation signals, δ values (δ_C 73.2), and *J* values (δ_H 4.02, d, *J* = 8.5 Hz, H-17a, δ_H 3.57, d, *J* = 8.5 Hz, H-17b) for key protons to compound 1 at C-17, indicating that they share the same relative configuration at C-16 and C-17. Compound 7 was named Eupfisenoid G, and the structure was assigned to be *ent*-2,13*α*-dihydroxyatisan-1-en-16*α*,17-acetonide-3-one.

Compound 8, white powder, has the molecular formula $C_{23}H_{34}O_4$ as determined from an HR-ESI-MS peak at *m/z* 375.2536 $[M+H]^+$ (calcd. $C_{23}H_{35}O_4$, 375.2530). The ^{13}C NMR and DEPT (Tables 2, 3) spectra of 8 displayed 23 carbon signals and resonances attributable to four methyls, eight methylenes (one oxygenated methylene), five methines, and six quaternary carbons (one carbonyl and two oxygenated carbons). Comparison of literature data analysis revealed that compound 8 is structurally similar to the known compound *ent*-atisane-16*β*,17-isopropylidenedioxy-19-ol-3-one (Kuang et al., 2017), the ^1H-^1H COSY (Figure 3B) spectra of H-1/H-2 correlation exists, and H-1/C-3, C-5, H-2/C-3, H2-19/C-4, C-5 correlation exists in the HMBC spectra (Figure 3B); combined with the HSQC (Tables 2, 3) (δ_C 47.2, C-5, δ_H 1.91, H-5; δ_C 67.9, C-19, δ_H 3.71, d, *J* = 10.7 Hz, 3.44, d, *J* = 10.9 Hz, H-19) and ^{13}C NMR spectra and DEPT spectra (Tables 2, 3) (δ_C 125.4, C-1; δ_C 160.1, C-2), an extra double bond at C-1 and C-2 can be seen. The relative configuration of compound 8 was determined by the ROESY spectrum, and the ROESY spectrum (Figure 3C) showed a correlation of H-12/H_2-17 and H_3-18/H_3-20, which shows the *α*-configuration of H_3-18, and the *β*-orientation of H_2-19 can be concluded. Compound 8 showed similar ROESY correlation signals, δ values (δ_C 74.2), and *J* values (δ_H 3.99, d, *J* = 8.4 Hz, H-17a, δ_H 3.62, d, *J* = 8.4 Hz, H-17b) for key protons to compound 1 at C-17, indicating that they share the same relative configuration at C-16 and C-17. Compound 8 was named Eupfisenoid H, and the structure was assigned to be *ent*-atisan-1-en-16*α*,17-acetonide-19-hydroxy-3-one.

Compound 9, white powder, has the molecular formula $C_{23}H_{34}O_4$ based on its HR-ESI-MS peak at *m/z* 399.2510 $[M+Na]^+$ (calcd. $C_{23}H_{36}NaO_4$, 399.2506). The ^{13}C NMR and DEPT (Tables 2, 3) spectra of 9 displayed 23 carbon signals. Analysis of the ^1H and ^{13}C NMR data shows differences between compounds 8 and 9 and the no double-bond carbon substitution at C-1 and C-2 of 9. The relative configuration was determined by the ROESY spectrum, and the ROESY spectrum (Figure 3C) showed a correlation between H_3-18/H_3-20, establishing the *α*-configuration of H_3-18 and suggesting the *β*-orientation of H_2-19. Compound 9 showed

TABLE 3 ^{1}H NMR spectroscopic data (δ) for compounds 5–9.

Position	5[a]	6[a]	7[a]	8[a]	9[a]
1a	2.81, m	1.81, m	6.31, s	5.83, d(10.1)	1.86, m
1b	1.68, d(4.4)	1.34, s			
2a	2.52, m	2.65, ddd (15.8, 13.3, 6.4)		7.02, d(10.1)	1.87, m
2b	2.40, m	2.31, m			0.81, m
3					
4					
5	1.34, m	1.24, dd (12.4, 2.5)	1.63, dt, (18.8, 8.4)	1.91, m	1.64, dd (11.9, 2.0)
6a	1.65, d(3.5)	1.63, m	1.49, m	1.53, m	1.49, m
6b	1.53, dd (13.5, 2.8)	1.48, ddd (17.9, 7.9, 4.5)		1.45, m	1.35, m
7a	2.39, m	2.31, m	1.52, m	1.46, m	1.50, m
7b	0.97, m	0.82, m	1.27, d, (13.9)	1.27, m	
8					
9	1.87, t(7.6)	1.43, m	1.6, m	1.64, dd(10.6, 7.5)	1.36, m
10					
11a		1.79, m	1.97, dd, (17.6, 6.7)	2.20, t(12.0)	2.01, dd (13.1, 11.6)
11b		1.62, m	1.87, d	1.41, m	1.23, m
12	2.80, m	2.20, m	1.90, t, (9.9)	1.81, m	1.74, m
13a	3.00, dd (19.5, 2.6)	2.76, dt (19.1, 2.8)	3.97, d, (8.9)	1.53, m	2.64, ddd (16.0, 13.9, 6.5)
13b	2.43, m	2.16, td (6.2, 2.9)			2.27, ddd (16.0,5.0,2.5)
14a			1.77, d, (14.1)	1.87, m	1.47, m
14b			1.36, s	0.87, m	1.23, m
15a	2.13, d(15.2)	1.90, d(14.8)	1.40, s	1.51, s	1.46, m
15b	1.86, t(7.6)	1.69, d(14.8)			
16					
17a	3.93, d(8.8)	4.06, d(8.5)	4.02, d, (8.5)	3.99, d(8.4)	3.97, d (8.3)
17b	3.64, d(8.8)	3.75, d(8.5)	3.57, d, (8.5)	3.62, d(8.4)	3.60, s
18	0.97, s	1.00, s	1.21, s	1.07, s	1.01, s
19a	1.07, s	1.07, s	1.11, s	3.71d(10.7)	3.63, m
19b				3.44, d(10.7)	3.37, dd (11.3, 6.7)
20	0.87, s	0.86, s	1.36, s	1.28, s	1.18, s
1’					
2’	1.33, s	1.33, s	1.39, s	1.37, s	1.36, s
3’	1.36, s	1.34, s	1.38, s	1.37, s	1.37, s

[a] Chemical shifts (ppm) referenced to solvent peak (δ_H7.26 in $CDCl_3$) at 500 MHz.

similar ROESY correlation signals, δ values (δ_C 74.4), and *J* values (δ_H 3.97, d, *J* = 8.3 Hz, H-17a, δ_H 3.60, s, H-17b) for key protons to compound 1 at C-17, indicating that they share the same relative configuration at C-16 and C-17. Compound 9 was named Eupfisenoid I, and the structure was assigned to be *ent*-atisan-16*α*,17-acetonide-19-hydroxy-3-one.

Given the presence of a ketal carbon in the compound, we initially suspected that it might be an artifact. To investigate further, using a Waters Acquity ultraperformance liquid chromatograph (UPLC) tandem Xevo TQ-S micro triple quadrupole mass spectrometer, we conducted ultraperformance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) (Sun et al., 2023) quantitative analysis assays for compound 1, methanol extract, and the ethyl acetate extraction layer (Figure 4).

The experimental findings revealed the presence of compound 1 in both the methanol and ethyl acetate extracts, with the ethyl acetate extraction layer containing a higher quantity of compounds than the methanol extract (Table 4). This substantiates that compound 1 is not an artificial product but rather a natural plant-derived compound.

2.2 Bioassay results for compound 1

We performed cytopathic effect assays of SARS-CoV-2 using Vero E6 cells with remdesivir as a positive control to further understand the biological activity of compound 1. In DMSO (Figure 5Ba) as naive control, the Vero E6 cells were morphologically intact. SARS-CoV-2-infected Vero E6 cells showed a morphological deformation with chromatin condensation and karyopycnosis (Figure 5Bb), demonstrating a significant cytopathic effect. The addition of compound 1 reduced the number of cells that developed lesions in comparison to before (Figure 5Bc), but not as much as remdesivir (Figure 5Bd). Thus, assays of cytopathic effect clearly show the anti-SARS-CoV-2

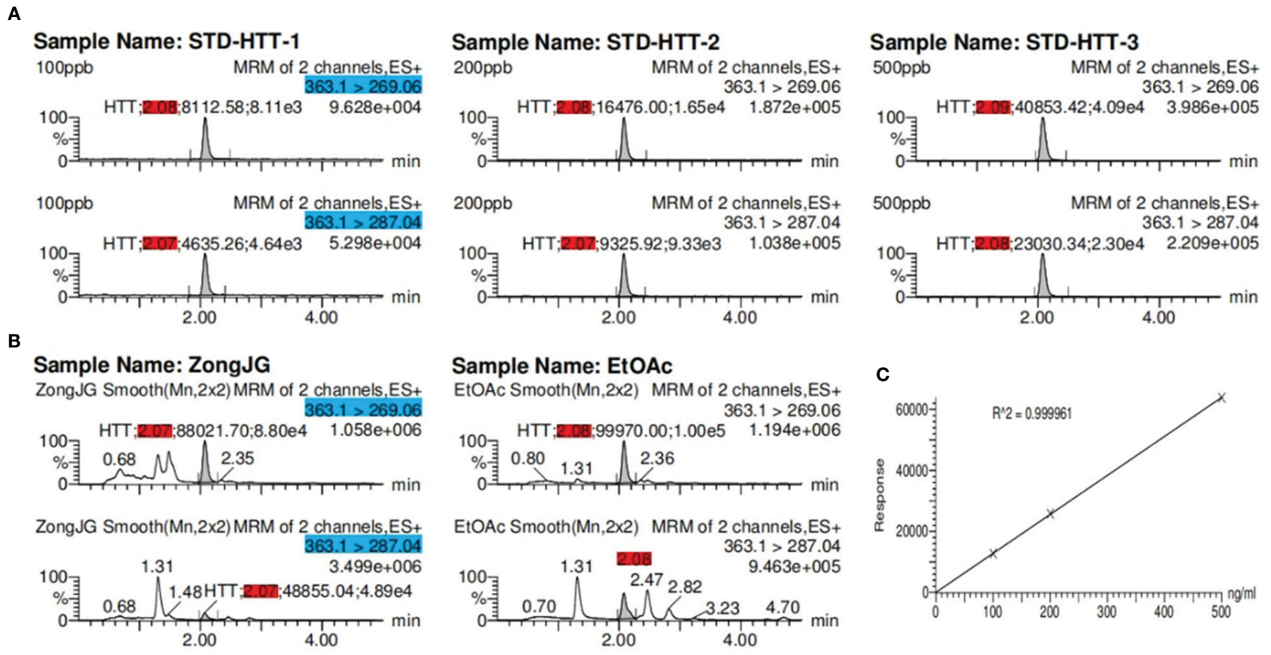
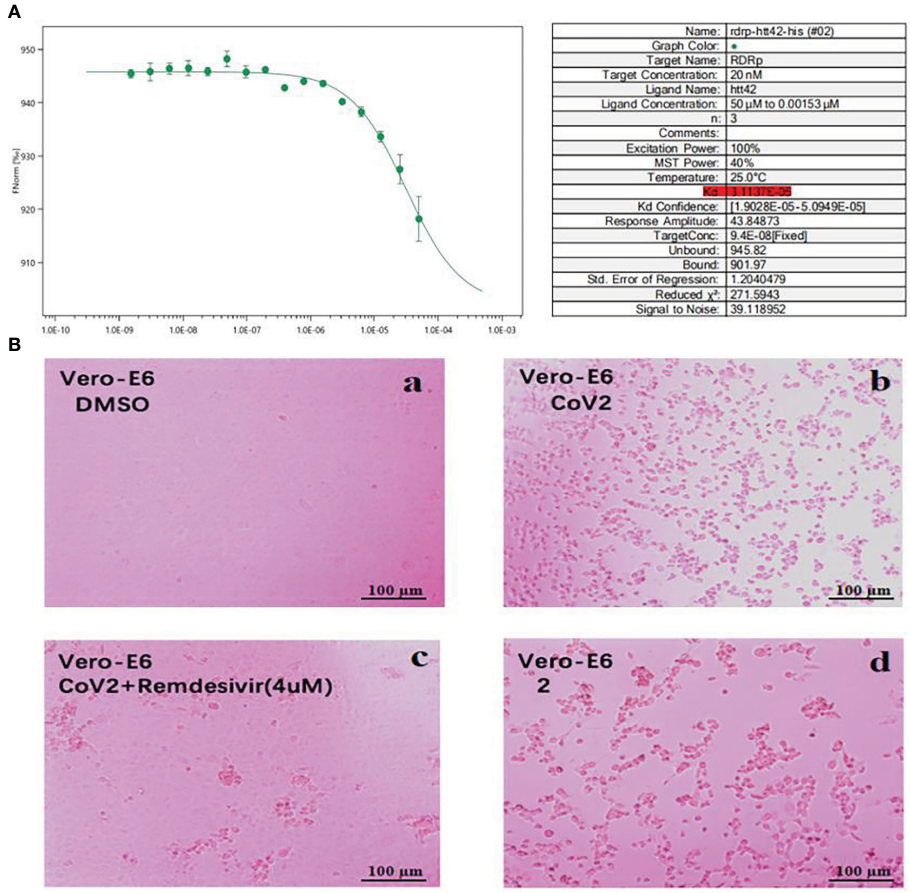


FIGURE 4
The UPLC-MS/MS results for sample. **(A)** Retention time (red) of compound at three concentrations; **(B)** retention time (red) of compound 1 at methanol extract and ethyl acetate extraction layer. The blue part represents the precursor and product ions of compound 1. **(C)** Coefficient of determination: R^2 = 0.999961 of compound 1.

TABLE 4 The UPLC-MS/MS results for sample.

Sample name	Type	Std. Conc	RT	Area	Response	ng/mL	%Dev
STD-HTT-1	Standard	100.00	2.08	8,112.583	12,747.847	99.7	−0.3
STD-HTT-2	Standard	200.00	2.08	16,476.002	25,801.927	201.7	0.8
STD-HTT-3	Standard	500.00	2.09	40,853.418	63,883.760	499.4	−0.1
ZongJG	Analyte		2.07	88,021.695	136,876.738	1070.0	
EtOAc	Analyte		2.08	99,970.000	156,843.273	1226.1	



Name:	rdrp-htt42-his (#02)
Graph Color:	●
Target Name:	RDRp
Target Concentration:	20 nM
Ligand Name:	htt42
Ligand Concentration:	50 μM to 0.00153 μM
n:	3
Comments:	
Excitation Power:	100%
MST Power:	40%
Temperature:	25.0°C
Kd:	3.1137E-05
Kd Confidence:	[1.9028E-05-5.0949E-05]
Response Amplitude:	43.84873
TargetConc:	9.4E-08[Fixed]
Unbound:	945.82
Bound:	901.97
Std. Error of Regression:	1.2040479
Reduced χ^2:	271.5943
Signal to Noise:	39.118952

FIGURE 5
(A) MST analysis results of compound 1. The linear fit is close to S, and the K_d = 31.13 μM. **(B)** Cytopathic effects of SARS-CoV-2-infected Vero E6 cells. (a) Vero E6 cells were treated with a DMSO solution. (b) Vero E6 cells were then infected with SARS-CoV-2; (c) SARS-CoV-2 + Remdesivir; (d) SARS-CoV-2 + compound 1 (because the compounds were screened in different batches, 2 in the figure actually represents compound 1; concentration was 20 μM).

activity of compound 1 (Zhao et al., 2023) (Figure 5B). We used molecular docking to screen for small-molecule ligands with better binding to SARS-CoV-2 RdRp. The molecular docking (Morris et al., 2008) results are as follows (Figure 6): compound 1 (−8.0 kcal/mol), compound 2 (−7.7 kcal/mol), compound 3 (−8.3 kcal/mol), compound 4 (−7.7 kcal/mol), compound 5 (−7.7 kcal/mol), compound 6 (−7.3 kcal/mol), compound 7 (−7.9 kcal/mol), compound 8 (−7.1 kcal/mol), and compound 9 (−7.4 kcal/mol). Compound 1 demonstrated tight binding to SARS-CoV-2 RdRp, with binding energies of −8.0 kcal/mol. This suggests their potential as foundational compounds for specific agents targeting the active site of SARS-CoV-2 RdRp. Considering RdRp as a primary target for anti-SARS-CoV-2 measures, it holds significant promise in discovering potential COVID-19 drugs. Thus, we further assessed the binding of compound 1 and SARS-CoV-2 Rdrp protein. Employing MST (Wienken et al., 2010), a method analyzing molecular movement in a microscopic temperature gradient field, we determined the dissociation constant K_d. The lower K_d (31.13 μM, Figure 5A) values for 1 indicate stronger affinity between compound 1 and SARS-CoV-2 RdRp, further confirming its binding to the SARS-CoV-2 Rdrp protein.

3 Materials and methods

3.1 General

One- and two-dimensional NMR spectra were determined by Bruker 500-MHz and 600-MHz NMR instruments with the internal standard: TMS. The chemical shifts δ were expressed in parts per million (ppm), and the coupling constant *J* was expressed in Hz. The CD spectra were measured on a photophysical circular dichroism spectrometer (Applied Photophysics, Leatherhead, Surrey, UK). HR-ESI-MS data acquisition was performed in positive mode on an Agilent 1290 UPLC/6540 Q-TOF mass spectrometer. The UV spectra were detected on a Shimadzu UV-2401A UV spectrometer, the specific spin data were detected on a JASCO DIP-370 digital spinometer, and the IR spectra were detected on a Tenor 27 infrared spectrometer with KBr pressurization assay. Single-crystal x-ray diffraction experiments were detected on a Bruker APEX DUO diffractometer with a copper target. Semipreparative HPLC was performed on an Agilent 1260 apparatus equipped with a UV detector and a Zorbax SB-C-18 (Agilent, 9.4 mm × 25 cm) column. Column chromatography (CC)

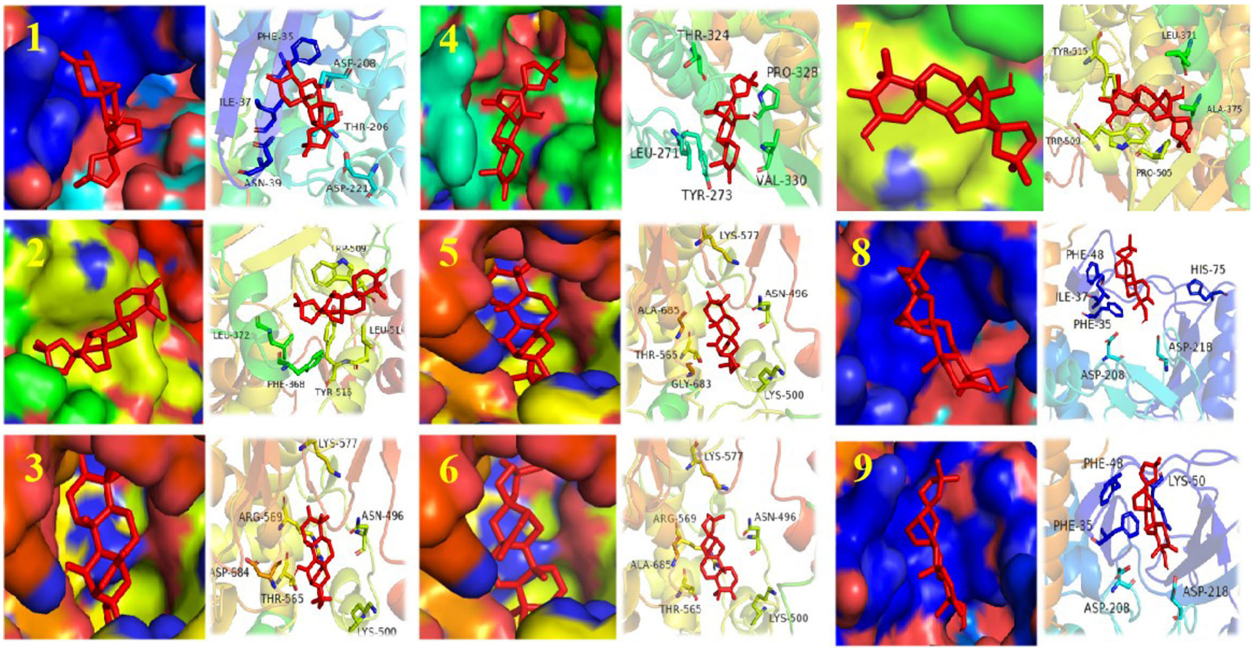


FIGURE 6
Structure of compounds 1–9 bound to SARS-CoV-2 Rdrp. The compound is denoted in red. The left shows the state of the compound in the substrate-binding pocket of the SARS-CoV-2 Rdrp. The right shows the co-crystal structures of the compound with SARS-CoV-2 Rdrp.

was performed using silica gel (200–300 mesh and H, Qingdao Marine Chemical Co. Ltd., Qingdao, China) and RP-C18 gel (40–63 μm, Merck, Darmstadt, Germany). Fractions were monitored by TLC (GF254, Qingdao Marine Chemical Co. Ltd., Qingdao, China), and spots were visualized by heating silica gel plates sprayed with 10% H_2SO_4 in EtOH. All solvents were distilled prior to use.

3.2 Plant material

The roots of *E. fischeriana* Steud were collected in September 2015 from Xianggelila City, Yunnan Province, People's Republic of China. The plant samples were identified by Prof. Xun Gong of the State Key Laboratory of Phytochemistry and Plant Resource in West China. Voucher specimens (HXJ20150915) were deposited at Kunming Institute of Botany (KIB), Chinese Academy of Sciences (CAS).

3.3 Isolation and purification phytochemicals of *Euphorbia fischeriana* Steud

The air-dried and powdered twigs and roots of *E. fischeriana* Steud were extracted three times with 75% EtOH at room temperature to give a crude extract (1.93 kg). The EtOAc fraction (540 g) was subjected to silica gel CC with a gradient elution of petroleum ether-acetone (50:1 to 20:1) to afford five fractions, A–E. Fraction D (140 g) was subjected to a silica gel column, eluted with petroleum ether-acetone (10:1, 8:2, 7:3, 6:4, 1:1, V/V), to yield 12 fractions (D1–D12), continuation of the fine purification of the compounds.

3.4 Purity of the isolated compounds

Fraction D3 (1.1 g) was chromatographed over a column of C18 reversed-phase silica gel (MeOH/H_2O, 20% to 100%, increase in 5% gradient) to give eight subfractions (D3a–D3h). Fraction D3e (300 mg) was purified by semipreparative HPLC with the mobile phase of 60% MeOH in H_2O to give compound **3** (40.5 mg, retention time: 35.0 min); using semipreparative HPLC with the mobile phase of 50% CH_3CN in H_2O led to the isolation of 6 (4.5 mg, retention time: 45.5 min) and 9 (9.7 mg, retention time: 55.5 min), and similarly, fraction D3h (30 mg) was purified by semipreparative HPLC with the mobile phase of 48% CH_3CN in H_2O to give compound 1 (7.3 mg, retention time: 50.0 min). Fraction D4 (10 g) was subjected to a silica gel column, eluted with acetic ether-acetone (60:1 to 1:1), to yield 11 factions (D4a–D4k). Fraction D4c (1.2 g) was chromatographed over a column of C18 reversed-phase silica gel (MeOH/H_2O, 20% to 100%, increase in 5% gradient) to give four subfractions (D4c1–D4c4). Fraction D4c1 (80 mg) was purified by semipreparative HPLC with the mobile phase of 42% CH_3CN in H_2O to give compounds 2 (1.7 mg, retention time: 37.0 min), **7** (10 mg, retention time: 48.6 min), and 8 (7.5 mg, retention time: 65.5 min). Fraction D4i (3 g) was chromatographed over a column of C18 reversed-phase silica gel (MeOH/H_2O, 30% to 100%, increase in 5% gradient) to give two subfractions (D4i1 and D4i2). Fraction D4i2 (40 mg) was purified by semipreparative HPLC with the mobile phase of 56% CH_3CN in H_2O to give compounds 4 (9.6 mg, retention time: 28 min) and 5 (6 mg, retention time: 42 min).

3.5 Compound naming rule

The compound name consists of a partial letter of the plant's Latin name (*E. fischeriana* Steud) and the suffixes of the terpenoids.

3.6 Spectral measurements

(4*R*,5*S*,8*S*,9*S*,10*R*,11*S*,12*S*,16*S*)-*ent*-11α-hydroxy-16α,17-acetonide-19-noratisan-3-one (Eupfisenoid A, compound 1), white crystals (MeOH); [α]24.6 D-17.96 (c 0.05, MeOH); UV (MeOH) λ_{max} (log ε) 203 (3.78) nm; IR (KBr) ν_{max} 3,438, 3,397, 2,922, 2,851, 1,734, 1,701, 1,646, 1,455, 1,371, 1,255, 1,171, 1,058, 721 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 1, 2; HR-ESI-MS peak at *m*/*z* 363.2529 $[M+H]^+$ (calcd. $C_{23}H_{35}O_4$, 363.2530).

Ent-11α-hydroxyatisan-16α,17-acetonide-3-one (Eupfisenoid B, compound 2), white amorphous powder; [α]23.9 D-48.96 (c 0.25, MeOH); UV (MeOH) λ_{max} (log ε) 203 (4.03) nm; IR (KBr) ν_{max} 3,457, 2,934, 2,855, 1,702, 1,631, 1,448, 1,370, 1,213, 1,121, 1,058, 722 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 1, 2; HR-ESI-MS peak at *m*/*z* 399.2512 $[M+Na]^+$ (calcd. $C_{23}H_{36}NaO_4$, 399.2506).

Ent-11α-hydroxyatisan-16α,17-acetonide-3,14-dione (Eupfisenoid C, compound 3), colorless crystals; [α]23.1 D-18.31 (c 0.13, MeOH); UV (MeOH) λ_{max} (log ε) 203 (4.03) nm; IR (KBr) ν_{max} 3,439, 2,945, 2,860, 1,711, 1,692, 1,440, 1,371, 1,216, 1,154, 1,060, 721 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 1, 2; HR-ESI-MS peak at *m*/*z* 391.2486 $[M+ H]^+$ (calcd. $C_{23}H_{35}O_5$,391.2479).

Ent-11β-hydroxyatisan-16α,17-acetonide-3,14-dione (Eupfisenoid D, compound 4), white crystals; [α]20 D-1 (c 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 203.5 (4.07) nm; IR (KBr) ν_{max} 3,493, 2,928, 2,869, 1,697, 1,452, 1,374, 1,256, 1,125, 1,074, 720 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 1, 2; HR-ESI-MS peak at *m*/*z* 391.2477 $[M+ H]^+$ (calcd. $C_{23}H_{35}O_5$,391.2477).

Ent-atisan-16α,17-acetonide-3,14-dione (Eupfisenoid E, compound 5), white powder; [α]19.9 D19.06 (c 0.11, MeOH); UV (MeOH) λ_{max} (log ε) 204 (4.36) nm; IR (KBr) ν_{max} 3,425, 2,922, 2,863, 1,716, 1,630, 1,458, 1,373, 1,261, 1,149, 1,060, 722 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 2, 3; HR-ESI-MS peak at *m*/*z* 433.2229 $[M+COOH]^-$ (calcd. $C_{24}H_{33}O_7$, 433.2232).

Ent-atisan-16α,17-acetonide-3-one (Eupfisenoid F, compound 6), colorless crystals; [α]25 D1 (c 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 203 (4.00) nm; IR (KBr) ν_{max} 3,432, 2,937, 2,860, 1,711, 1,638, 1,459, 1,369, 1,249, 1,152, 1,060, 723 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 2, 3; HR-ESI-MS peak at *m*/*z* 373.2382 $[M-H]^-$ (calcd. $C_{23}H_{33}O_4$, 373.2384).

Ent-2,13α-dihydroxyatisan-1-en-16α,17-acetonide-3-one (Eupfisenoid G, compound 7), colorless needle crystals; [α]24.6 D-44.62 (c 0.13, MeOH); UV (MeOH) λ_{max} (log ε) 270 (3.50) nm; IR (KBr) ν_{max} 3,439, 2,982, 1,717, 1,648, 1,572, 1,250, 714 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 2, 3; HR-ESI-MS peak at *m*/*z* 435.2387 $[M+COOH]^-$ (calcd. $C_{24}H_{35}O_7$,435.2388).

Ent-atisan-1-en-16α,17-acetonide-19-hydroxy-3-one (Eupfisenoid H, compound 8), white powder; [α]23.3 D-11.86 (c 0.07, MeOH); UV (MeOH) λ_{max} (log ε) 228 (3.91) nm; IR (KBr) ν_{max} 3,437, 2,934, 2,870, 1,667, 1,369, 1,252, 1,156, 1,057, 715 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 2, 3; HR-ESI-MS peak at *m*/*z* 375.2536 $[M+H]^+$ (calcd. $C_{23}H_{35}O_4$, 375.2530).

Ent-atisan-16α,17-acetonide-19-hydroxy-3-one (Eupfisenoid I, compound 9), white powder; [α]24.2 D-20.71 (c 0.14, MeOH); UV (MeOH) λ_{max} (log ε) 203.5 (3.98) nm; IR (KBr) ν_{max} 3,426, 2,929, 2,868, 1,734, 1,451, 1,370, 1,252, 1,057, 724 cm^{-1}; ^{1}H and ^{13}C NMR, see Tables 2, 3; HR-ESI-MS peak at *m*/*z* 399.2510 $[M+Na]^+$ (calcd. $C_{23}H_{36}NaO_4$, 399.2506).

3.7 X-ray crystallographic data for compound 1

Colorless crystals of 1 were obtained by recrystallization in MeOH at room temperature. X-ray crystal data were acquired on a Bruker APEX-II CCD detector with graphite monochromated Cu Kα radiation (λ = 1.541, 78 Å). The structure of 1 was directly elucidated using SHELXL-97 (Sheldrick 2008) and refined by the full-matrix least-squares difference Fourier method. The x-ray data of 1 have been deposited at the Cambridge Crystallographic Data Center.

3.8 Assays of cytopathic effect of compound 1

Vero E6 cells were seeded in 96-well plates and grown overnight. Cells were incubated with SARS-CoV-2 at 37°C for 2 h and infected at a multiplicity of infection of 0.1. Then, the cells were incubated with the maintenance medium in compound **1** (1 mg, 20 μM). Remdesivir (4 μM) was used as positive controls and DMSO solution was used as naive control. After 72 h, cell viability was then assessed using colorimetric MTS assays (Promega Corp.) as described by the manufacturer. Thereafter, the cells were photographed using a microscope.

3.9 Molecular docking of compound 1

Autodock tool 1.5.6 software was used to perform operations such as acceptor polarization of hydrogen, Gasteiger charge distribution, and removal of water molecules for 7BV1 and small-molecule ligands. Set the docking central coordinates of the Autodock Vina software to center_x = 131.622, center_y = 135.777, center_z = 121.114. The docking box size is 126 Å, the exhaustiveness value of the search parameter is 10, the top nine conformations are output according to the docking score, and the default value is selected for the rest of the parameters. Finally, the docking results were visualized by PyMOL software.

3.10 Microscale thermophoresis of compound 1

Purified SARS-CoV-2 RdRp protein was subjected to NHS (lysine labeling method) labeling. One hundred microliters of 10 μM protein and lysine labeling reagent was incubated in a dark environment for 30 min. The initial concentration of compound 1 (1 mg) was set at 20 mM, and it was subsequently diluted to a concentration of 100 μM using PBS-T buffer, with 16 gradient dilutions of 100 μM. Twenty microliters of compound 1 solution was added to PCR tube 1 and 10 μL of PBS-T buffer was added to the remaining 15 PCR tubes. Ten microliters of solution was pipetted from tube 1 and added to tube 2 and mixed; this process was repeated for tubes 3 to 16, the serial dilution was completed sequentially, and 10 μL was discarded in the last tube. Ten microliters of SARS-CoV-2 RdRp protein solution was added to each PCR tube and mixed well, a capillary was used to aspirate the sample, and the MST experiment was performed. The instrument was set to a medium MST power, and the K_d values and binding curves of compound 1 and the SARS-CoV-2 RdRp protein were ultimately obtained. The dilution and assay steps were repeated on two separate occasions.

4 Discussion

Plants, with their complex secondary metabolism, produce a wide range of compounds and offer significant advantages in the treatment of infectious diseases. Throughout history, phytotherapy has been utilized during epidemics such as the Black Death, smallpox, tuberculosis, malaria, and Spanish flu, providing valuable references for mankind on the safety and effectiveness of plant-based treatments (Garcia, 2020). Developing anti-SARS-CoV-2 drugs based on existing antiviral plants that have a proven track record could streamline the clinical trial process and expedite the identification of potential plant inhibitors (Pandey et al., 2020). In this study, we focused on extracting and isolating compounds from the roots of *E. fischeriana* Steud with potential activity against SARS-CoV-2. Nine undescribed *ent*-atisane type diterpenoids were successfully isolated from this plant. The elucidation of their configurations was achieved through a comprehensive suite of 1D and 2D NMR spectroscopic analyses as well as x-ray diffraction. Atisane-type diterpenoids belong to the tetracyclic diterpenoid family. They possess a bicyclo[2.2.2]octane ring system, decorated with methyl groups at C-4, C-10, and C-16. The most frequently oxidized positions of the *ent*-atisane skeleton are C-3, C-16, and C-17; C-16 and C-17 are typically in the form of an olefin (Drummond et al., 2021). The double bond at C-16 and C-17 is oxidized to a terthydroxyl group at C-16, which undergoes further oxidation to form an acetone dimethyl acetal to finally form the compound we obtained in this study. Many of the ent-atisane diterpenoids have antiviral activity, including anti-influenza A virus (Zhang et al., 2024), anti-HIV-1 (Yan et al., 2018), and anti-human rhinovirus 3 (Wang et al., 2018). Notably, in our study, one of these compounds exhibited promising anti-SARS-CoV-2 activity, and cytopathic effect assays confirmed the anti-SARS-CoV-2 activity of compound 1.

Molecular docking plays a crucial role in the search for antiviral compounds within various plant extracts (Pandey et al., 2020). Using this way of thinking, based on existing antiviral plants, scientists have already employed an integrated approach combining network pharmacology analysis, molecular docking, LC-MS analysis, and bioassays to uncover the potential ingredients of Scutellariae radix for SARS-CoV-2 (Liu et al., 2022). In our study, molecular docking predicted that compound 1 has an affinity for RdRp, with a binding energy of −8.0 kcal/mol. Based on the molecular docking results, our study further investigated the affinity between compound 1 and RdRp using MST, and these findings suggest that compound 1 could serve as a potential therapeutic target against SARS-CoV-2 RdRp. The search for potential antiviral agents targeting SARS-CoV-2 RdRp remains a subject of ongoing research. For instance, the polyphenolic compound gossypol could directly block SARS-CoV-2 RdRp, thereby inhibiting SARS-CoV-2 replication in cellular and mice models of infection (Wang et al., 2022). Other compounds like quercetin and procyanidins have also demonstrated promising inhibitory effects on SARS-CoV-2 RdRp through *in vitro* enzyme assays (Jin et al., 2022; Munafò et al., 2022). The diversity of plant-derived compounds is illustrated by the fact that different types of compounds can act on the same targets. Based on our work, compound 1 also appears promising as a SARS-CoV-2 RdRp inhibitor. In addition, structural modifications may further enhance the anti-SARS-CoV-2 activity of compound 1.

Medicinal plants have long been used to treat infectious diseases and have been vital to human society. As our understanding of plant science grows, we will continue to discover new plant-derived drugs. In order to fully utilize the role of medicinal plants, we need to apply modern technology and a great deal of multidisciplinary research. The current study, which is a tiny portion of plant science research, aims to offer some new scientific foundation.

5 Conclusions

In conclusion, this study conducted a preliminary exploration of the material basis, active compounds, and related targets of *E. fischeriana* Steud against SARS-CoV-2. Nine previously unreported *ent*-atisane-type diterpenoid compounds were isolated from the roots of *E. fischeriana* Steud, and their activities, targets, and mechanisms against SARS-CoV-2 were investigated. Cell pathology experiments confirmed that *ent*-atisane-type diterpenoid compound 1 exhibited certain anti-SARS-CoV-2 activity. Compound 1 was predicted to bind to RdRp through high-throughput virtual screening. Subsequently, using MST technology, the affinity between compound 1 and RdRp was tested, revealing that compound 1 could form a stable complex with RdRp protein, with a K_d value of 31.13 μM. RdRp protein was preliminarily identified as the target of compound 1 against SARS-CoV-2. This work expands the research achievements of *E. fischeriana* Steud and its diterpenoid components in the field of antiviral research, providing valuable references for the discovery of potential anti-SARS-CoV-2 targets and mechanisms. Additionally, it enriches the library of antiviral active compounds against SARS-CoV-2 and clarifies the potential of *ent*-atisane-type diterpenoid compounds in antiviral research, offering new insights for COVID-19 drug development.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

Ethical approval was not required for the studies on animals in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

TR: Conceptualization, Writing – original draft, Methodology. Z-RX: Conceptualization, Methodology, Writing – original draft. Y-WZ: Formal analysis, Writing – review & editing. S-RF: Formal analysis, Writing – review & editing. JR: Formal analysis, Writing – review & editing. QZ: Writing – review & editing, Methodology. X-LS: Writing – review & editing, Formal analysis. S-LW: Writing – review & editing. L-LX: Writing – review & editing. MQ: Writing – review & editing, Methodology. C-XJ: Methodology, Writing – review & editing. X-JH: Writing – review & editing, Funding acquisition, Supervision. D-ZC: Conceptualization, Funding acquisition, Methodology, Software, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the National Natural Science Foundation of China (81973212, 82293683, 82293680); Key Research and Development Project of Yunnan Province (202203AC100009, 202003AD150012),; Foundation of Central Asian Drug Discovery and Development Center of Chinese Academy of Sciences (CAM202103, China); Project of Yunnan Characteristic Plant Screening and R&D Service CXO Platform (2022YKZY001).

Acknowledgments

We deeply thank all the authors and reviewers who have participated in this Research Topic. We thank Haiyang Liu (Kunming Institute of Botany) for plant photos.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1425759/full#supplementary-material

References

Dai, Y.-J., Wan, S.-Y., Gong, S.-S., Liu, J.-C., Li, F., and Kou, J.-P. (2020). Recent advances of traditional Chinese medicine on the prevention and treatment of COVID-19. *Chin. J. Nat. Med.* 18, 881–889. doi: 10.1016/S1875-5364(20)60031-0

Drummond, G.-J., Grant, P.-S., and Brimble, M.-A. (2021). *ent*-Atisane diterpenoids: isolation, structure and bioactivity. *Nat. Prod. Rep.* 38, 330–345. doi: 10.1039/D0NP00039F

Feikin, D.-R., Higdon, M.-M., Abu-Raddad, L.-J., Andrews, N., Araos, R., Goldberg, Y., et al. (2022). Duration of effectiveness of vaccines against SARS-CoV-2 infection and covid-19 disease: results of a systematic review and meta-regression. *Lancet* 399, 924–944. doi: 10.1016/S0140-6736(22)0015-0

Flora of China Editorial Committee (1997). *Flora of China* (Beijing: Science Press).

Garcia, S. (2020). Pandemics and traditional plant-based remedies. A historical-botanical review in the era of COVID-19. *Front. Plant sci.* 11. doi: 10.3389/fpls.2020.571042

Gorbalenya, A.-E., Baker, S.-C., Baric, R.-S., de Groot, R.-J., Drosten, C., Gulyaeva, A.-A., et al. (2020). Severe acute respiratory syndrome-related coronavirus: The species and its viruses - a statement of the Coronavirus Study Group. *BioRxiv.* 8, 1–20. doi: 10.1101/2020.02.07.937862

Hu, B., Guo, H., Zhou, P., and Shi, Z.-L. (2021). Characteristics of SARS-coV-2 and COVID-19. *Nat. Rev. Microbiol.* 19, 141–154. doi: 10.1038/s41579-020-00459-7

Huang, Y.-P. (2021). Studies on the chemical constituents of the small polar parts of *Euphorbia fischeriana* Steud. Yunnan Normal University, Yunnan.

Jin, Y.-H., Lee, J., Jeon, S., Kim, S., Min, J.-S., and Kwon, S. (2022). Natural polyphenols, 1, 2, 3, 4, 6-O-pentagalloylglucose and proanthocyanidins, as broad-spectrum anticoronaviral inhibitors targeting Mpro and RdRp of SARS-CoV-2. *Biomedicines* 10, 1170. doi: 10.3390/biomedicines10051170

Kuang, Y., Fu, S.-Y., Wang, F., Ren, F.-C., Yang, D.-F., Yang, S.-X., et al. (2017). Two new *ent*-atisane diterpenoids from the whole plants of *Euphorbia wallichii*. *Nat. Prod. Res.* 31, 849-852. doi: 10.1080/14786419.2016.1250088

Li, X.-L., Li, Y., Wang, S.-F., Zhao, Y.-L., Liu, K.-C., Wang, X.-M., et al. (2009). Ingol and ingenolditerpenes from the aerial parts of Euphorbia royleana and their antiangiogenic activities. *J. Nat. Prod.* 72, 1001–1005. doi: 10.1021/np800816n

Li, Y.-N., He, J., Zhang, J., Shi, Y.-X., Guo, L.-B., Peng, Z.-C., et al. (2021). Existing knowledge on *Euphorbia fischeriana* Steud. (Euphorbiaceae): Traditional uses,clinical applications, phytochemistry, pharmacology and toxicology. *J. Ethnopharmacol.* 275, 114095. doi: 10.1016/j.jep.2021.114095

Liu, J.-Z., Meng, J.-R., Li, R.-F., Jiang, H.-M., Fu, L., Xu, T., et al. (2022). Integrated network pharmacology analysis, molecular docking, LC-MS analysis and bioassays revealed the potential active ingredients and underlying mechanism of Scutellariae radix for COVID-19. *Front. Plant sci.* 13. doi: 10.3389/fpls.2022.988655

Morris, G.-M., Huey, R., and Olson, A.-J. (2008). Using autodock for ligand-receptor docking. *Curr. Protoc. Bioinf.* 24, 8–14. doi: 10.1002/0471250953.bi0814s24

Munafò, F., Donati, E., Brindani, N., Ottonello, G., Armirotti, A., De Vivo, M., et al. (2022). Quercetin and luteolin are single-digit micromolar inhibitors of the SARS-CoV-2 RNA-dependent RNA polymerase. *Sci. Rep.* 12, 10571. doi: 10.1038/s41598-022-14664-2

Pandey, A., Khan, M.-K., Hamurcu, M., and Gezgin, S. (2020). Natural plant products: A less focused aspect for the COVID-19 viral outbreak. *Front. Plant sci.* 11. doi: 10.3389/fpls.2020.568890

Pharmacopoeia Commission of PRC (2020). *Pharmacopoeia of People's Republic of China* Vol. 298. Part 1 (Beijing: China Medical Science Press).

Sievers, B.-L., Chakraborty, S., Xue, Y., Gelbart, T., Gonzalez, J.-C., Cassidy, A.-G., et al. (2022). Antibodies elicited by SARS-CoV-2 infection or mRNA vaccines have reduced neutralizing activity against Beta and Omicron pseudoviruses. *Sci. Transl. Med.* 14, eabn7842. doi: 10.1126/scitranslmed.abn7842

Snijder, E.-J., Decroly, E., and Ziebuhr, J. (2016). The nonstructural proteins directing coronavirus RNA synthesis and processing. *Adv. Virus Res.* 96, 59–126. doi: 10.1016/bs.aivir.2016.08.008

Sun, Z., Liu, X., Zuo, W., Fu, Q., Xu, T.-T., Cui, L.-Y., et al. (2023). Development of a robust UPLC-MS/MS method for the quantification of riluzole in human plasma and its application in pharmacokinetics. *Front. Pharmacol.* 14. doi: 10.3389/fphar.2023.1227354

te Velthuis, A. J. (2014). Common and unique features of viral RNA-dependent polymerases. *Cell. Mol. Life Sci.* 71, 4403–4420. doi: 10.1007/s00018-014-1695-z

Wang, B., Wei, Y., Zhao, X., Tian, X., Ning, J., Zhang, B., et al. (2018). Unusual *ent*-atisane type diterpenoids with 2-oxopropyl skeleton from the roots of Euphorbia ebracteolata and their antiviral activity against human rhinovirus 3 and enterovirus 71. *Bioorg. Chem.* 81, 234–240. doi: 10.1016/j.bioorg.2018.08.029

Wang, W.-J., Li, W.-K., Wen, Z.-Y., Wang, C., Liu, W.-L., Zhang, Y.-F., et al. (2022). Gossypol broadly inhibits coronaviruses by targeting RNA-dependent RNA polymerases. *Adv. Sci.* 9, 2203499. doi: 10.1002/advs.202203499

Wienken, C.-J., Baaske, P., Rothbauer, U., Braun, D., and Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* 1, 100. doi: 10.1038/ncomms1093

World Health Organization (2024). WHO coronavirus (COVID-19) dashboard. Available online at: https://covid19.who.int/.

Yan, S.-L., Li, Y.-H., Chen, X.-Q., Liu, D., Chen, C.-H., and Li, R.-T. (2018). Diterpenes from the stem bark of *Euphorbia neriifolia* and their in *vitro* anti-HIV activity. *Phytochemistry* 145, 40–47. doi: 10.1016/j.phytochem.2017.10.006

Zhang, J.-G., Xiao, J., Chen, X.-J., Zhang, R.-R., Ma, B.-P., and Liu, Z.-Q. (2024). *Ent*-atisane diterpenoids from Euphorbia wallichii and their anti-influenza A virus activity. *Phytochemistry* 220, 113996. doi: 10.1016/j.phytochem.2024.113996

Zhao, D.-D., Qin, Y.-H., Liu, J.-Q., Tang, K.-G., Lu, S.-Y., Liu, Z.-R., et al. (2023). Orally administered BZL-sRNA-20 oligonucleotide targeting TLR4 effectively ameliorates acute lung injury in mice. *Sci. China Life Sci.* 66, 1589–1599. doi: 10.1007/s11427-022-2219-0

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 September 2024
DOI 10.3389/fpls.2024.1464388

OPEN ACCESS

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Tuo Li,
Dalian University, China
Liang Xiangxiu,
South China Agricultural University, China

*CORRESPONDENCE
Qi Shen
shenqi@gzucm.edu.cn
Tianyue An
antianyue2007@126.com

†These authors have contributed equally to this work

RECEIVED 14 July 2024
ACCEPTED 19 August 2024
PUBLISHED 10 September 2024

CITATION
Wu Z, Li M, Liang X, Wang J, Wang G, Shen Q and An T (2024) Crucial amino acids identified in Δ12 fatty acid desaturases related to linoleic acid production in *Perilla frutescens*.
Front. Plant Sci. 15:1464388.
doi: 10.3389/fpls.2024.1464388

Crucial amino acids identified in Δ12 fatty acid desaturases related to linoleic acid production in *Perilla frutescens*

Zhenke Wu[1†], Mingkai Li[1†], Xiqin Liang[1], Jun Wang[1],
Guoli Wang[1], Qi Shen[2*] and Tianyue An[1*]

[1]Featured Laboratory for Biosynthesis and Target Discovery of Active Components of Traditional Chinese Medicine, School of Traditional Chinese Medicine, Binzhou Medical University, Yantai, China,
[2]Institute of Medical Plant Physiology and Ecology, School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China

Perilla oil from the medicinal crop *Perilla frutescens* possess a wide range of biological activities and is generally used as an edible oil in many countries. The molecular basis for its formation is of particular relevance to perilla and its breeders. Here in the present study, four *PfFAD2* genes were identified in different perilla cultivars, PF40 and PF70, with distinct oil content levels, respectively. Their function was characterized in engineered yeast strain, and among them, PfFAD2-1^{PF40}, PfFAD2-1^{PF70} had no LA biosynthesis ability, while PfFAD2-2^{PF40} in cultivar with high oil content levels possessed higher catalytic activity than PfFAD2-2^{PF70}. Key amino acid residues responsible for the enhanced catalytic activity of PfFAD2-2^{PF40} was identified as residue R221 through sequence alignment, molecular docking, and site-directed mutation studies. Moreover, another four amino acid residues influencing PfFAD2 catalytic activity were discovered through random mutation analysis. This study lays a theoretical foundation for the genetic improvement of high-oil-content perilla cultivars and the biosynthesis of LA and its derivatives.

KEYWORDS

***Perilla frutescens*, catalytic activity, molecular docking, mutation, FAD2 desaturase**

1 Introduction

Linoleic acid (LA) is an ω-6 polyunsaturated fatty acid (FA) and is indispensable for human health (Belury, 2023; Saini and Keum, 2018). LA is pivotal in health supplements and food products, and exhibits notable pharmacological activities, such as antioxidant and anti-inflammatory effects. Its applications extend to metabolic regulation, blood sugar management, and cardiovascular disease prevention (Yang et al., 2020; Alarcon-Gil et al., 2022; Hamilton and Klett, 2021; Marangoni et al., 2020). Furthermore, LA serves as a vital precursor for α-linolenic acid synthesis, endowing it with considerable economic significance and development potential (Choudhary and Mishra, 2021). Currently, large-scale LA production is predominantly achieved through extraction from natural oilseed crops. Consequently, breeding these crops is essential for the sustainable advancement of the LA synthesis industry (Sana et al., 2018; Kanbar et al., 2023; Yang et al., 2023; Pathak et al., 2014). Perilla seeds, harvested from the medicinal crop *Perilla frutescens*, are rich in unsaturated fatty acids, including oleic acid (OA), LA, and α-linolenic acid, rendering them a significant oilseed crop (Kim et al., 2020). Nonetheless, the diverse perilla cultivars and the varying LA content among them constrain the large-scale cultivation and improvement of these varieties (Park et al., 2022; Lee et al., 2019). Therefore, investigating the genetic basis of LA phenotypes and identifying Δ12 fatty acid desaturases (FAD2) with high catalytic efficiency for LA synthesis is crucial for perilla breeding.

Omics technologies offer an efficient approach to studying cellular phenotypes and gene functions comprehensively (Fagerberg et al., 2014). Through omics sequencing and bioinformatics analysis, plant genes can be scrutinized to pinpoint relevant functional genes and elucidate the molecular mechanisms underlying specific biological processes (Su et al., 2023). Consequently, high-throughput sequencing technologies, known for their substantial data output and accuracy, are extensively employed to analyze secondary metabolite variations between plant varieties and to identify key enzymes. For example, transcriptome sequencing analysis of different *Ziziphus jujuba* Mill. varieties can identify key genes in the flavonoid biosynthesis pathway and elucidate the molecular mechanisms affecting epicatechin content across different varieties (Wang et al., 2023). Similarly, transcriptomic and phenotypic analyses of differential genes associated with unsaturated fatty acid content in soybeans can provide insights into enhancing fatty acid composition in soybean seeds (Liu et al., 2022a).

The eukaryotic model organism *Saccharomyces cerevisiae*, renowned for its well-characterized genetic background and ease of genetic manipulation, is widely utilized for enzyme functional verification, including terpene synthases and fatty acid desaturases (Deng et al., 2022; Runguphan and Keasling, 2014). Yeast cells possess a natural fatty acid synthesis pathway capable of generating pyruvate through glycolysis, which subsequently undergoes decarboxylation, dehydrogenation, and a series of reactions to produce malonyl-CoA. Using malonyl-CoA as a basic unit, yeast cells can synthesize various free fatty acids (FAs), among which OA, a monounsaturated fatty acid, serves as a crucial precursor for LA synthesis. Consequently, employing OA-overproducing *S. cerevisiae* strains for the functional characterization of fatty acid desaturases is currently an effective strategy (Kuziora et al., 1983; Miao et al., 2019). Furthermore, *S. cerevisiae* is extensively applied in the construction of microbial cell factories for fatty acids and their derivatives (Guo et al., 2022).

Previous research has reported the genome data of perilla and conducted transcriptome analyses on various perilla cultivars (Wu et al., 2021). In this study, two perilla cultivars with distinct oil content levels, PF40 and PF70, were selected (Zou et al., 2024). Four *PfFAD2* genes were identified from their transcriptome data and designated as PfFAD2-1^{PF40}, PfFAD2-1^{PF70}, PfFAD2-2^{PF40} and PfFAD2-2^{PF70}, respectively. Then these genes were functionally characterized using our constructed OA-overproducing *S. cerevisiae* strain. The results revealed that PfFAD2-1^{PF40} and PfFAD2-1^{PF70} lacked the capability to synthesize LA, while only PfFAD2-2^{PF40} and PfFAD2-2^{PF70} could catalyze the conversion of OA to LA. Furthermore, PfFAD2-2^{PF40}, derived from the high oil-producing perilla cultivar, exhibited superior catalytic activity. Through sequence alignment, molecular docking, and site-directed mutation studies, key amino acid residue responsible for the enhanced catalytic activity of PfFAD2-2^{PF40} compared to PfFAD2-2^{PF70} were identified. Additional amino acid residues influencing PfFAD2 catalytic activity were discovered through random mutation analysis. This study lays a theoretical foundation for the genetic improvement of high-oil-content perilla cultivars and provides optimal microbial elements for the biosynthesis of LA and its derivatives.

2 Materials and methods

2.1 Plant materials and yeast strains

Perilla seeds were collected from their natural habitat, specifically the plant material cultivation greenhouse of Guangzhou University of Chinese Medicine, during their developmental period (late August to mid-September). Seeds from various perilla cultivars were harvested within one-month post-flowering, with calyxes removed before storage in liquid nitrogen for future use. The *S. cerevisiae* strain BY4741, used as the chassis strain in this study, was purchased from Shanghai Yuanye Bio-Technology Co., Ltd.

2.2 Gene cloning

RNA was extracted from perilla seeds using the FastPure Plant Total RNA Isolation Kit (Vazyme, China). Subsequently, 1 μL of RNA was reverse transcribed into cDNA using the HiScript III 1st Strand cDNA Synthesis Kit (+gDNA wiper) (Vazyme, China). Following the manufacturer's protocol, the full-length sequence of the target gene was cloned, ligated into a Blunt vector (TransGen Biotech, China), and sequenced. Primers used in this part were listed in Supplementary Table S1.

2.3 Construction of engineered yeast strain

Firstly, the GAL10p and GAL1p promoters of pESC-URA/HIS/LEU (Miaoling Biology, China) were replaced by TEF1p and PGK1p to get glucose induced vector pTP-URA/HIS/LEU. *TGL1*, *TGL3* and *TGL5* were cloned from the genomic DNA of *S. cerevisiae* strain BY4741. *TGL1* and *TGL5* were ligated into pTP-URA, and *TGL3* was ligated into pTP-HIS. Then the expression cassette of URA maker and *TGL1* and *TGL5*, and the expression cassette of HIS maker and *TGL3*, were cloned from the constructed pTP vectors. These two expression cassettes were integrated into the *FAA1* and *FAA4* sites in the yeast genome by using the method reported in the previous work (Jiang et al., 2022). The schematic diagram illustrating the construction of the yeast strain was shown in Supplementary Figure S1. Primers used in this part were listed in Supplementary Table S1.

2.4 Construction and transformation of yeast expression vectors

The cloned *PfFAD2* gene was inserted into the yeast expression vector pTP-LEU between the Sac I and Not I restriction sites in the multiple cloning sites, and verified by sequencing. The transformation plasmid was mixed with yeast competent cells in a solution containing 50% PEG3350 (Solarbio, China), 1 mol/L lithium chloride (Solarbio, China), and salmon sperm DNA (Solarbio, China), followed by incubation at 42°C for 1 h. The transformed yeast cells were centrifuged at 9500 rpm for 1 min, and the supernatant was discarded. The cell pellet was resuspended in YPD medium and cultured at 30°C with shaking at 220 rpm for 3 h. After washed twice with distilled water, the samples were cultured on appropriately labeled solid media. Primers used in this part were listed in Supplementary Table S1.

2.5 Shake flask cultivation

A single colony of positive yeast strains was cultured in 5 mL of SD medium (6.67 g/L yeast nitrogen base without amino acids (BD Difco, America), 20 g/L glucose (Macklin, China), supplemented with corresponding amino acids) at 30°C and 220 rpm. Upon reaching an OD_{600} of 1.5-2.5, an appropriate volume of the seed culture was transferred to 50 mL of induction medium (6.67 g/L yeast nitrogen source without amino acids, 20 g/L glucose, supplemented with appropriate amino acids). The culture was then incubated at 30°C and 220 rpm for 5 days.

2.6 Extraction and detection of compounds

After fermentation, 600 μL of the cultured broth was collected and combined with 30 μL of 40% tetrabutylammonium hydroxide (Macklin, China) and 200 μL of dichloromethane (Macklin, China) solution containing 200 mM methyl iodide (Sigma-Aldrich, America). The mixture was vortexed at 2000 rpm for 30 min and then centrifuged at 5000 g. The lower dichloromethane layer was transferred to a glass vial, dried, and resuspended in 200 μL of n-hexane. The prepared sample was subsequently analyzed by gas chromatography-mass spectrometry (GC-MS).

GC-MS analysis was conducted using an Agilent 7890 gas chromatograph equipped with an HP-5MS column and a Saturn 2100 ion trap mass spectrometer. Samples were injected in splitless mode with an injection volume of 1 μL. The temperature program was as follows: initial temperature of 50°C held for 2 min, increased to 140°C at a rate of 30°C/min, then raised to 280°C at a rate of 10°C/min, and held for 3 min.

2.7 Site-directed and random mutation

The selected *PfFAD2-2*^{PF70} containing yeast expression vector served as a template to generate the *PfFAD2-2*^{PF70} mutants using the Mut Express II Fast Mutagenesis Kit V2 (Vazyme, China). The amplified plasmid underwent digestion with the restriction endonuclease DpnI (Thermo Fisher Scientific, America) for one hour, followed by transformation into competent *Escherichia coli* DH5α cells ((TransGen Biotech, China)). Positive colonies were identified and sequenced to confirm the site-directed mutation. Primers used in this part were listed in Supplementary Table S1.

Random mutation was conducted using the *PfFAD2-2*^{PF40} gene as a template. The fragment was amplified with the Instant Error-prone PCR Kit. Subsequently, the amplified fragment was ligated into the pTP-LEU vector, and positive *E. coli* DH5α colonies were selected for sequencing. Primers used in this part were listed in Supplementary Table S1.

2.8 Statistical analysis

The data analysis was done by using GraphPad Prism 8 software. The significant differences were performed by unpaired two-tailed Student's t-tests. P-value < 0.05 was considered as statistically significant. The experiments were performed in triplicate, and the data was presented as the mean ± standard deviation (SD).

3 Results

3.1 Cloning and functional analysis of PfFAD2 from different perilla cultivars

Our previous study has found that the oil contents in the seeds of perilla cultivars PF40 and PF70 was different, and 343.51 mg/g and 281.34 mg/g of total oil contents were detected in these two cultivars, respectively (Zou et al., 2024). Meanwhile, the LA content in PF40 was higher than that in PF70 (Zou et al., 2024). By blasting the genome and transcriptome data of perilla using the previously identified *P. frutescens* fatty acid desaturase 2.1 (GenBank accession No. MZ747489.1) and 2.2 (GenBank accession No. MZ747499.1) as query, a total of four *FAD2* transcripts were identified from the high oil-producing perilla cultivar PF40 and the low oil-content perilla cultivar PF70. Their full-length gene sequences were cloned using 5'/3'-RACE PCR and designated as *PfFAD2-1*^{PF40}, *PfFAD2-1*^{PF70}, *PfFAD2-2*^{PF40} and *PfFAD2-2*^{PF70}. The full length of *PfFAD2-1*^{PF40} and *PfFAD2-1*^{PF70} are 927 bp, encoding a protein of 309 aa, while *PfFAD2-2*^{PF40} and *PfFAD2-2*^{PF70} are 1149 bp long, encoding a protein of 383 aa. Protein sequence alignment revealed that the amino acid sequences of PfFAD2-1^{PF40} and PfFAD2-1^{PF70} are identical, whereas PfFAD2-2^{PF40} and PfFAD2-2^{PF70} differ in two amino acids, located at positions 243 and 221 (Figure 1A). To preliminarily explore the function of these cloned genes, we screened for identical or similar sequences from different species by Protein BLAST program. Then multiple sequence alignment was performed on the selected sequences and cloned genes, and the phylogenetic tree was constructed by MEGA software (Tamura et al., 2021). The phylogenetic analysis indicated that all four genes were predicted to encode Δ12 fatty acid desaturases (Figure 1B).

3.2 Functional characterization of PfFAD2

OA, which constitutes approximately 80% of free FAs in the cytoplasm of *S. cerevisiae*, serves as the precursor for LA, indicating that intracellular free FA level is the primary limiting factor for LA synthesis (Klug and Daum, 2014). Previous research has demonstrated that lipid droplets in *S. cerevisiae* generate free FAs through the activity of steryl ester hydrolase Tgl1 and TAG lipases Tgl3 and Tgl5 (Athenstaedt and Daum, 2005). Additionally, free FAs within the cytoplasm of *S. cerevisiae* can re-enter the β-oxidation pathway catalyzed by fatty acyl-CoA synthetases Faa1 and Faa4, resulting in OA consumption (Leber et al., 2015).

Here, the *S. cerevisiae* strain was employed as a chassis for the functional verification of the cloned *PfFAD2* genes. Due to the relatively low OA content, the substrate for PfFAD2, in *S. cerevisiae*, metabolic engineering optimization was performed on *S. cerevisiae* BY4741 to ensure an adequate supply of OA as a precursor for PfFAD2. This was achieved by overexpressing *TGL1*, *TGL3*, and *TGL5* to promote the release of free FAs from lipid droplets and by knocking out *FAA1* and *FAA4* to reduce the activation of free FAs (Figure 2A). Consequently, according to the standard curve of OA (Supplementary Figure S2), the resulting engineered strain PF produced 70.55 ± 6.27 mg/L of OA, rendering it suitable as a chassis strain for FAD2 functional verification. Subsequently, the four cloned *PfFAD2* genes were individually constructed into the yeast expression vector pTP-LEU and transformed into the aforementioned engineered yeast strain PF. After fermentation, the target products were detected by using GC-MS equipment. By comparing the retention times of GC peaks and MS spectra with standards (Supplementary Figure S3), LA was detected in strains transformed with plasmids carrying *PfFAD2-2*^{PF40} and *PfFAD2-2*^{PF70} genes (Figure 2B). Conversely, no LA production was observed in strains transformed with *PfFAD2-1*^{PF40} and *PfFAD2-1*^{PF70} contained plasmids (Figure 2B). This indicates that *PfFAD2-2*^{PF40} and *PfFAD2-2*^{PF70} can catalyze LA production, while *PfFAD2-1*^{PF40} and *PfFAD2-1*^{PF70} lack this catalytic ability. Previous studies have shown that the deletion of consecutive amino acids near the active site can lead to reduced enzyme activity, whereas the deletion of consecutive amino acids at the C-terminus can result in a complete loss of enzyme activity (Arjomand et al., 2016; Zhu et al., 2016). Therefore, it is hypothesized that the functional loss of PfFAD2-1^{PF40} and PfFAD2-1^{PF70} may be related to the absence of a highly relevant functional region at its C-terminus associated with LA synthase catalytic function (Figure 1A).

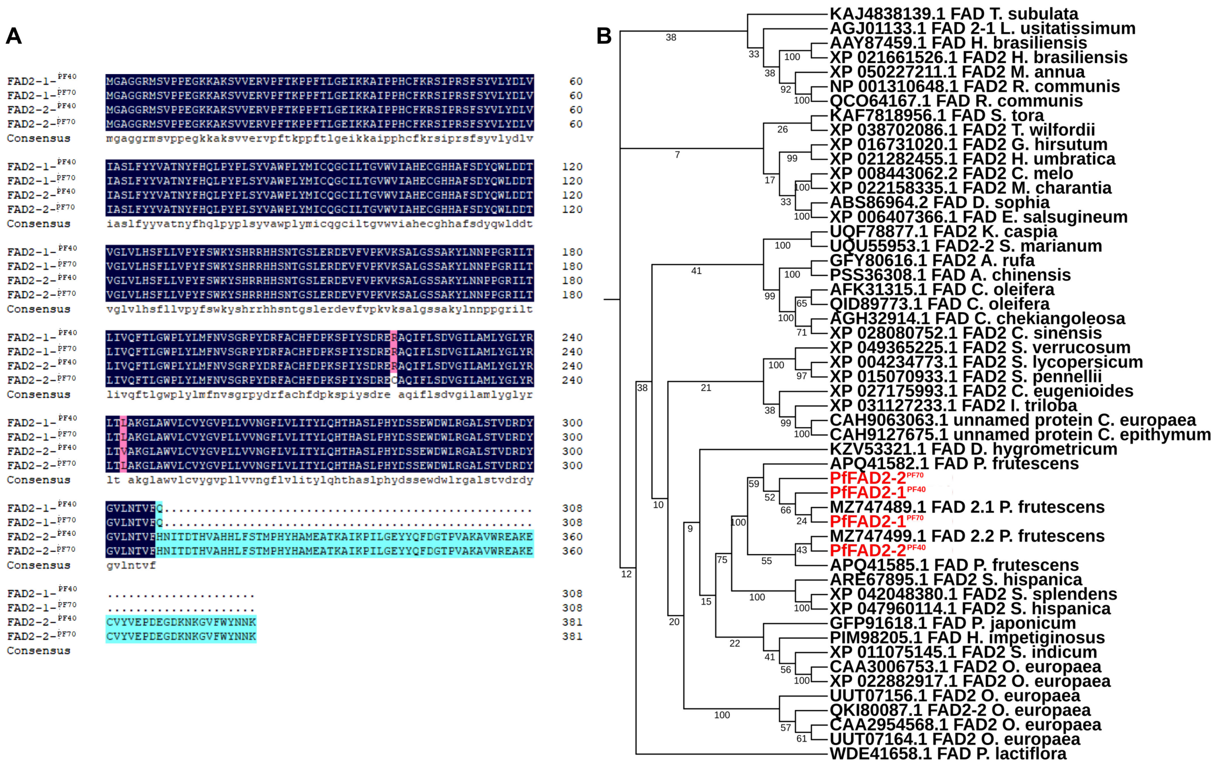
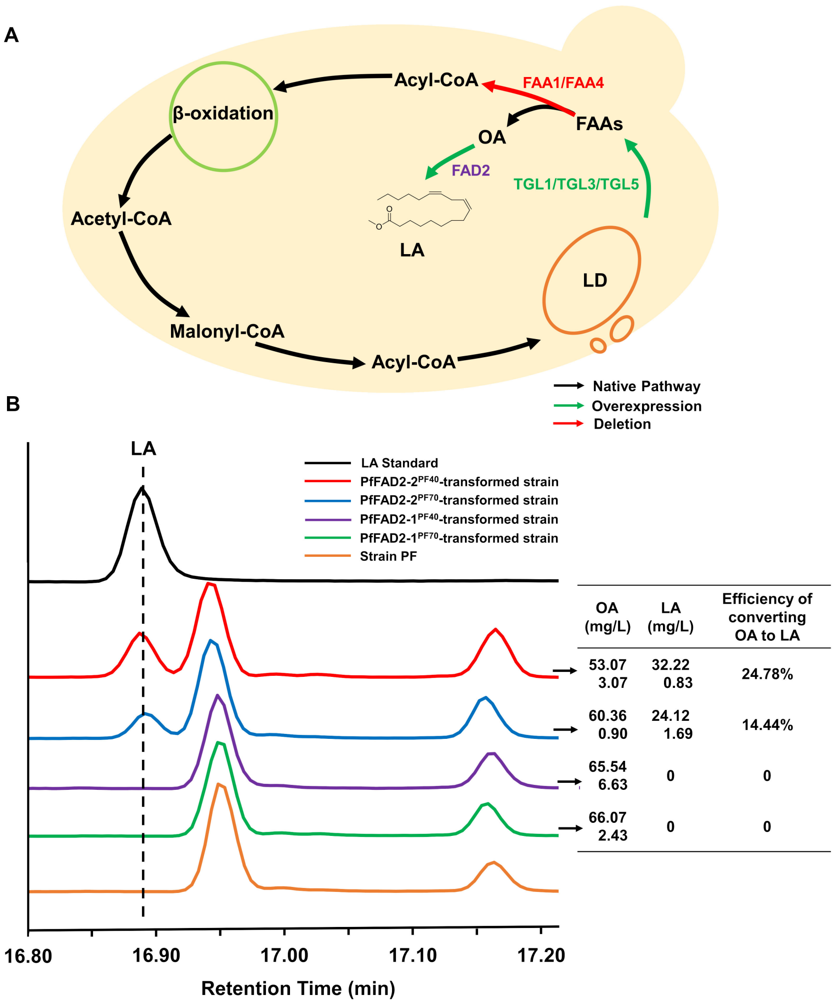


FIGURE 1
The sequence alignment and phylogenetic analysis of the cloned PfFAD2. **(A)** The sequence alignment of different PfFAD2 in PF40 and PF70 cultivars. The amino acids in blue color and cyan were identical, and the positions of different amino acids were shown in in purple color. **(B)** The phylogenetic analysis of PfFAD2.

3.3 Key amino acids residues affecting the catalytic ability of PfFAD2-2 in different perilla cultivars

During the functional characterization of PfFAD2-2, it was observed that PfFAD2-2^{PF40}, derived from the high oil-producing perilla cultivar PF40, catalyzed a higher yield of LA (32.22 ± 0.83 mg/L) compared to PfFAD2-2^{PF70} (24.12 ± 1.69 mg/L) from the low oil-producing cultivar PF70 (Figure 3). This suggests that PfFAD2-2^{PF40} possesses higher catalytic ability. This result was consistent with the finding that PF40 contained more LA than PF70 (Zou et al., 2024). Sequence alignment previously revealed two residues' differences between PfFAD2-2^{PF40} and PfFAD2-2^{PF70} protein sequences, located at positions 221 and 243 (Figure 1A). To investigate whether these two sites were responsible for the difference in catalytic ability, the two amino acid sites in the PfFAD2-2^{PF70} were mutated to match those of PfFAD2-2^{PF40}, resulting in two variants: PfFAD2-2^{L243V} and PfFAD2-2^{C221R}. These variants were then transformed into the engineered yeast strain to measure LA production. Results indicated that the strain transformed with PfFAD2-2^{L243V} exhibited no significant change in LA production (Figure 3). We guessed the main reason for this was that both Leu and Val are non-polar hydrophobic amino acids with similar side chain groups. And also, the L243V site mutation did not significantly change the protein's steric hindrance and surface activity. The amino acid properties (side chain structure, acidity, hydrophobicity) before and after the mutation were similar (Supplementary Figure S4). However, the strain transformed with PfFAD2-2^{C221R} produced LA increased to a comparable level (31.01 ± 0.13 mg/L) with those of PfFAD2-2^{PF40} (Figure 3). This suggests that the amino acid variation at position 221 is the primary factor contributing to the difference in catalytic ability between the LA synthases from the two perilla cultivars.

To further elucidate the mechanism by which the amino acid variation at position 221 affects the catalytic efficiency of PfFAD2-2, homology modeling and molecular docking were employed. Given



OA (mg/L)	LA (mg/L)	Efficiency of converting OA to LA
53.07 3.07	32.22 0.83	24.78%
60.36 0.90	24.12 1.69	14.44%
65.54 6.63	0	0
66.07 2.43	0	0

FIGURE 2
The functional characterization of PfFAD2 by yeast expression system. **(A)** The metabolic engineering of yeast for the accumulation of OA. The green color represented the overexpressed enzymes and the red color represented the deleted enzymes. **(B)** The detection and contents of OA and LA in different yeast strains. The left represented the GC profiles of LA standard and different engineered yeast strains, and the right represented the contents of OA, LA and the efficiency of converting oleic acid to LA in the four engineered yeast strains carrying different PfFAD2 enzymes.

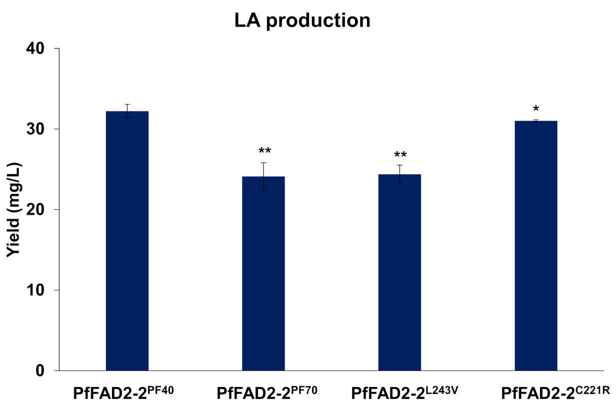


FIGURE 3
The LA production in yeast strains with different PfFAD2 transformation. PfFAD2-2^{L243V} and PfFAD2-2^{C221R} were mutated from PfFAD2-2^{PF70}. The asterisks indicate significant differences (**$p < 0.01$, *$p < 0.05$).

the absence of crystal structure template for FAD2, Alphafold 2 was utilized to construct the three-dimensional structure of PfFAD2-2 (Bryant et al., 2022). Following evaluation by SAVEs 6.1 (https://saves.mbi.ucla.edu/) (Supplementary Figure S5), OA was docked into the protein model (Figure 4A). The docking analysis and multiple sequences alignment (Supplementary Figure S6) revealed that the active site pocket of the PfFAD2-2 structure consists primarily of the β-strand beginning with the conserved region TXSXXXDEVFVP, the α-chain terminating with the conserved region GWPXYL, and the conserved amino acid region REXXEC at the C-terminus (Figure 4B), and the active pocket was shown in Figure 4C. The amino acid at position 221 is situated on the first α-chain near the active site pocket (Figure 4D). Compared to the cysteine at position 221 in PfFAD2-2^{PF70} (Figure 4E), the arginine in PfFAD2-2^{PF40} exhibits greater rigidity, enhancing the protein's structural stability (Figure 4F) (Mendes et al., 2015). Consequently, it is hypothesized that R221 in PfFAD2-2^{PF40} may augment its catalytic capability by stabilizing the protein. Additionally, homology modeling indicated that the active site pocket of PfFAD2-1^{PF40} and PfFAD2-1^{PF70} lack the conserved region REXXEC (Figure 1A; Supplementary Figure S7), which likely accounts for its lack of catalytic function.

3.4 Identification of additional amino acid sites affecting the catalytic ability of PfFAD2-2

Based on the above results, variations in amino acid sites were identified as the primary factors causing differences in the catalytic function of PfFAD2 among various perilla cultivars. To further elucidate the catalytic mechanism of PfFAD2, random mutation and functional verification methods were employed to identify additional amino acid sites affecting PfFAD2 catalytic efficiency. Initially, an error-prone PCR kit was utilized to amplify the *PfFAD2-2^{PF40}* gene, constructing a random mutation gene library (Lu et al., 2023). These mutated *PfFAD2-2^{PF40}* genes were then sequentially transformed into the engineered strain PF, followed by fermentation and LA production detection. The detection result revealed that among the thirty-two *PfFAD2-2^{PF40}* mutant transformants obtained, most showed no LA production or no significant change in LA yield, with only four strains exhibiting notable yield changes. Specifically, the mutants P41H, L171R, R177S, and Q115H demonstrated reduced LA production than the WT PfFAD2-2^{PF40} (Figure 5A), indicating that P41, Q115, L171, and R177 are crucial residues affecting PfFAD2-2^{PF40} catalytic capacity. Subsequently, the position of these four residues were checked (Figure 5B) and the reasons for the decreased FAD2 catalytic efficiency caused by these mutations were hypothesized.

By analyzing the distribution of these amino acid sites in the molecular docking model, it was found that L171 and R177 are located in the active pocket of PfFAD2-2^{PF40}. The L171R mutation (Figures 6A, B) narrowed the channel entering the active pocket due to arginine having a larger side chain than leucine, thereby affecting the entry of OA molecules. The R177S mutation (Figures 6C, D) was unfavorable for active pocket stability, as serine has a smaller side chain than arginine. P41 (Figures 6E, F) and Q115 (Figures 6G, H) are both distributed on the enzyme surface in key loop regions connecting α-chains. The P41H mutation reduced protein stability, as histidine is less rigid than proline. The Q115H mutation decreased protein solubility, given that histidine is less hydrophilic than glutamic acid. Consequently, these mutations all adversely affected the catalytic ability of FAD2.

4 Discussion

Over recent decades, crop breeding research has transitioned from traditional breeding methods to molecular breeding approaches (Bai et al., 2018). Currently, the primary objectives of oilseed crop breeding encompass enhancing seed yield, quality, and stress tolerance. Specifically, increasing seed yield has evolved from genomic screening and gene editing to the application of molecular breeding techniques (Stewart-Brown et al., 2019; Anders et al., 2021; Guo, 2021). Molecular breeding emphasizes identifying trait-influencing genes, examining the relationships between these genes and known genes, and investigating the molecular mechanisms by which these genes affect traits. Advances in high-throughput sequencing, big data mining, and synthetic biology have provided valuable tools for molecular breeding (Liu et al., 2022b). In this study, *FAD2* genes were extracted from existing genomic and transcriptomic data, and their functions were verified using engineered strains. Key amino acid sites affecting PfFAD2 function were identified through sequence alignment and molecular docking analysis. This research provides a reference for the cultivation and precision breeding of perilla varieties with high LA and oil yields.

To date, FAD2 enzymes involved in LA synthesis have been identified in various organisms, including animals, plants, fungi, and algae (Jiao and Zhang, 2013). The FAD2 enzyme was initially

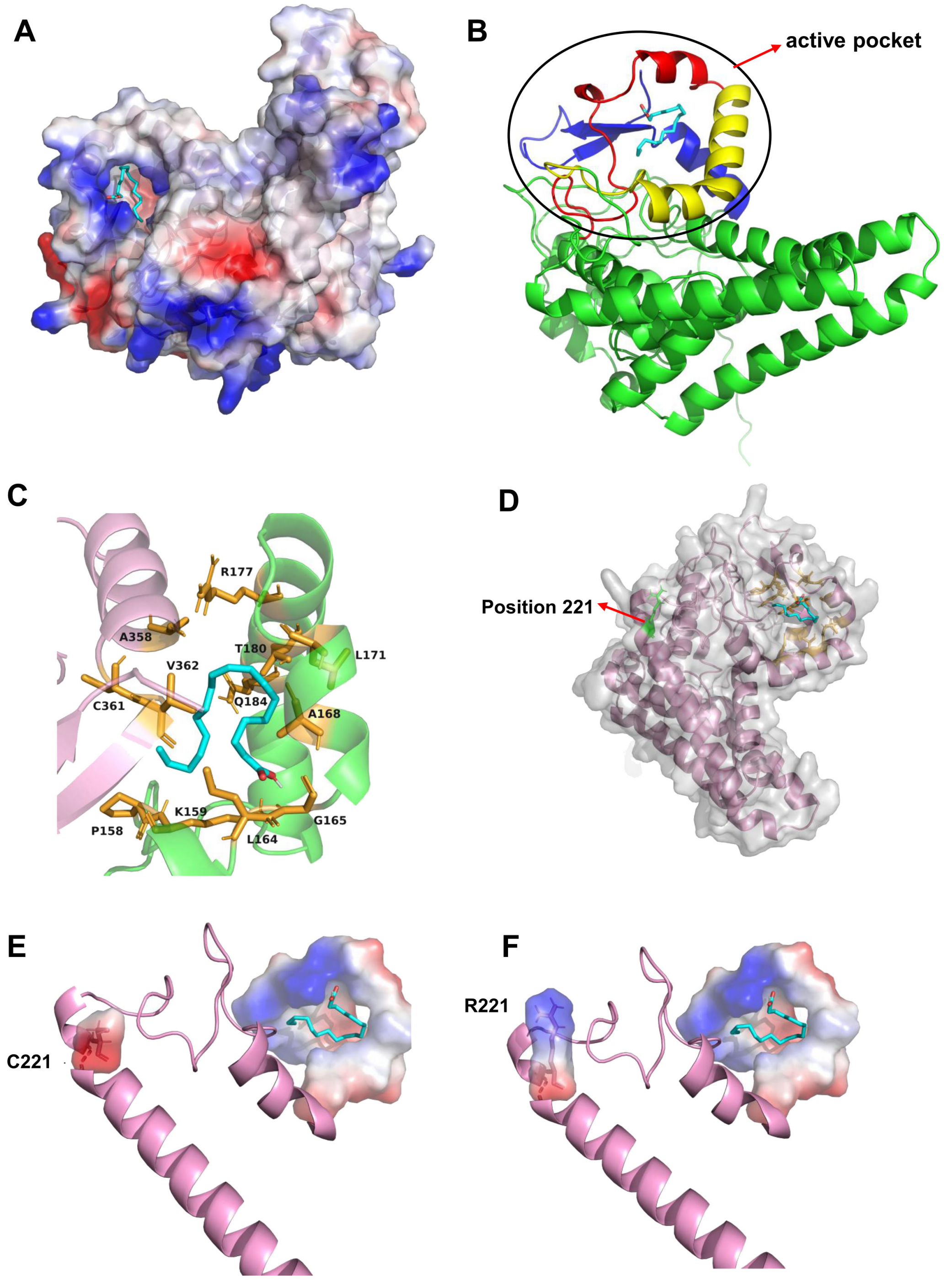


FIGURE 4
The molecular docking and residue 221 analysis of PfFAD2-2. **(A)** The docking of OA to the active pocket of PfFAD2-2. **(B)** The constitution of the active pocket of PfFAD2-2. The red represented the motif of TXSXXXDEVFVP, the yellow represented the motif of GWPXYL, and the blue represented the motif of REXXEC. **(C)** The active pocket of PfFAD2-2. **(D)** The position of residue 221 in PfFAD2-2. **(E, F)** The conformational difference of C221 and R221 in PfFAD2-2.

discovered in the model plant *Arabidopsis thaliana*. It is located on the endoplasmic reticulum and catalyzes the desaturation of OA using a phospholipid-derived acyl group, with the assistance of nicotinamide adenine dinucleotide (NADH), NADH-cytochrome b5 reductase, and cytochrome b5 (Okuley et al., 1994). Differences in *FAD2* gene sequences among species and cultivars result in variations in the expression patterns and functions of LA synthases (Jiao and Zhang, 2013). Therefore, it is essential to investigate the molecular factors affecting the catalytic efficiency of LA synthases. This study explored gene variations in *FAD2* from two different perilla cultivars. However, further research is needed to understand the expression and function of other unsaturated fatty acid

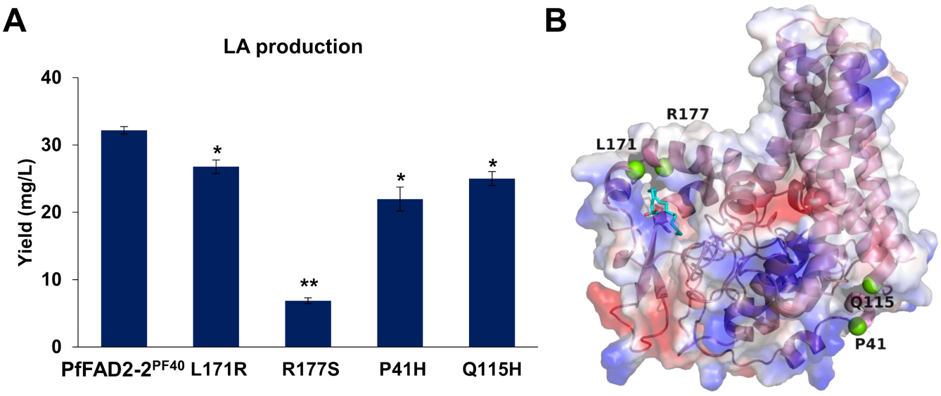


FIGURE 5
The effect of mutations in PfFAD2-2^{PF40} in LA production. **(A)** The LA production of different mutants of PfFAD2-2. All the mutants were mutated from PfFAD2-2^{PF40}. The asterisks indicate significant differences (**p < 0.01, *p < 0.05). **(B)** The position of P41, Q115, L171, and R177 residues in PfFAD2-2.

desaturases, such as FAD3 and FAD7, in perilla seeds (Arondel et al., 1992; Ma et al., 2015).

Amino acids form the foundation of protein structure and function. Throughout evolution, amino acids crucial to enzyme activity and stability have been conserved. Consequently, mutations at key amino acid residues can effectively alter the catalytic function of proteins (Li et al., 2024; Cai et al., 2024). For instance, manipulating the repetitive sequences of the *FAD2* gene can enhance the conversion rate of OA to LA in mouse cells (Chen et al., 2009). In this study, the conserved functional sequences of FAD2 from different perilla cultivars were compared to identify key amino acid residues responsible for variations in LA catalytic efficiency. Subsequently, homology modeling and molecular docking were employed to elucidate the structural basis for these changes in catalytic efficiency. Furthermore, to identify additional amino acid residues critical to the catalytic function of LA synthase, various FAD2-2 mutants were generated through random mutation. Functional validation and molecular docking analysis identified four important amino acid sites affecting LA production. This research provides crucial theoretical support for the genetic breeding of high oil-yielding perilla and the biosynthesis of LA.

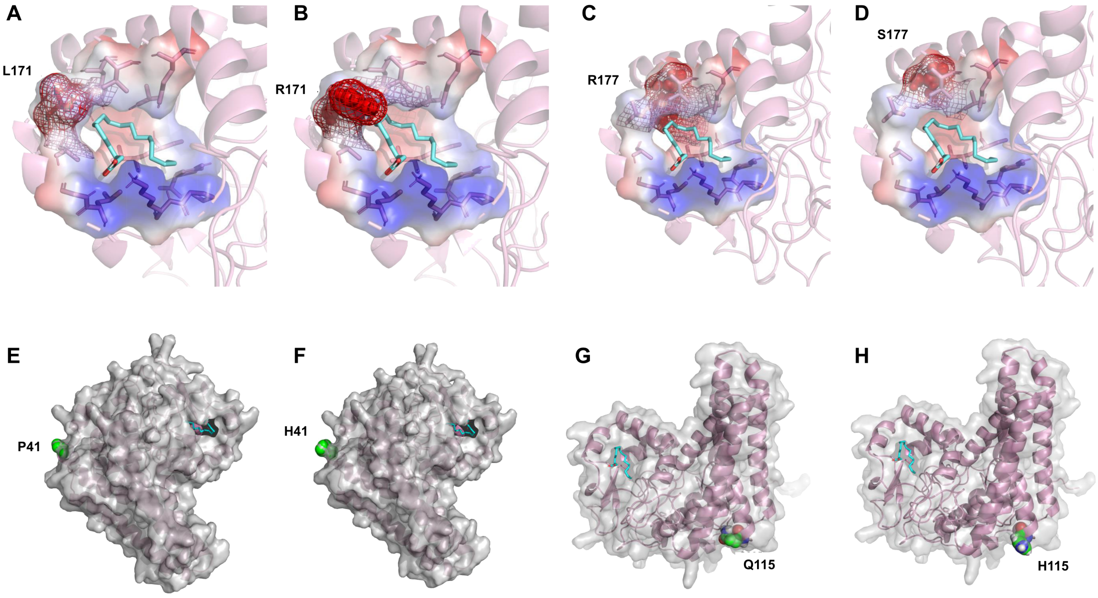


FIGURE 6
Mechanistic analysis of the catalytic activity of different PfFAD2-2 mutants. **(A–H)** Comparison of the conformational difference of L171, R177, P41 and Q115 mutants with WT.

Data availability statement

The datasets presented in this study can be found in NCBI GenBank database. The accession number(s) can be found below: PfFAD2-1^{PF40}, PQ178165; PfFAD2-1^{PF70}, PQ178166; PfFAD2-2^{PF40}, PQ178167; PfFAD2-2^{PF70}, PQ178168.

Author contributions

ZW: Investigation, Methodology, Visualization, Writing – original draft. ML: Methodology, Writing – original draft. XL: Investigation. JW: Investigation. GW: Investigation, Funding acquisition. QS: Resources, Writing – review & editing. TA: Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by Shandong Provincial Natural Science Foundation (No. ZR2021QC097) and Introduction and Cultivation Project for Young Creative Talents of Higher Education of Shandong Province.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1464388/full#supplementary-material

References

Alarcon-Gil, J., Sierra-Magro, A., Morales-Garcia, J. A., Sanz-SanCristobal, M., Alonso-Gil, S., Cortes-Canteli, M., et al. (2022). Neuroprotective and anti-inflammatory effects of linoleic acid in models of parkinson's disease: the implication of lipid droplets and lipophagy. *Cells* 11, 2297. doi: 10.3390/cells11152297

Anders, S., Cowling, W., Pareek, A., Gupta, K. J., Singla-Pareek, S. L., and Foyer, C. H. (2021). Gaining acceptance of novel plant breeding technologies. *Trends Plant Sci.* 26, 575–587. doi: 10.1016/j.tplants.2021.03.004

Arjomand, M. R., Habibi-Rezaei, M., Ahmadian, G., Hassanzadeh, M., Karkhane, A. A., Asadifar, M., et al. (2016). Deletion of loop fragment adjacent to active site diminishes the stability and activity of exo-inulinase. *Int. J. Biol. Macromol.* 92, 1234–1241. doi: 10.1016/j.ijbiomac.2016.08.039

Arondel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H. M., and Somerville, C. R. (1992). Map-based cloning of a gene controlling omega-3 fatty acid desaturation in Arabidopsis. *Science.* 258, 1353–1355. doi: 10.1126/science.1455229

Athenstaedt, K., and Daum, G. (2005). Tgl4p and Tgl5p, two triacylglycerol lipases of the yeast *Saccharomyces cerevisiae* are localized to lipid particles. *J. Biol. Chem.* 280, 37301–37309. doi: 10.1074/jbc.M507261200

Bai, S., Yu, H., Wang, B., and Li, J. (2018). Retrospective and perspective of rice breeding in China. *J. Genet. Genomics* 45, 603–612. doi: 10.1016/j.jgg.2018.10.002

Belury, M. A. (2023). Linoleic acid, an omega-6 fatty acid that reduces risk for cardiometabolic diseases: premise, promise and practical implications. *Curr. Opin. Clin. Nutr. Metab. Care* 26, 288–292. doi: 10.1097/MCO.0000000000000919

Bryant, P., Pozzati, G., and Elofsson, A. (2022). Improved prediction of protein-protein interactions using AlphaFold2. *Nat. Commun.* 13, 1265. doi: 10.1038/s41467-022-28865-w

Cai, X., Shi, X., Wang, J. Y., Hu, C. H., Shen, J. D., Zhang, B., et al. (2024). Enhancing the thermal stability and enzyme activity of ketopantoate hydroxymethyltransferase through interface modification engineering. *J. Agric. Food Chem.* 72, 13186–13195. doi: 10.1021/acs.jafc.3c09589

Chen, Q., Liu, Q., Wu, Z., Wang, Z., and Gou, K. (2009). Generation of fad2 transgenic mice that produce omega-6 fatty acids. *Sci. China C Life Sci.* 52, 1048–1054. doi: 10.1007/s11427-009-0143-z

Choudhary, A. K., and Mishra, G. (2021). Functional characterization and expression profile of microsomal FAD2 and FAD3 genes involved in linoleic and α-linolenic acid production in *Leucas cephalotes*. *Physiol. Mol. Biol. Plants.* 27, 1233–1244. doi: 10.1007/s12298-021-01016-z

Deng, X., Shi, B., Ye, Z., Huang, M., Chen, R., Cai, Y., et al. (2022). Systematic identification of Ocimum sanctum sesquiterpenoid synthases and (-)-eremophilene overproduction in engineered yeast. *Metab. Eng.* 69, 122–133. doi: 10.1016/j.ymben.2021.11.005

Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell Proteomics.* 13, 397–406. doi: 10.1074/mcp.M113.035600

Guo, Y. (2021). Molecular design for rice breeding. *Nat. Food* 2, 849. doi: 10.1038/s43016-021-00410-w

Guo, J., Zhou, W., Li, Y., Qin, W., Wei, D., Wang, F., et al. (2022). Combination of protein engineering and metabolic engineering to enhance (+)-nootkatone production in *Saccharomyces cerevisiae*. *Food Bioengineering* 1, 192–202. doi: 10.1002/fbe2.12021

Hamilton, J. S., and Klett, E. L. (2021). Linoleic acid and the regulation of glucose homeostasis: A review of the evidence. *Prostaglandins Leukot. Essent. Fatty Acids* 175, 102366. doi: 10.1016/j.plefa.2021.102366

Jiang, S., Tang, Y., Xiang, L., Zhu, X., Cai, Z., Li, L., et al. (2022). Efficient *de novo* assembly and modification of large DNA fragments. *Sci. China Life Sci.* 65, 1445–1455. doi: 10.1007/s11427-021-2029-0

Jiao, J., and Zhang, Y. (2013). Transgenic biosynthesis of polyunsaturated fatty acids: a sustainable biochemical engineering approach for making essential fatty acids in plants and animals. *Chem. Rev.* 113, 3799–3814. doi: 10.1021/cr300007p

Kanbar, A., Beisel, J., Gutierrez, M. T., Graeff-Hönninger, S., and Nick, P. (2023). Peruvian amaranth (kiwicha) accumulates higher levels of the unsaturated linoleic acid. *Int. J. Mol. Sci.* 24, 6215. doi: 10.3390/ijms24076215

Kim, T. J., Park, J. G., Kim, H. Y., Ha, S. H., Lee, B., Park, S. U., et al. (2020). Metabolite Profiling and Chemometric Study for the Discrimination Analyses of Geographic Origin of Perilla (*Perilla frutescens*) and Sesame (*Sesamum indicum*) Seeds. *Foods.* 9, 989. doi: 10.3390/foods9080989

Klug, L., and Daum, G. (2014). Yeast lipid metabolism at a glance. *FEMS Yeast Res.* 14, 369–388. doi: 10.1111/1567-1364.12141

Kuziora, M. A., Chalmers, J. H. J. R., Douglas, M. G., Hitzeman, R. A., Mattick, J. S., and Wakil, S. J. (1983). Molecular cloning of fatty acid synthetase genes from *Saccharomyces cerevisiae*. *J. Biol. Chem.* 258, 11648–11653. doi: 10.1016/S0021-9258(17)44277-3

Leber, C., Polson, B., Fernandez-Moya, R., and Da Silva, N. A. (2015). Overproduction and secretion of free fatty acids through disrupted neutral lipid recycle in *Saccharomyces cerevisiae*. *Metab. Eng.* 28, 54–62. doi: 10.1016/j.ymben.2014.11.006

Lee, K. R., Kim, K. H., Kim, J. B., Hong, S. B., Jeon, I., Kim, H. U., et al. (2019). High accumulation of γ-linolenic acid and Stearidonic acid in transgenic Perilla (*Perilla frutescens* var. *frutescens*) seeds. *BMC Plant Biol.* 19, 120. doi: 10.1186/s12870-019-1713-2

Li, J., Wang, L., Zhang, N., Cheng, S., Wu, Y., and Zhao, G. R. (2024). Enzyme and pathway engineering for improved betanin production in *saccharomyces cerevisiae*. *ACS Synth. Biol.* 13, 1916–1924. doi: 10.1021/acssynbio.4c00195

Liu, J., Dong, L., Duan, R., Hu, L., Zhao, Y., Zhang, L., et al. (2022a). Transcriptomic analysis reveals the regulatory networks and hub genes controlling the unsaturated fatty acid contents of developing seed in soybean. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.876371

Liu, S., Raman, H., Xiang, Y., Zhao, C., Huang, J., and Zhang, Y. (2022b). *De novo* design of future rapeseed crops: Challenges and opportunities. *Crop J.* 10, 587–596. doi: 10.1016/j.cj.2022.05.003

Lu, S., Deng, H., Zhou, C., Du, Z., Guo, X., Cheng, Y., et al. (2023). Enhancement of β-Caryophyllene Biosynthesis in *Saccharomyces cerevisiae* via Synergistic Evolution of β-Caryophyllene Synthase and Engineering the Chassis. *ACS Synth. Biol.* 12, 1696–1707. doi: 10.1021/acssynbio.3c00024

Ma, Q. P., Li, C., Wang, J., Wang, Y., and Ding, Z. T. (2015). Analysis of synonymous codon usage in FAD7 genes from different plant species. *Genet. Mol. Res.* 14, 1414–1422. doi: 10.4238/2015.February.13.20

Marangoni, F., Agostoni, C., Borghi, C., Catapano, A. L., Cena, H., Ghiselli, A., et al. (2020). Dietary linoleic acid and human health: Focus on cardiovascular and cardiometabolic effects. *Atherosclerosis* 292, 90–98. doi: 10.1016/j.atherosclerosis.2019.11.018

Mendes, M. I., Smith, D. E., Vicente, J. B., Tavares De Almeida, I., Ben-Omran, T., Salomons, G. S., et al. (2015). Small aminothiol compounds improve the function of Arg to Cys variant proteins: effect on the human cystathionine β-synthase p.R336C. *Hum. Mol. Genet.* 24, 7339–7348. doi: 10.1093/hmg/ddv431

Miao, X., Zhang, L., Hu, X., Nan, S., Chen, X., and Fu, H. (2019). Cloning and functional analysis of the *FAD2* gene family from desert shrub *Artemisia sphaerocephala*. *BMC Plant Biol.* 19, 481. doi: 10.1186/s12870-019-2083-5

Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E., and Browse, J. (1994). Arabidopsis *FAD2* gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* 6, 147–158. doi: 10.1105/tpc.6.1.147

Park, H., Sa, K. J., Lee, S., and Lee, J. K. (2022). Genetic variation of seed oil characteristics in native Korean germplasm of Perilla crop (*Perilla frutescens* L.) using SSR markers. *Genes Genomics* 44, 1159–1170. doi: 10.1007/s13258-022-01289-y

Pathak, N., Rai, A. K., Kumari, R., and Bhat, K. V. (2014). Value addition in sesame: A perspective on bioactive components for enhancing utility and profitability. *Pharmacogn. Rev.* 8, 147–155. doi: 10.4103/0973-7847.134249

Runguphan, W., and Keasling, J. D. (2014). Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid-derived biofuels and chemicals. *Metab. Eng.* 21, 103–113. doi: 10.1016/j.ymben.2013.07.003

Saini, R. K., and Keum, Y. S. (2018). Omega-3 and omega-6 polyunsaturated fatty acids: Dietary sources, metabolism, and significance - A review. *Life Sci.* 203, 255–267. doi: 10.1016/j.lfs.2018.04.049

Sana, M. A., Neila, B., and Nabil, K. (2018). Chemical composition and bioactive compounds of Cucumis melo L. seeds: Potential source for new trends of plant oils. *Process Saf. Environ. Prot.* 113, 68–77. doi: 10.1016/j.psep.2017.09.016

Stewart-Brown, B. B., Song, Q., Vaughn, J. N., and Li, Z. (2019). Genomic selection for yield and seed composition traits within an applied soybean breeding program. *G3 (Bethesda).* 9, 2253–2265. doi: 10.1534/g3.118.200917

Su, J., Wang, Y., Bai, M., Peng, T., Li, H., Xu, H. J., et al. (2023). Soil conditions and the plant microbiome boost the accumulation of monoterpenes in the fruit of Citrus reticulata 'Chachi'. *Microbiome* 11, 61. doi: 10.1186/s40168-023-01504-2

Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120

Wang, W., Pu, Y., Wen, H., Lu, D., Yan, M., Liu, M., et al. (2023). Transcriptome and weighted gene co-expression network analysis of jujube (*Ziziphus jujuba* Mill.) fruit reveal putative genes involved in proanthocyanin biosynthesis and regulation. *Food Sci. Hum. Wellness* 12, 1557–1570. doi: 10.1016/j.fshw.2023.02.003

Wu, D., Yang, S., Shang, Z., Xu, J., Zhao, D., Wang, H., et al. (2021). Genome-wide analysis of the fatty acid desaturase gene family reveals the key role of PfFAD3 in α-linolenic acid biosynthesis in perilla seeds. *Front. Genet.* 12. doi: 10.3389/fgene.2021.735862

Yang, K., Tang, Y., Xue, H., Ji, X., Cao, F., Li, S., et al. (2023). Enrichment of linoleic acid from yellow horn seed oil through low temperature crystallization followed by urea complexation method and hypoglycemic activities. *Food Sci. Biotechnol.* 33, 145–157. doi: 10.1007/s10068-023-01327-9

Yang, B., Zhou, Y., Wu, M., Li, X., Mai, K., and Ai, Q. (2020). [amp]]omega;-6 Polyunsaturated fatty acids (linoleic acid) activate both autophagy and antioxidation in a synergistic feedback loop via TOR-dependent and TOR-independent signaling pathways. *Cell Death Dis.* 11, 607. doi: 10.1038/s41419-020-02750-0

Zhu, S., Canales, A., Bedair, M., and Vik, S. B. (2016). Loss of Complex I activity in the *Escherichia coli* enzyme results from truncating the C-terminus of subunit K, but not from cross-linking it to subunits N or L. *J. Bioenerg. Biomembr.* 48, 325–333. doi: 10.1007/s10863-016-9655-y

Zou, X., Zhang, K., Wu, D., Lu, M., Wang, H., and Shen, Q. (2024). Integrated analysis of miRNA, transcriptome, and degradome sequencing provides new insights into lipid metabolism in perilla seed. *Gene.* 895, 147953. doi: 10.1016/j.gene.2023.147953

Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 September 2024
DOI 10.3389/fpls.2024.1418957

OPEN ACCESS

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Penghua Shu,
Xuchang University, China
Juanjuan Yu,
Henan Normal University, China

*CORRESPONDENCE
Zhimao Chao
✉ chaozhimao@163.com

RECEIVED 17 April 2024
ACCEPTED 23 August 2024
PUBLISHED 26 September 2024

CITATION
Liu Z, Cheng Y, Xiang Y and Chao Z (2024) Unveiling the potential applications of buds of *Lonicera japonica* Thunb. var. *chinensis* (Wats.) Bak based on *in vitro* biological activities, bio-active components, and potential applications coupled to targeted metabolomics.
Front. Plant Sci. 15:1418957.
doi: 10.3389/fpls.2024.1418957

Unveiling the potential applications of buds of *Lonicera japonica* Thunb. var. *chinensis* (Wats.) Bak based on *in vitro* biological activities, bio-active components, and potential applications coupled to targeted metabolomics

Zhenying Liu[1,2], Yunxia Cheng[1], Yaoting Xiang[1] and Zhimao Chao[1*]

[1]Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China,
[2]Graduate School of China Academy of Chinese Medical Science, Beijing, China

Introduction: The buds of *Lonicera japonica* Thunb. var. *chinensis* (Wats.) Bak, commonly named red honeysuckle, have attracted attention because of their bright colors. However, owing to the lack of systematic studies, the potential applications of red honeysuckle are not clear, and its development and utilization have not been well known.

Methods: In this study, compared with the buds of *L. japonica* Thunb. (honeysuckle), the potential applications of red honeysuckle were explored based on biological activities, bio-active components, and sensory flavor combined with widely targeted metabolomics.

Results: As a result, *in vitro* tests showed that it had a stronger antioxidant and a stronger inhibitory effect on the growth of *Escherichia coli* and *Staphylococcus aureus*. There was no cytotoxicity on LPS-induced RAW264.7 cells in its aqueous extract using the CCK-8 method. Moreover, it also had a stronger effect on inhibiting the expression of inflammatory factors such as interleukin-6 (IL-6), tumor necrosis factor-α (TNF-α), and interleukin-1β (IL-1β). The content of its bio-active components chlorogenic acid and cynaroside was significantly higher ($p \leq 0.001$) than that of green honeysuckle. Widely targeted metabolomics analysis revealed that 4 volatile metabolites, such as (*E*)-4-hexene-1-ol and pyrazole, and 21 non-volatile metabolites, such as macranthoside B and oleanolic acid-3-*O*-glc(1-2)-(ara)-28-*O*-glucoside ester, were specific in red honeysuckle. Interestingly, 14 specific terpenoid metabolites were triterpenoid saponins, indicating a stronger biological activity in red honeysuckle. The sensory flavor analysis showed that the red honeysuckle had a stronger herbal and lighter floral flavor.

Discussion: In conclusion, red honeysuckle had great development value with potential applications in medicines, foods, beverages, pigment additives, and health products.

KEYWORDS

red honeysuckle, biological activities, bio-active components, metabolomics, sensory flavor

1 Introduction

The bud of *Lonicera japonica* Thunb. (Caprifoliaceae), named green honeysuckle, is known for its medicinal and edible properties. Many studies demonstrated that it has a series of biological (antioxidant, antibacterial, and anti–inflammatory) activities (Wang M. et al., 2021; Mu et al., 2022). It has gained great popularity with the increase of people's awareness of its healthcare applications. For example, it could be made into functional foods and dietary supplements to prevent chronic diseases such as arthritis, diabetes mellitus, and cardiopathy (Kirina et al., 2021; Feng et al., 2022). It could be made into herbal teas, beverages, juices, and wines because of its strong fragrance and miraculous healing effect, especially in summer (Belyaeva et al., 2021; Fang et al., 2022).

In recent years, mutants of the *Lonicerae* plants have been gradually developed and are favored by consumers (Carrillo-Galván et al., 2020; Eltaib and Alzain, 2023). Among them, *L. japonica* Thunb. var. *chinensis* (Wats.) Bak is a representative variety that has been cultivated in large areas. Its bud, named red honeysuckle, has been getting more and more attention because of its bright colors and strong fragrance. For example, Yu made a preliminary study on the red honeysuckle anthocyanin and formation mechanism (Yu, 2013). Tan et al. explored the role of anthocyanin synthase coding genes in the synthesis of red honeysuckle anthocyanins (Tan et al., 2022). Li et al. studied the breeding system and pollination biology of red honeysuckle (Li et al., 2019). Li et al. determined the optimal harvesting time of red honeysuckle by measuring the flowers' thousand bud weight of different flowering phase (Li et al., 2013). Yuan et al. compared the anthocyanin content of red and traditional green honeysuckle (Yuan et al., 2014). Hu et al. determined the active compounds in different organs of red honeysuckle cultivated in saline soil by high-performance liquid chromatography (HPLC) (Hu et al., 2017). However, its biological activities and chemical composition have been less studied, resulting in its potential development value remaining unclear. Therefore, systematic studies on the biological activities and chemical composition of red honeysuckle are necessary for its development and commercialization as a new food raw material.

In this study, in order to comprehensively explore the potential applications of red honeysuckle and provide data support for the development of related products, its biological activities (such as antioxidant, antibacterial, and anti-inflammatory), bio-active components (such as chlorogenic acid, cynaroside, and rutin), metabolites (such as volatile and non-volatile metabolites), and sensory flavor were analyzed and compared with those of green honeysuckle.

2 Materials and methods

2.1 Chemicals and reagents

6-Hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox), ferric chloride hexahydrate, 2,2′-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt (ABTS), Folin–Ciocalteu phenol, and 1,1-diphenyl-2-picrylhydrazyl (DPPH) were obtained from Sigma–Aldrich (St. Louis, MO, USA). All standards (purities ≥ 98%) of the HPLC method were purchased from Push Bio–Technology Co., Ltd. (Chengdu, China). The internal standard (*p*-xylene-d10) of the gas chromatography–mass spectrometry (GC-MS) method was purchased from Yien Chemical Technology Co., Ltd. (Shanghai, China). HPLC- and UPLC-grade solvents were from Fisher Scientific (Waltham, MA, USA). Analytical-grade solvents were purchased from Fuyu Fine Chemical Co., Ltd. (Tianjin, China). The mouse interleukin-6 (IL-6) (SEKM-0007), interleukin-1β (IL-1β) (SEKM-0002), tumor necrosis factor-α (TNF-α) (SEKM-0034) ELISA kit, and PBS buffer were purchased from Solarbio Science & Technology Co., Ltd. (Beijing, China). The anthocyanin content detection kit was from Zike Biological Technology Co., Ltd (Shenzhen, China). Dexamethasone, fetal bovine serum (FBS), lipopolysaccharides (LPS), and Griess reagent were purchased from Sigma–Aldrich Co., Ltd. DMEM basic (1×) and PBS pH 7.4 basic (1×) were from ThermoFisher Biochemical Products (Beijing) Co., Ltd. The Cell Counting Kit-8 (CCK-8, Biosharp) was from Beijing Labgic Technology Co., Ltd. Penicillin–streptomycin solution was manufactured by HyClone Laboratories.

2.2 Bacterial strains and culture media

All bacterial strains, including *Escherichia coli* (ATCC 25922), *Staphylococcus aureus* (ATCC 29213), and *Bacillus subtilis* (ATCC 6633), were acquired from Guangdong Microbial Culture Collection Center (Guangzhou, China).

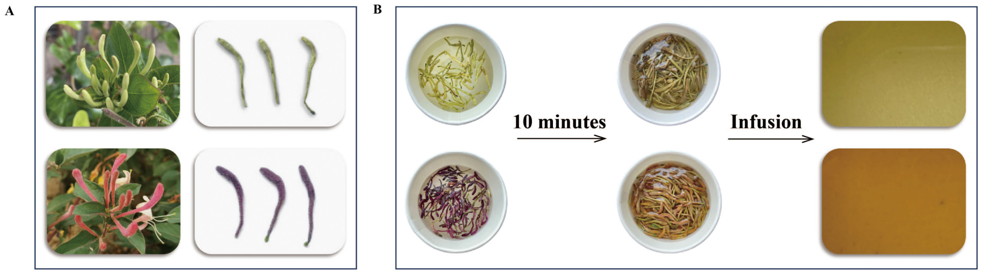


FIGURE 1
The appearance **(A)** and tea infusion **(B)** of green and red honeysuckles.

LB broth culture medium and agar were obtained from Sangon Biotech (Shanghai) Co., Ltd. (Shanghai, China).

2.3 Sample preparation

As shown in Figure 1, the buds of *L. japonica* Thunb. (green honeysuckle) were collected from a planting base located in Julu County of Hebei Province, and the buds of *L. japonica* Thunb. var. *chinensis* (Wats.) Bak (red honeysuckle) were obtained from a planting base located in Henan Academy of Agricultural Sciences. The buds were collected in the morning of flowering and then dried so that their moisture content met the requirements of the Chinese Pharmacopoeia. Then, they were crushed into powder (particle size: 40 mesh) by a high-speed crusher and placed in a dryer for subsequent use.

2.4 *In vitro* antioxidant activity determination

The *in vitro* antioxidant activities were determined by three assays based on a previous study (Liu et al., 2023), namely, ABTS and DPPH radical scavenging activity and ferric reducing antioxidant power (FRAP). Correspondingly, measurement of the absorbance of the solution was performed at 734, 517, and 593 nm, respectively. Their results were calculated by a calibration curve and expressed as µTrolox equivalent per gram (TE µ/g).

2.5 *In vitro* antibacterial activity determination

The *in vitro* antibacterial activity of the aqueous extracts was estimated against three bacteria strains listed above using the agar plate-counting technique.

2.5.1 Cell suspension preparation

LB liquid medium (3 mL) was added to four bacterial culture tubes. Single colonies of the above three bacteria strains were selected and transformed to the liquid medium, and the other one was chosen as a blank control. Then, they were incubated for 15 h in an oscillator (37°C, 200 rpm).

2.5.2 Aqueous solution extraction

Samples (30 g) and 300 mL of water were added to a casserole, soaked for 40 min, and boiled for approximately 30 min. Then, the mixture was filtered with gauze. Next, the filtrate was poured out and repeatedly decocted. After the two filtrates were combined, they were freeze-dried to obtain aqueous extracts. Then, the extracts were placed in centrifuge tubes and placed under ultraviolet lamps for UV sterilization for more than 30 min for later use.

2.5.3 Plate coating count

Firstly, sterile PBS solution was added to dilute the bacterial solution to 1×10^6 colony-forming units (CFU)/mL. Next, 2 mL of diluted bacterial solution was added into the corresponding centrifuge tubes, while the extracts were not added in the blank group. They were then incubated at 37°C for 24 h. Similarly, the bacterial solution was diluted 10 times. After uniformly coating LB solid medium with 100 µL of diluent, 18 h of incubation at 37°C was performed. Finally, the colonies were photographed and counted, and the antibacterial rate was calculated by the following formula:

$$\text{Bacterial concentration (bc)} = \text{colony number} \times \text{dilution factor} \times 10$$

$$\text{Antibacterial rate (\%)} = [1 - (\text{experimental}/\text{control})_{\text{bc}} \times 100\,\%]$$

2.6 *In vitro* anti-inflammatory activity determination

2.6.1 Cell culture and viability

CCK-8 was used to detect the cell viability of two aqueous extracts of honeysuckle at different concentrations in RAW264.7 cells. In detail, RAW264.7 cells were cultured in DMEM (Dulbecco's minimal essential medium) containing 10% FBS (fetal bovine serum) and 1% double antibiotic in an incubator with 37°C and 5% CO_2 for 24 h. Then, the different concentrations of two aqueous extracts (1,000, 800, 400, 200, 100, 50, 25, 12.5, and 0 µg/mL) were added for the next 24 h. Subsequently, CCK-8 (10 µL/

well) was added and incubated for 2 h. After shaking, the absorbance of each well was measured at 450 nm and cell viability was calculated.

$$\textbf{Cell viability}\ \% = (A_S - A_B)/(A_C - A_B)$$

where A_S indicates the absorbance of the sample, A_B indicates the absorbance of the blank, and A_C indicates the absorbance of the control.

2.6.2 NO production by Griess

The NO inhibitory effects of two aqueous extracts in LPS-induced RAW264.7 cells were detected by Griess assay. Specifically, the control group, model group (LPS), DXMS group (dexamethasone), and sample groups (400, 800, and 1,000 μg/mL of two aqueous extracts were named G-L, G-M, G-H, R-L, R-M, and R-H, respectively) were set up, respectively. The RAW264.7 cells (5×10^4 cells/well) were cultured for 24 h. Next, according to the setting, DMEM and drug solutions were given, respectively. Then, 2 h later, LPS (last concentration was 1 μg/mL) was added to the model and administration groups and then cultured for 24 h. Finally, NO production in the supernatant solution was measured using the Griess assay.

2.6.3 Inflammatory factor determination by ELISA

The experimental group was the same as that in section 2.6.2. The RAW264.7 cells (5×10^5 cells/well) were cultured for 24 h. Then, the different concentrations of two aqueous extracts and DXMS solutions were added. After 2 h, the LPS solution (last concentration was 1 μg/mL) was added to the model and administration groups for 24-h culturing. Next, the supernatants were collected and used for ELISA to determine the concentrations of three inflammatory factors, namely, IL-6, IL-1β, and TNF-α. According to the instructions of the ELISA kit, the sample and the standard, the biotinized antibody working solution, the enzyme binding working solution, the chromogenic substrate, and the termination solution were added successively and incubated according to the requirements, respectively. Finally, the absorbance was measured at 450 nm using an enzyme-labeled instrument.

2.6.4 Real time-quantitative polymerase chain reaction

The groups were the same as those in section 2.6.2. Firstly, RAW264.7 cells (1×10^6 cells/well) were cultured for 24 h. Then, the different solutions were added according to the group setting. After 24 h, LPS (last concentration was 1 μg/mL) was added to the groups except for the control group for the next 6 h. Finally, the supernatant was removed and 500 μL of TRIzol reagent was added to extract the total RNA of cells, and the relative mRNA expressions of IL-6, IL-1β, and TNF-α were measured by real time-quantitative polymerase chain reaction (rt-qPCR).

Then, the samples were reverse-transcribed and amplified according to the reverse transcription kit SynScript®III RT SuperMix for qPCR. The cDNA product obtained by reverse transcription was threefold diluted and amplified as a template for rt-qPCR. The primer sequences are shown in Supplementary Table S1.

2.7 Bio-active components analysis

The sample powder was extracted with 50% methanol and analyzed using an HPLC system (Shimadzu, Kyoto, Japan). The chromatography condition was consistent with previous studies (Liu et al., 2023). Next, 10 bio-active components, namely, chlorogenic acid, cynaroside, neochlorogenic acid, caffeic acid, isochlorogenic acid A, isochlorogenic acid B, isochlorogenic acid C, cryptochlorogenic acid, lutin, and quercetin, were analyzed qualitatively and quantitatively by reference to the standard solution.

2.8 GC-MS analysis

2.8.1 Sample preparation

The 0.5-g honeysuckle powder was weighed and transferred to a head-space vial. Then, the vial was sealed and placed at 60°C for 5 min. Next, a 120-μm divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) filter (Agilent, Palo Alto, CA, USA) was exposed to the head space for 15 min at 60°C. Subsequently, the samples were analyzed by the GC-MS system.

The GC-MS conditions were similar to those in another study (Wang H. et al., 2021), except for the heating procedure. In this study, the initial temperature setting was 40°C and was held for 3.5 min. Then, the temperature was increased to 100°C at 10°C/min, 180°C at 7°C/min, 280°C at 25°C/min, and held for 5 min.

2.8.2 Qualitative and quantitative analyses

Volatile metabolites were identified using the NIST 14 standard library with a similarity of >80% based on the characteristic ion chromatogram and mass spectrum value (Li et al., 2023). Then, the concentration of each metabolite was calculated by reference to the internal standard (*p*-xylene-d10).

2.8.3 Sensory flavor annotation

In order to fully compare the difference in flavor between the two kinds of honeysuckle, the flavor sensory annotation was carried out on the different metabolites selected. With reference to a series database, including the TGSCIS database (The Good Scents Company Information System, http://www.thegoodscentscompany.com), the perflavory database (http://perflavory.com/), the odour database (http://www.odour.org.uk/odour/index.html), or a related database (http://foodflavorlab.cn/#/home), combined with the comparison of the relative content of metabolites focused on the same flavor, the odor difference of the two kinds of honeysuckle was finally explored.

2.9 UPLC-QTRAP-MS/MS analysis

2.9.1 Sample analysis

The 0.5-g honeysuckle powder was extracted with 70% methanol. Then, the extract was centrifuged and filtered for subsequent analysis.

The ExionLC™ AD UPLC system (AB, Framingham, MA, USA) coupled to an Agilent SB–C18 column (100 mm × 2.1 mm, 1.8 μm) was used to achieve chromatographic separation. The solvent system was water (A) and acetonitrile (B) with 0.1% formic acid. The gradient program was as follows: 0 min, 95% A, 0–9 min, 95%–5% A, 9–10 min, 5% A, 10–11.10 min, 5%–95% A, and 11.10–14 min, 95% A. The column temperature was 40°C. The injection volume was 2 μL. Finally, the effluent was further analyzed by electron spray ionization (ESI)–triple quadrupole–linear ion trap (QTRAP)–MS. Moreover, the MS condition was similar to the other report (Wang H. et al., 2021).

2.9.2 Qualitative and quantitative analyses

The data were processed as described elsewhere (Zou et al., 2020). All metabolites were annotated combined with the Metware in–house MS^2 spectral tag (MS2T) database (Wuhan Metware Biotechnology Co., Ltd., Wuhan, China) (http://www.metware.cn/) and public metabolite databases including MassBank (http://www.massbank.jp), KNAPSAcK (http://kanaya.naist.jp/KNApSAcK), human metabolome database (HMDB, http://www.hmdb.ca), MoTo DB (http://www.ab.wur.nl/moto), and METLIN (http://metlin.scripps.edu/index.php). For the quantitative analysis, the result was expressed as the mass spectrum peak area.

2.10 Statistical analysis

The multivariate statistical analysis was performed using the SIMCA 14.1 software (Umerics, Umeå, Sweden). All results were visualized by R (http://www.r-project.org), chiplot (https://www.chiplot.online/), and GraphPad Prism 9.0.0 (GraphPad, San Diego, CA, USA).

3 Results

3.1 *In vitro* antioxidant activity analysis

Three methods, including ABTS, DPPH, FRAP, based on different principles were used to evaluate the antioxidant activities due to the efficiency of antioxidants in biological systems may differ. Their results are shown in Figure 2.

For red honeysuckle, the average antioxidant activity evaluated by ABTS, DPPH, and FRAP was 93.44 TE μM/g, 114 TE μM/g, and 6,195 TE μM/g, respectively, while for green honeysuckle, the corresponding values were 77.58 TE μM/g, 105.1 TE μM/g, and 5,766 TE μM/g. The results of FRAP were significantly higher than those of ABTS and DPPH, which may be due to the different principles of these methods. The FRAP method aims to determine the ability of a substance to reduce iron ions, whereas ABTS and DPPH aim to determine the ability to remove corresponding free radicals.

3.2 *In vitro* antibacterial activity analysis

The ability of the aqueous extract of two kinds of honeysuckle to inhibit the growth of *E. coli*, *S. aureus*, and *B. subtilis* was evaluated. The average colony number and the antibacterial rate at the corresponding dilution ratio are shown in Table 1, and the inhibitory effect on colony growth at the same dilution ratio (10^3) is shown in Figure 3.

Compared with the control group, for *S. aureus*, the antibacterial rate of green honeysuckle was 75.52%, and that of red honeysuckle was as high as 99.27%, and there was a significant difference ($p < 0.01$) between them. For *E. coli*, the antibacterial rate of green honeysuckle was 98.69% and that of red honeysuckle was 99.89%, but there was no significant difference. For *B. subtilis*, both of them showed significant antibacterial effect, and the antibacterial rate was 100%.

This result showed that there were fewer colonies in the medium corresponding to the red honeysuckle under the same dilution ratio. *B. subtilis* had no colony growth under the inhibition of the two extracts. Combining the antibacterial rate of *S. aureus* and *E. coli*, it could be speculated that the inhibitory effects of red honeysuckle were stronger than those of green honeysuckle.

3.3 *In vitro* anti-inflammatory activities analysis

Using the CCK-8 method (Figures 4A, B), the green and red honeysuckle aqueous extracts showed no cytotoxicity to RAW264.7 cells. All the cell viabilities of different groups were more than 100% when the drug concentrations were 12.5–1,000

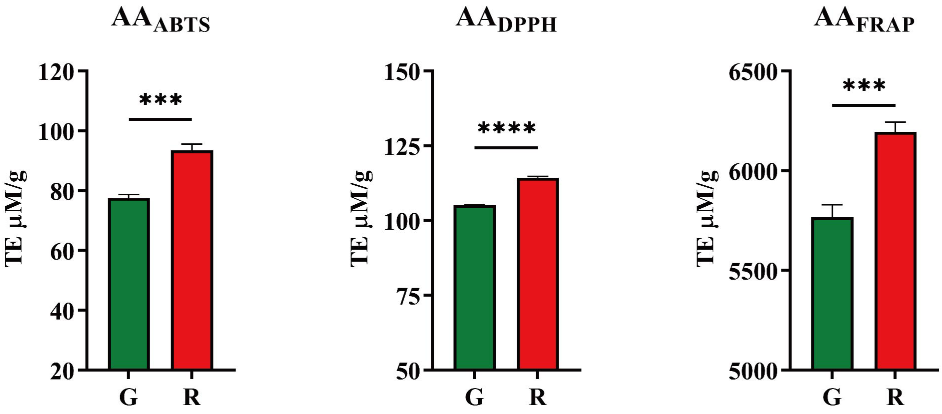


FIGURE 2
Antioxidant activities of honeysuckles. G, green honeysuckle; R, red honeysuckle. ***$p \leq 0.001$, ****$p \leq 0.0001$.

TABLE 1 Plate counting test results ($n = 3$).

Strain	Group	Colony number	Dilution ratio	Concentration(CFU/mL)	Antibacterial rate (%)
Staphylococcus aureus	Control	143	10^3	1.4×10^6	/
	G	35	10^3	3.5×10^5	75.52
	R	105	10^1	1.1×10^4	99.27
Escherichia coli	Control	253	10^3	2.5×10^6	/
	G	33	10^2	3.3×10^4	98.69
	R	27	10^1	2.7×10^3	99.89
Bacillus subtilis	Control	60	10^3	6.0×10^5	/
	G	0	10^1	$<10^2$	100.00
	R	0	10^1	$<10^2$	100.00

The number of colonies was the counting result of each group at the corresponding dilution ratio. The counting principle was based on the national standard GB4789.2-2016, and the number of colonies was selected to be between 30 and 300 CFU for colony counting.

μg/mL, which indicated that the two aqueous extracts were safe to a certain extent.

Moreover, the anti-inflammatory effects on LPS-induced RAW264.7 cells are shown in Figures 4C–I. NO secretion results (Figure 4C) showed that compared with the control group, when LPS activated inflammatory response, the concentration of NO in the model group increased to 19.29 μM significantly. However, the concentrations of NO in the administration groups decreased significantly ($p < 0.01$), which indicated that the two honeysuckle aqueous extracts could inhibit the secretion of NO and had certain anti-inflammatory effects, with a certain concentration dependence. Additionally, the inhibitory effect of red honeysuckle was stronger than that of the traditional green honeysuckle.

The concentrations of major inflammatory factors IL-6, TNF-α, and IL-1β in the supernatant are shown in Figures 4D–F. After LPS induction, they were significantly higher in the model group than in the control group, and inflammation occurred. Their concentrations decreased and showed a certain concentration-dependent relationship ($p < 0.01$). Moreover, the relative mRNA expressions of IL-6, TNF-α, and IL-1β were measured by rt-qPCR, and the results are shown in Figures 4G–I. Similarly, when LPS was induced, the expression of inflammatory factors increased in varying degrees. There was a significant increase in the model group when compared with the control. Their expressions in G-L, G-M, G-H, R-L, R-M, and R-H were decreased after administration ($p < 0.01$). Interestingly, these two extracts showed clear

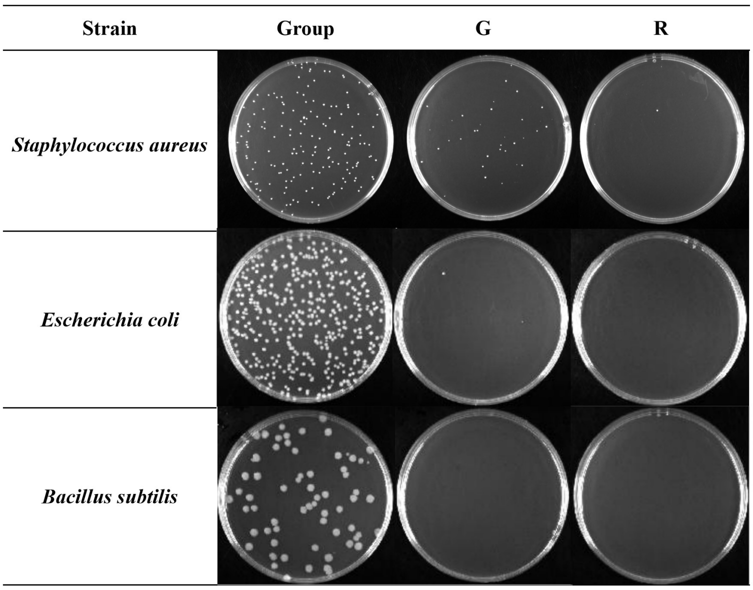


FIGURE 3
Antibacterial effect of two kinds of honeysuckle under the same dilution ratio (10^3). G, green honeysuckle; R, red honeysuckle.

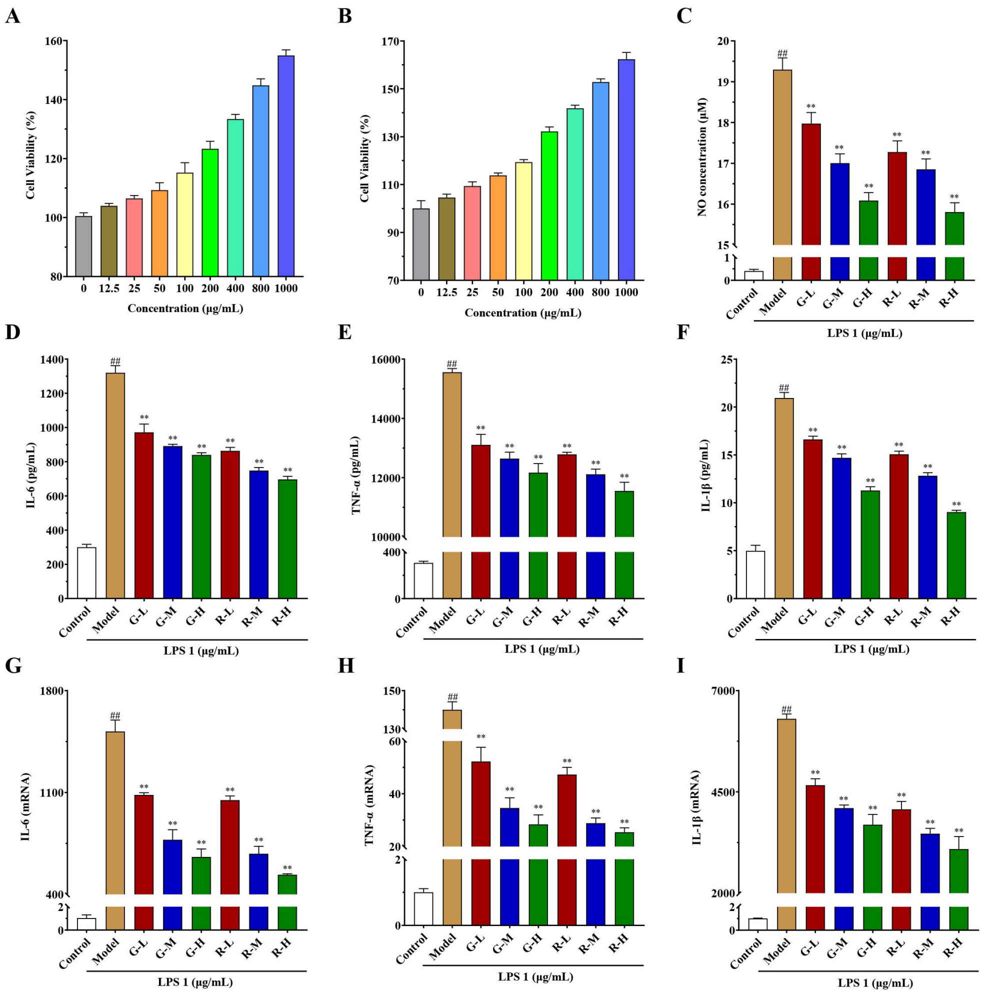


FIGURE 4
In vitro anti-inflammatory activity results of honeysuckle aqueous extracts on the LPS-induced RAW 264.7 cells. **(A)** Cell viability of green honeysuckle. **(B)** Cell viability of red honeysuckle. **(C)** The concentrations of NO in different groups. **(D–F)** The concentrations of inflammatory factors in the supernatant, IL-6, TNF-α, and IL-1β, respectively. **(G–I)** The relative mRNA expression of IL-6, TNF-α, and IL-1β, respectively. G-L, G-M, and G-H: the concentrations of green honeysuckle were 400, 800, and 1,000 μg/mL, respectively; R-L, R-M, and R-H: the concentrations of red honeysuckle were 400, 800, and 1,000 μg/mL, respectively; $n = 3$; #$p < 0.05$, ##$p < 0.01$, compared to the control group; **$p < 0.01$ compared to the model group.

concentration-dependent relationships. The higher the drug concentration, the lower levels of mRNA expression.

These results suggested that two honeysuckle aqueous extracts could inhibit the expression of NO, IL-6, TNF-α, and IL-1β induced by LPS dose dependently, which was consistent with previous studies (Guo et al., 2021; Su et al., 2021). It also showed that they had good anti-inflammatory activity, and red honeysuckle showed better inhibitory effect.

3.4 Bio-active component analysis

The results of physio–chemical indicators showed that there might be a significant difference in chemical composition. In this study, 10 bio-active components were analyzed and their content is shown in Figure 5. Compared with green honeysuckle, red honeysuckle had a lower content of neochlorogenic acid, rutin, and isochlorogenic acids A, B, and C, and a higher content of chlorogenic acid, cryptochlorogenic acid, caffeic acid, cynaroside, and quercetin. There were significant differences in these components except for isochlorogenic acid C. For chlorogenic acid, its content (6.45% ± 0.06%) in red honeysuckle was higher than that (4.57% ± 0.03%) in green honeysuckle. Similarly, the average content of cynaroside (0.16%) in red honeysuckle was higher than that (0.13%) in green honeysuckle. It was indicated that red honeysuckle had higher medicinal value. In particular, the quercetin content of red honeysuckle was significantly higher ($p \leq 0.0001$) than that of green honeysuckle.

3.5 Volatile metabolite analysis

3.5.1 Metabolite profiling analysis

The separate (Supplementary Figure S1A) and overlaid (Supplementary Figure S1B) analysis of the total ion

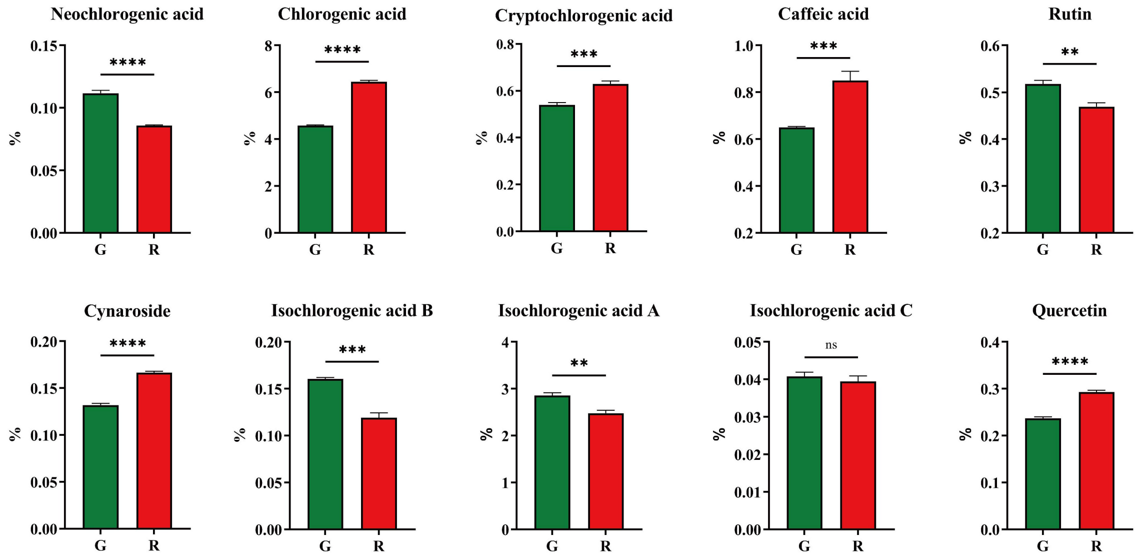


FIGURE 5
Ten main compounds of honeysuckles. G, green honeysuckle; R, red honeysuckle. ns, no significance, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$.

chromatogram (TIC) showed that the instrument had a good stability, which improved repeatability and reliability.

A total of 605 volatile metabolites were identified (Figure 6A), namely, 103 terpenoids, 102 heterocyclic compounds, 100 esters, 62 hydrocarbons, 56 ketones, 48 aldehydes, 48 alcohols, 36 aromatics, 13 amines, 11 acids, 8 phenols, 6 nitrogen compounds, 4 halogenated hydrocarbons, 4 sulfur compounds, and 4 others.

3.5.2 PCA and OPLS–DA

Principal component analysis (PCA) could preserve as much of the original information as possible with fewer comprehensive variables (Clark et al., 2019). Based on these metabolites, a non-supervised PCA model was used for comprehensive analysis. The total contribution rate of 86.03% indicated the validity and stability of this model. As shown in Figure 6B, green and red honeysuckles could be distinguished well, which indicated that their volatile metabolites were significantly different.

However, PCA cannot ignore within-group errors (Liu et al., 2022). Therefore, supervised orthogonal partial least squares discrimination analysis (OPLS–DA) was used to further analyze the difference of this variables. It is known that Q^2 is an important parameter in determining the predictive capability of the model. The model is excellent and stable when Q^2 is >0.9. The validity of the OPLS–DA model was assessed due to high Q^2 (0.991), R^2X (0.756), and R^2Y (0.998) (Supplementary Figures S2A, B). In Figure 6C, it could be seen that green and red honeysuckles were clearly distinguished and this model was stable, reliable, and effective. The above result provided an excellent explanation for the metabolic variations.

3.5.3 Differential metabolite screening

The above results revealed the different accumulation patterns of volatile metabolites. To further identify the differential volatile metabolites between green and red honeysuckles, 605 metabolites were analyzed according to the principle of variable important in projection (VIP) > 1 and $|\log_2(\text{fold change})| \geq 1$ (Figure 6D). A total of 178 differential metabolites were screened out (Supplementary Table S2), of which 81 metabolites such as Δ–limonene, copaene, and (*Z*)–4–heptenal were upregulated, but 97 metabolites such as benzyl benzoate, methyl dodecanoate, and 3–nonen–5–one were downregulated (Figure 6E).

These differential metabolites were classified into 12 categories (Figure 6F). Most of the upregulated metabolites were terpenoids (32.10%) and heterocyclic compounds (20.99%), and most of the downregulated metabolites were esters (23.71%) and terpenoids (14.43%). This comparative analysis revealed that heterocyclic compounds were significantly more abundant and esters were significantly less abundant in red honeysuckle than those in green honeysuckle. For terpenoids, their upregulation and downregulation in red honeysuckle might lead to a significant flavor difference.

As shown in Table 2 and Figure 7, some differential metabolites were found only in green or red honeysuckles. For instance, 1,4–pentanediol and DL–camphoroquinone were only detected in green honeysuckle, and (*E*)–4–hexen–1–ol,pyrazole, 1,2,4–trimethyl-benzene, and (3R,6R)–2,2,6–trimethyl–6–vinyltetrahydro–2H–pyran–3–ol were only detected in red honeysuckle.

3.5.4 Sensory flavor analysis

Aroma is closely related to the taste and nutritional quality of plants and foods and is an important factor affecting its characteristics (Banožić et al., 2020). According to sensory analysis, differential metabolites can be characterized based on their flavor characteristics. In this study, the sensory flavor characteristics of 178 differential metabolites were annotated (Figure 6G). The top 10 sensory flavor characteristics of differential metabolites were fruity,

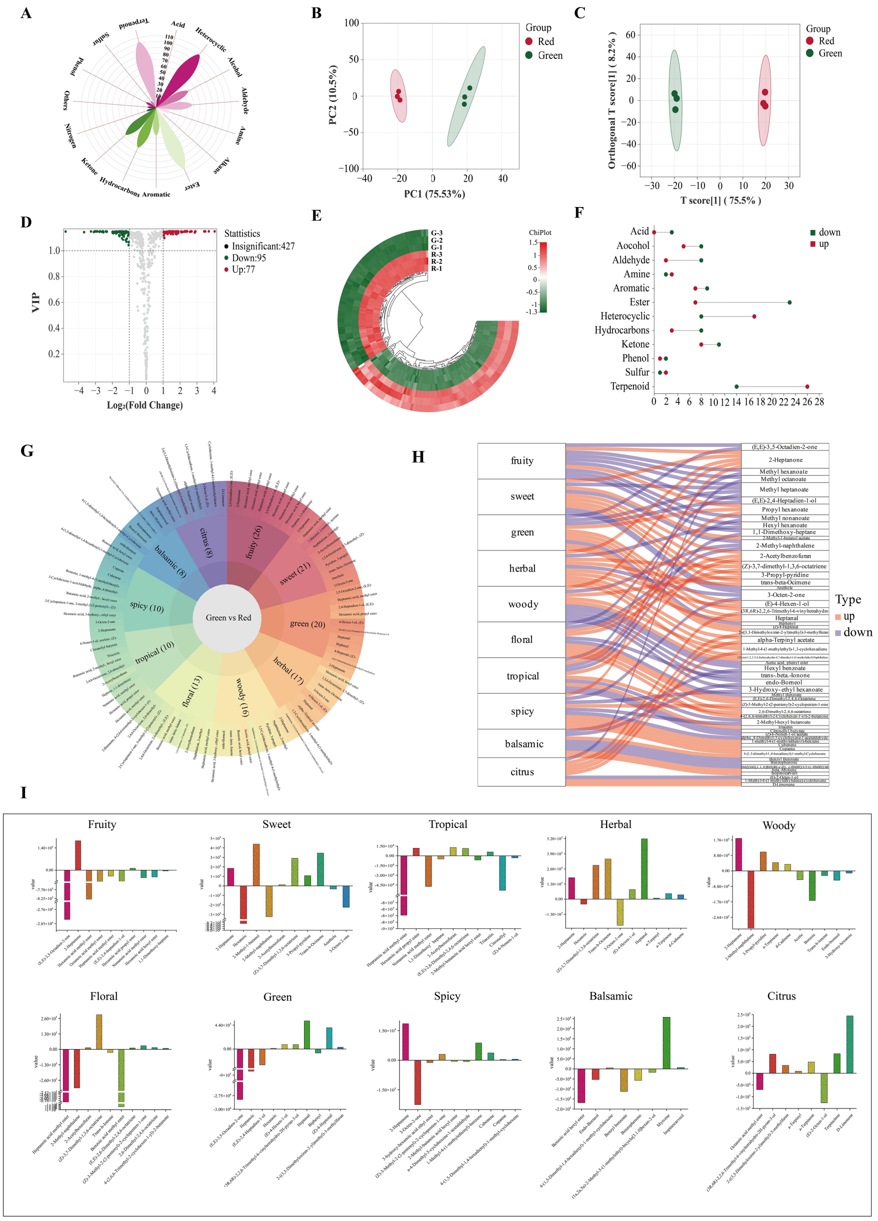


FIGURE 6
Volatile metabolite analysis. **(A)** Classification of all metabolites. **(B)** PCA score plot. **(C)** OPLS–DA score plot. **(D)** Volcano plot. **(E)** Heatmap of differential metabolites; **(F)** Up– and downregulation of differential metabolites. **(G)** Flavor sunburst. **(H)** Flavor sankey. **(I)** Metabolite content variation among flavors.

sweet, green, herbal, woody, floral, tropical, spicy, balsamic, and citrus. Fruity, sweet, and green were annotated by 26, 21, and 20 differential metabolites, respectively.

The association was analyzed between the top 10 differential metabolites (ranked by VIP value) and the corresponding flavor characteristics (Figure 6H). For downregulated metabolites, (*E*,*E*)−3,5−octadien−2−one decreased fruity and green; methyl heptanoate decreased fruity, sweet, green, and floral; and 2−methyl−naphthalene decreased fruity, sweet, green, and floral. For upregulated metabolites, 2−heptanone increased fruity, sweet, herbal, woody, and spicy; propyl hexanoate increased fruity, green, and tropical; and (*Z*)−3,7−dimethyl−1,3,6−octatriene increased sweet, herbal, and floral.

Each flavor characteristic was associated with both up− and downregulated metabolites; thus, it was impossible to assess the final difference. Therefore, it was necessary to study the specific content difference of each metabolite corresponding to the same flavor. In Figure 6I, red honeysuckle had stronger herbal, spicy, and citrus flavors, and lighter fruity, sweet, green, woody, floral, tropical, and balsamic flavors. These results fully explained the reason why the herbal flavor of red honeysuckle was stronger than that of green honeysuckle.

TABLE 2 Unique volatile metabolites of green and red honeysuckle.

Group	Compounds	Class	Formula	CAS
Green	1,4-Pentanediol	Alcohol	$C_5H_{12}O_2$	626-95-9
	dl-Camphoroquinone	Terpenoids	$C_{10}H_{14}O_2$	10373-78-1
Red	(*E*)-4-Hexen-1-ol	Alcohol	$C_6H_{12}O$	928-92-7
	Pyrazole	Heterocyclic compound	$C_3H_4N_2$	288-13-1
	1,2,4-Trimethyl-benzene	Aromatics	C_9H_{12}	95-63-6
	(3*R*,6*R*)-2,2,6-Trimethyl-6-vinyltetrahydro-2H-pyran-3-ol	Heterocyclic compound	$C_{10}H_{18}O_2$	14009-71-3

3.6 Non-volatile metabolite analysis

3.6.1 Metabolite profiling analysis

The TIC of green and red honeysuckles was performed in both positive and negative ion modes. The results showed the characteristic metabolites of both red and green honeysuckles and are found in Supplementary Figures S3A, B. It could be seen that the signal stability was good according to the overlap TIC of QC samples (Supplementary Figures S3C, D), which provided an important guarantee for repeatability and reliability.

There were 1,971 non–volatile metabolites identified (Figure 8A), namely, 426 flavonoids, 284 phenolic acids, 205 lipids, 186 amino acid and derivatives, 161 alkanoids, 153 terpenoids, 120 lignans and coumarins, 100 organic acids, 89 nucleotides and derivatives, 12 quinones, 6 tannins, and 229 others.

3.6.2 PCA and OPLS–DA

The PCA score (Figure 8B) showed good separation and indicated a significant difference. The permutation plot of OPLS–DA showed that this model had effective predictive ability and did not experience overfitting (Supplementary Figures S4A, B). The OPLS–DA score plots (Figure 8C) also showed significant differences similar to PCA.

3.6.3 Differential metabolite screening

VIP > 1 and $|\log_2(\text{fold change})| \geq 1$ were used as reference to screen different metabolites (Figure 8D). A total of 691 differential metabolites were screened out (Figure 8E). Flavonoids and phenolic acids accounted for 25.04% and 15.77% in honeysuckle, respectively. There were 388 upregulated metabolites such as peonidin–3–*O*–glucoside, vanillic acid, and tryptamine, and 303 downregulated metabolites such as secologanin, methyl *p*–coumarate, and geniposidic acid.

Notably, the content of procyanidin B5, a proanthocyanidin metabolite, in green honeysuckle was significantly higher ($p \leq 0.05$) than that in red honeysuckle. In contrast, the content of peonidin–3–*O*–glucoside, an anthocyanin metabolite, in red honeysuckle was significantly higher ($p \leq 0.05$) than that in green honeysuckle. Plants produce proanthocyanidins as polyphenols in their flowers, fruit cores, and leaves. Anthocyanins are water–soluble natural pigments in plants. As a result of their excellent anti-oxidation properties and health benefits such as preventing cancer, protecting vision, slowing aging, and beautifying the skin, proanthocyanidins and anthocyanins are becoming more popular (Zhao et al., 2023). Anthocyanins are colored such as purple centaurin, orange geranium, and blue–purple delphinium, but proanthocyanidins are colorless (Khoo et al., 2017). Peonidin–3–*O*–glucoside, a red anthocyanin metabolite, is methylated centaurea. In this study, it was considered to be a key metabolite to show the significant difference between green and red honeysuckles due to its high accumulation pattern. In addition, this study has demonstrated its ability to induce anabolic impacts on bone by increasing osteoblast proliferation and differentiation and altering the osteoblast epigenome (Ren et al., 2021). Therefore, it is suggested that the

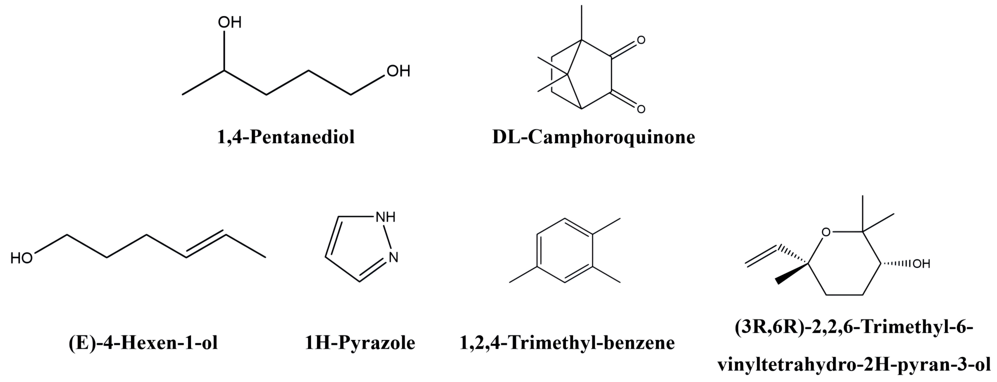


FIGURE 7
The chemical structure of unique volatile metabolites in green and red honeysuckles.

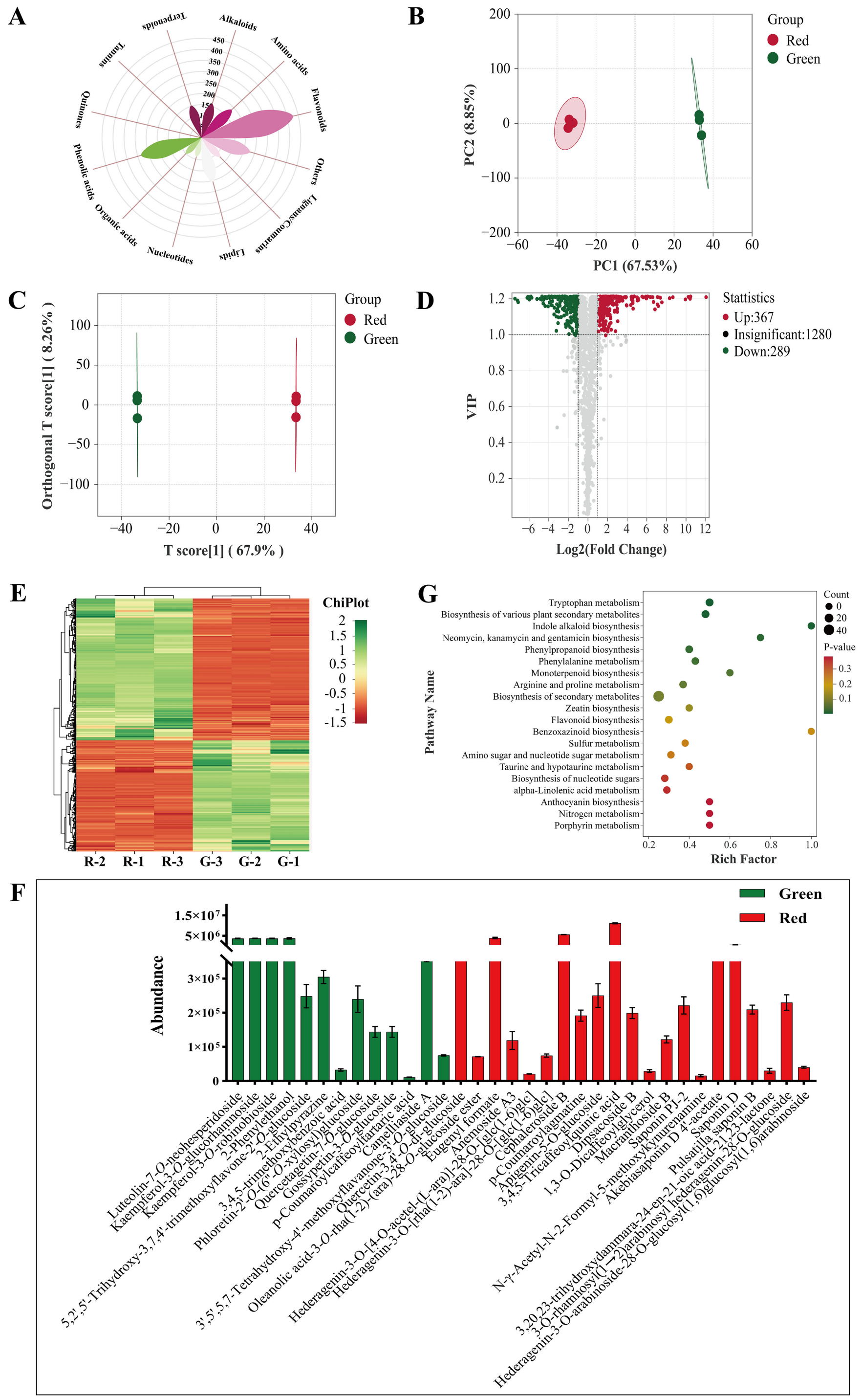


FIGURE 8
Non-volatile metabolite analysis. **(A)** Classification of all metabolites. **(B)** PCA score plot. **(C)** OPLS–DA score plot. **(D)** Volcano plot. **(E)** Heatmap of differential metabolites. **(F)** Specific metabolites. **(G)** KEGG enrichment analysis.

red honeysuckle can be further developed into a series of osteoporosis medicines and health products.

Moreover, the content of some differential metabolites in red honeysuckle was higher than that in green honeysuckle, and their pharmacological activities significantly enhance the development value of red honeysuckle. For instance, verbascoside has been shown to alleviate pneumococcal pneumonia and inhibit PLY-mediated cytotoxicity (Zhao et al., 2016). It could also effectively inhibit the activity of SARS-CoV-2 main protease and protect against COVID-19 (Xiao et al., 2022). Similarly, gallocatechin has significant anti-inflammatory activity in a pleurisy model in mice (Siebert et al., 2021). Protocatechuic acid methyl ester has been found to effectively attenuate the F-induced changes in oxidative stress, inflammation, and apoptosis markers (Ameeramja and

TABLE 3 Unique non–volatile metabolites of green or red honeysuckle.

Group	Compounds	Class	Formula	CAS
Green	Luteolin-7-*O*-neohesperidoside	Flavonoids	$C_{27}H_{30}O_{15}$	25694-72-8
	Kaempferol-3-*O*-glucorhamnoside	Flavonoids	$C_{27}H_{30}O_{15}$	40437-72-7
	Kaempferol-3-*O*-robinobioside	Flavonoids	$C_{27}H_{30}O_{15}$	17297-56-2
	2-Phenylethanol	Phenolic acids	$C_8H_{10}O$	60-12-8
	5,2',5'-Trihydroxy-3,6,7,4'-tetramethoxyflavone-2'-*O*-glucoside	Flavonoids	$C_{25}H_{28}O_{14}$	–
	2-Ethylpyrazine	Alkaloids	$C_6H_8N_2$	13925-00-3
	3,4,5-Trimethoxybenzoic acid	Phenolic acids	$C_{10}H_{12}O_5$	118-41-2
	5,7-Dihydroxy-2-(4-hydroxyphenyl)-6-[(2S,4R,5S)-3,4,5-trihydroxy-6-[[(2R,3S,5R)-3,4,5-trihydroxy-6-methyloxan-2-yl]oxymethyl]oxan-2-yl]chromen-4-one	Flavonoids	$C_{27}H_{30}O_{14}$	–
	Phloretin-2'-*O*-(6''-*O*-xylosyl)glucoside	Flavonoids	$C_{26}H_{32}O_{14}$	–
	Quercetagetin-7-*O*-glucoside	Flavonoids	$C_{21}H_{20}O_{13}$	548-75-4
	Gossypetin-3-*O*-glucoside	Flavonoids	$C_{21}H_{20}O_{13}$	777080-67-8
	p-Coumaroylcaffeoyltartaric acid	Phenolic acids	$C_{22}H_{18}O_{11}$	–
	Kaempferol-3-*O*-(6''-rhamnosyl-2''-glucosyl)glucoside	Flavonoids	$C_{33}H_{40}O_{20}$	135095-52-2
	3',5',5,7-Tetrahydroxy-4'-methoxyflavanone-3'-*O*-glucoside	Flavonoids	$C_{22}H_{24}O_{12}$	–
Red	Quercetin-3,4'-*O*-di-glucoside	Flavonoids	$C_{27}H_{30}O_{17}$	29125-80-2
	Oleanolic acid-3-*O*-glc(1-2)-(ara)-28-*O*-glucoside ester	Terpenoids	$C_{47}H_{76}O_{17}$	–
	Eugenyl formate	Others	$C_{11}H_{12}O_3$	10031-96-6
	3-*O*-Rhamnosyl(1→2)arabinosyl-23-hydroxylup-20(29)-en-28-oic acid	Terpenoids	$C_{41}H_{66}O_{12}$	129724-84-1
	Hederagenin-3-*O*-[4-*O*-acetel-(*L*-ara)]-28-*O*-[glc(1-6)glc]	Terpenoids	$C_{49}H_{78}O_{19}$	–
	Gypsogenin-3-*O*-[rha(1-2)-ara]-28-*O*-[glc(1-6)glc]	Terpenoids	$C_{53}H_{84}O_{22}$	–
	Cephaleroside B	Terpenoids	$C_{53}H_{86}O_{22}$	–
	p-Coumaroylagmatine	Alkaloids	$C_{14}H_{20}N_4O_2$	7295-86-5
	Apigenin-5-*O*-glucoside	Flavonoids	$C_{21}H_{20}O_{10}$	28757-27-9
	3,4,5-Tricaffeoylquinic acid	Phenolic acids	$C_{34}H_{30}O_{15}$	86632-03-3
	Dipsacoside B	Terpenoids	$C_{53}H_{86}O_{22}$	33289-85-9
	1,3-*O*-Dicaffeoylglycerol	Phenolic acids	$C_{21}H_{20}O_9$	–
	Macranthoside B	Terpenoids	$C_{53}H_{86}O_{22}$	146100-02-9
	Saponin PJ-2	Terpenoids	$C_{53}H_{86}O_{22}$	–
	N-γ-Acetyl-*N*-2-formyl-5-methoxykynurenamine	Amino acids and derivatives	$C_{13}H_{16}N_2O_4$	52450-38-1
	Akebia saponin D 4'-acetate	Terpenoids	$C_{49}H_{78}O_{19}$	126778-93-6
	Akebia saponin D	Terpenoids	$C_{47}H_{76}O_{18}$	39524-08-8
	3-*O*-α-*L*-Arabinopyranosylhederagenin-28-β-*D*-glucopyranoside	Terpenoids	$C_{41}H_{66}O_{13}$	39524-13-5
	3,20,23-Trihydroxydammara-24-en-21-oic acid-21,23-lactone-3-O-glc(1-2)-{ara(1-3)}-rha	Terpenoids	$C_{47}H_{76}O_{17}$	–

(Continued)

TABLE 3 Continued

Group	Compounds	Class	Formula	CAS
	3-*O*-rhamnosyl(1→2)arabinosyl hederagenin-28-*O*-glucoside	Terpenoids	$C_{47}H_{76}O_{17}$	–
	Hederagenin-3-*O*-arabinoside-28-*O*-glucosyl(1,6)glucosyl (1,6)arabinoside	Terpenoids	$C_{52}H_{84}O_{22}$	–

Perumal, 2017; Ameeramja et al., 2018). For other different metabolites, the content of red honeysuckle can also effectively indicate that the red honeysuckle has stronger antibacterial, anti-inflammatory, and antioxidant activities.

Additionally, some non–volatile metabolites were found only in green or red honeysuckles (Table 3). For example, there were 14 differential metabolites (such as lonicerin, biorobin, and kaempferol –3–*O*–glucorhamnoside) found only in green honeysuckle and 21 differential metabolites (such as quercetin–3,4'–*O*–di–glucoside, cephaleroside B, and apigenin–5–*O*–glucoside) found only in red honeysuckle. Their relative contents are shown in Figure 8F. Their chemical structures are shown in Figures 9 and 10, respectively. It is because of these unique metabolites that the difference in efficacy and function of green and red honeysuckles is recognized.

3.6.4 KEGG classification and enrichment analysis

Plants and foods contain metabolites that coordinate with one another and carry out biological functions together. Analyzing and annotating their metabolic pathways could further explain their corresponding functions. In this study, 691 differential non-volatile metabolites were mapped to the KEGG database and carried out pathway enrichment analysis. In Supplementary Figure S4C, these metabolites were annotated to 65 metabolic pathways that were mainly involved in three primary functions: metabolism, environmental information processing, and genetic information processing. These pathways promoted biosynthesis of secondary metabolites, biosynthesis of secondary metabolites, and tryptophan metabolism (Supplementary Figure S4D). Cluster analysis was performed for all differential metabolites in enriched KEGG

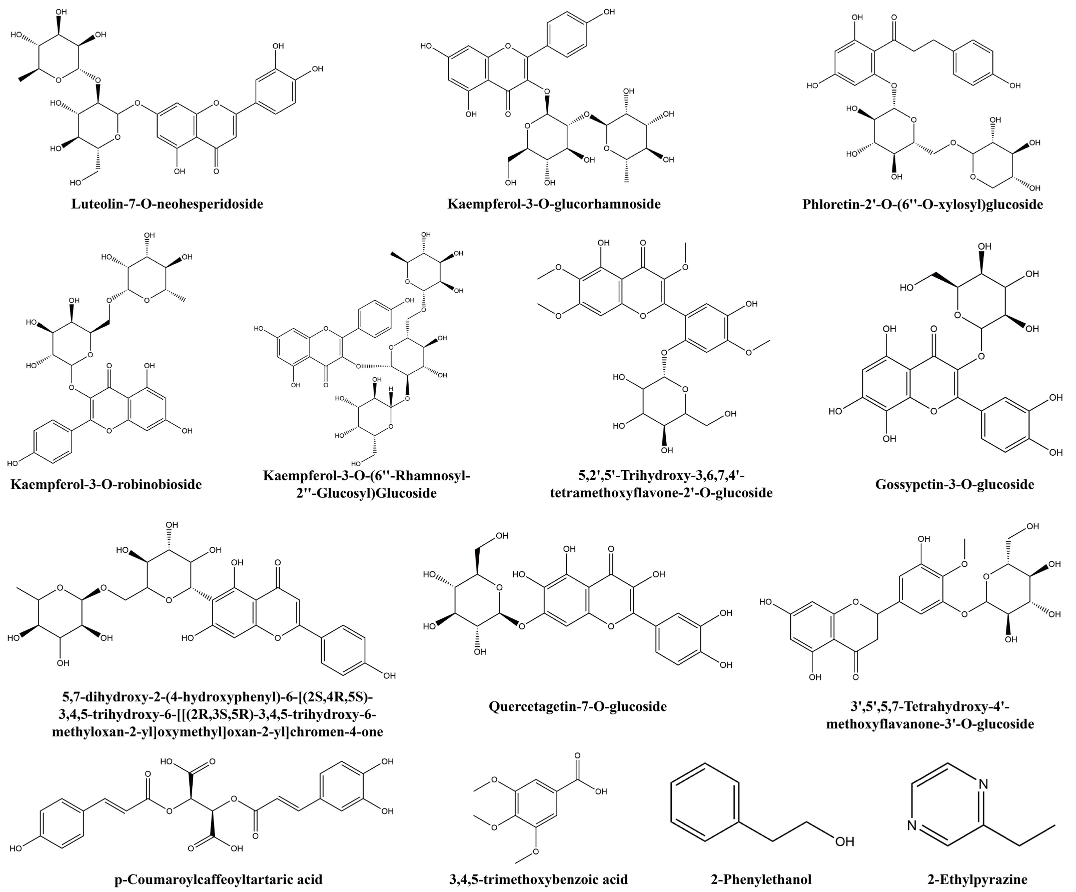


FIGURE 9
The chemical structure of unique non-volatile metabolites of green honeysuckle.

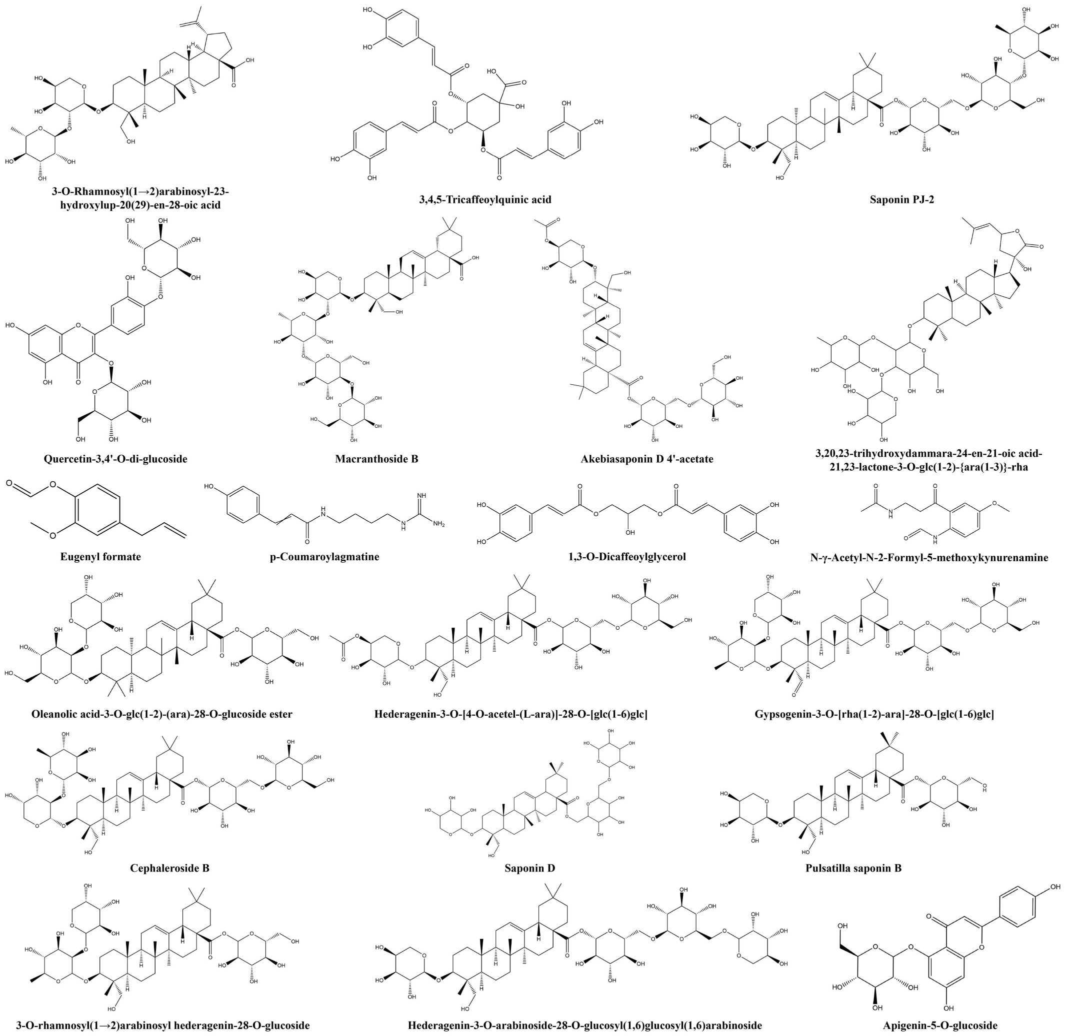


FIGURE 10
The chemical structure of unique non-volatile metabolites of red honeysuckle.

metabolic pathways. In Figure 8G, compared with green honeysuckle, red honeysuckle mainly enhanced tryptophan metabolism, phenylpropanoid biosynthesis, and various plant secondary metabolites' biosynthesis due to the high accumulation pattern of differential metabolites such as tryptamine, caffeic acid, and 2−hydroxycinnamic acid.

3.7 Correlation analysis of differential metabolites

Significant differential metabolic relationships were investigated using Spearman's paired rank correlation analysis, and the molecular networks are shown in Figure 11A (volatile metabolites) and Figure 11B (non−volatile metabolites). Red, blue, and green dots indicated strong, medium, and weak correlations, respectively. The larger the dot, the more related the metabolites are. Positive and negative correlations may represent synthesis and conversion, respectively (Markkinen et al., 2022).

As shown in Figure 11A, most of the correlations were positive. For example, (*E*,*E*)−2,4−heptadien−1−ol was positively correlated with 1−octanol, 2,7−dimethyl−3,6−bis(methylene)−1,7−octadiene, and 2−ethenyl−1,1−dimethyl−3−methylene−cyclohexane, and negatively correlated with β−hellandrene, β−myrcene, Δ−limonene, and trans−β−ocimene. In Figure 11B, carthamone, kaempferol−7−*O*−glucoside, and peonidin−3−*O*−glucoside were the top three metabolites that showed the most correlation with other metabolites. Kaempferol−7−*O*−glucoside, an important flavonoid compound of honeysuckle, has anti−HSV activity and can be considered as a new potential drug candidate for the treatment of HIV infection (Tang et al., 2024). Peonidin−3−*O*−glucoside can not only increase osteoblast differentiation but also serve as a virtual predictor of the anti-inflammatory activity of TNF−α signaling (Ren et al., 2021). These results again showed the potential application value of red honeysuckle.

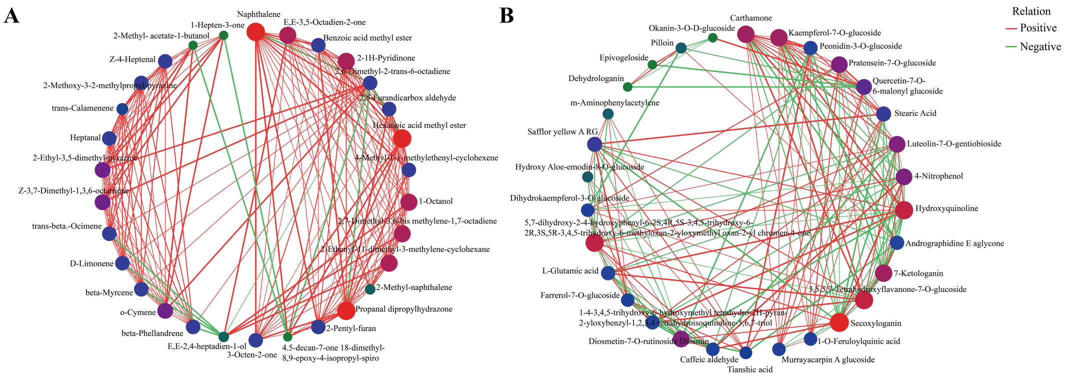


FIGURE 11
Correlation network plots for significant differential metabolites (top 30) of green and red honeysuckles. **(A)** Volatile metabolites. **(B)** Non −volatile metabolites.

4 Discussion

The determination of main biological activities was the main entry point for most scholars to study the targets and mechanisms of honeysuckle in the treatment of diseases (Yang et al., 2014; Lv et al., 2021; Yeh et al., 2022). As more and more plants are found to have antioxidant properties, there is increasing interest in measuring total antioxidant activity rather than simply the contents of different antioxidants (Ramón-Sierra et al., 2022). In this study, red honeysuckle had a higher activity based on three essays, which indicated its potential in cosmetics, pharmaceuticals, and functional foods. Moreover, red honeysuckle also had a higher quercetin content, which further demonstrated its development value.

Similarly, *in vitro* experiments, including cell activity and gene expression, proved that red honeysuckle had higher antibacterial and anti-inflammatory effects. Because its aqueous extracts could inhibit the growth of *S. aureus*, *E. coli*, and *B. subtilis*, producers could use the extract of honeysuckle to prepare AgNP-coated silks for colorful and long-term multifunctional textiles (Zhou and Tang, 2018). In terms of cytotoxicity, our study found that neither of the two water extracts was toxic to RAW264.7 at the same concentration. The traditional green honeysuckle has been recognized as safe, reliable, and edible (Kucharska et al., 2017; Ma et al., 2022). The safety of red honeysuckle, which can be used as a food raw material, is also shown. Furthermore, its aqueous extract could inhibit the expression of IL-6, IL-1β, and TNF-α at both cellular and gene levels. It was found that this inhibition may be due to the *in situ* enzymatic hydrolysis of RAW264.7 cells (Cao et al., 2018). On the whole, red honeysuckle could reduce inflammation by controlling inflammatory factors.

In addition, the bio-active components chlorogenic acid has a series of biological activities at the gastrointestinal level, in addition to decreasing inflammatory cytokines by inhibiting the NF-B pathway (Luo et al., 2021; La Rosa et al., 2023). Cynaroside could reduce body weight, total cholesterol, triglyceride, and LDL-C levels; increase blood HDL-C levels; and reduce liver steatosis (Sun et al., 2021). Caffeic acid could control MAPK and NF-κB activation and significantly reduce serum TNF-α and IL-6 levels when administered during LPS stimulation (Choi et al., 2017). Furthermore, studies have found that quercetin and its derivative quercetin caprylate are proteasome activators that have anti-aging and antioxidant properties (Chondrogianni et al., 2010). Interestingly, compared with green honeysuckle, red honeysuckle had a higher content of these compounds, which were significantly higher than the relevant limit in Chinese Pharmacopoeia. These results provided data support for the application of red honeysuckle in medicines, functional foods, and cosmetics from the perspective of biological activity.

For metabolites, many terpenoids with herbal flavor such as copaene, β−ocimene, and δ−cadinene were identified, and their content was significantly higher in red honeysuckle than in green honeysuckle. Some upregulated metabolites such as 2−heptanone and heptanal and specific metabolites such as (*E*)−4−hexen−1−ol and 1,2,4−trimethyl−benzene also have a herbal flavor. These compounds were the material basis of the rich aroma of red honeysuckle, and were also the main reason why the public prefers it. Moreover, 67% of the specific non-volatile metabolites of red honeysuckle were triterpenoid saponins, which were regarded as potential antitumor agents. For example, dipasperoside B has shown inhibitory activity against NO production in an LPS-stimulated murine macrophage RAW264.7 cell line (Li et al., 2016). Macranthoside B could inhibit cell growth especially in human acute promyelocytic leukemia HL-60 cells (Guan et al., 2011). Akebia saponin D could significantly decrease NO production and iNOS expression, and has anti-nociceptive and anti-inflammatory effects (Gong et al., 2019). These findings provided sufficient theoretical basis for the development of red honeysuckle in health products.

As for the metabolic pathway, some differential metabolites were mainly enriched in tryptophan metabolism pathway by KEGG analysis. Six upregulated differential metabolites related to this pathways were valine−tryptophan, *N*−(1−deoxy−1−fructosyl) −tryptophan, *L*−tryptophan, 5−hydroxy−*L*−tryptophan, glycyl −tryptophan, and *N*−acetyl−*L*−tryptophan. Tryptophan (Trp) is an essential amino acid and plays an important role in physiology and pathophysiology (Seo and Kwon, 2023). The results revealed the extensive development potential of red honeysuckle from another dimension. On the whole, the data obtained in this study could prove that red honeysuckle was better than green honeysuckle, and it was worthy of comprehensive and in-depth development and utilization.

5 Conclusion

In this study, compared with traditional green honeysuckle, the advantages and resource development potential of red honeysuckle were explored. This study provided novel evidence of the quality characterization and material basis of the unique flavor of red honeysuckle. Furthermore, combined with its safe properties, bright color, and effective biological activities, it can be used as a raw material that can be processed into medicines, foods, beverages, biscuits, pigment additives, and even functional foods in the future. In conclusion, our study contributed to reveal the potential applications of red honeysuckle and could help advance such applications.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

ZL: Data curation, Investigation, Methodology, Software, Visualization, Writing – original draft. YC: Formal analysis, Investigation, Software, Validation, Writing – review & editing. YX: Data curation, Investigation, Writing – review & editing. ZC: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The study was funded by the Scientific and Technological Innovation Project of China Academy of Chinese Medical Sciences (No. CI2021A04512, CI2023E001TS) and China Agriculture Research System of the Ministry of Finance and the Ministry of Agriculture and Rural Areas (CARS–21).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1418957/full#supplementary-material

References

Ameeramja, J., Kanagaraj, V. V., and Perumal, E. (2018). Protocatechuic acid methyl ester modulates fluoride induced pulmonary toxicity in rats. *Food. Chem. Toxicol.* 118, 235–244. doi: 10.1016/j.fct.2018.05.031

Ameeramja, J., and Perumal, E. (2017). Protocatechuic acid methyl ester ameliorates fluoride toxicity in A549 cells. *Food. Chem. Toxicol.* 109, 941–950. doi: 10.1016/j.fct.2016.12.024

Banožić, M., Jokić, S., Ačkar, Đ., Blažić, M., and Šubarić, D. (2020). Carbohydrates-key players in tobacco aroma formation and quality determination. *Molecules* 25, 1734. doi: 10.3390/molecules25071734

Belyaeva, O. V., Sergeeva, I. Y., Belyaeva, E. E., and Chernobrovkina, E. V. (2021). Study of antioxidant activity of juices and beverages from blue honeysuckle and black chokeberry. *Iop Conf. Series. Earth Environ. Sci.* 640, 52008. doi: 10.1088/1755-1315/640/5/052008

Cao, Y., Zhou, X., Yin, Z., Yu, X., Yang, Q., Guo, Q., et al. (2018). The anti-inflammatory effect of BML-111 on COPD may be mediated by regulating NLRP3 inflammasome activation and ROS production. *Prostaglandins Other Lipid Mediat.* 138, 23–30. doi: 10.1016/j.prostaglandins.2018.08.001

Carrillo-Galván, G., Bye, R., Eguiarte, L. E., Cristians, S., Pérez-López, P., Vergara-Silva, F., et al. (2020). Domestication of aromatic medicinal plants in Mexico: Agastache (Lamiaceae)-an ethnobotanical, morpho-physiological, and phytochemical analysis. *J. Ethnobiol. Ethnomed.* 16, 22. doi: 10.1186/s13002-020-00368-2

Choi, K. C., Son, Y. O., Hwang, J. M., Kim, B. T., Chae, M., and Lee, J. (2017). Antioxidant, anti-inflammatory and anti-septic potential of phenolic acids and flavonoid fractions isolated from *Lolium* multiflorum. *Pharm. Biol.* 55, 611–619. doi: 10.1080/13880209.2016.1266673

Chondrogianni, N., Kapeta, S., Chinou, I., Vassilatou, K., Papassideri, I., and Gonos, E. S. (2010). Anti-ageing and rejuvenating effects of quercetin. *Exp. Gerontol.* 45, 763–771. doi: 10.1016/j.exger.2010.07.001

Clark, C. C. T., Barnes, C. M., Duncan, M. J., Summers, H. D., and Stratton, G. (2019). Physical activity, motor competence and movement and gait quality: A principal component analysis. *Hum. Mov. Sci.* 68, 102523. doi: 10.1016/j.humov.2019.102523

Eltaib, L., and Alzain, A. A. (2023). Targeting the omicron variant of SARS-CoV-2 with phytochemicals from Saudi medicinal plants: molecular docking combined with molecular dynamics investigations. *J. Biomol. Struct. Dyn* 41, 9732–9744. doi: 10.1080/07391102.2022.2146203

Fang, L., Long, N., Li, Y., Liao, X., Shi, L., Zhao, H., et al. (2022). Transfer behavior of pesticides from honeysuckle into tea infusions: Establishment of an empirical model for transfer rate prediction. *Ecotoxicol. Environ. Saf.* 234, 113377. doi: 10.1016/j.ecoenv.2022.113377

Feng, G., Li, D., Liu, J., Sun, S., Zhang, P., Liu, W., et al. (2022). The herbal combination of *Radix astragali*, *Radix angelicae sinensis*, and *Caulis lonicerae* regulates the functions of type 2 innate lymphocytes and macrophages contributing to the resolution of collagen-induced arthritis. *Front. Pharmacol.* 13. doi: 10.3389/fphar.2022.964559

Gong, L. L., Yang, S., Liu, H., Zhang, W., Ren, L. L., Han, F.F., et al. (2019). Antinociceptive and anti-inflammatory potentials of Akebia saponin D. *Eur. J. Pharmacol.* 845, 85–90. doi: 10.1016/j.ejphar.2018.11.038

Guan, F., Shan, Y., Zhao, X., Zhang, D., Wang, M., Peng, F., et al. (2011). Apoptosis and membrane permeabilisation induced by macranthoside B on HL-60 cells. *Nat. Prod. Res.* 25, 332–340. doi: 10.1080/14786411003752086

Guo, X., Yu, X., Zheng, B., Zhang, L., Zhang, F., Zhang, Y., et al. (2021). Network pharmacology-based identification of potential targets of *Lonicerae japonicae* flos acting on anti-inflammatory effects. *BioMed. Res. Int.* 2021, 5507003. doi: 10.1155/2021/5507003

Hu, Y., Li, T., Gou, Q., and Zhang, X. (2017). Determination of five active substances in different organs of *Lonicera japonica* var. *chinensis* (wats.) Bak. and *Lonicera japonica* Thunb. by HPLC method. *J. Anhui Agric. Sci.* 45, , 126–127,147. doi: 10.3969/j.issn.0517-6611.2017.34.040

Khoo, H. E., Azlan, A., Tang, S. T., and Lim, S. M. (2017). Anthocyanidins and anthocyanins: Colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutr. Res.* 61, 1361779. doi: 10.1080/16546628.2017.1361779

Kirina, I. B., Titova, L. V., Popova, E. I., Grigoreva, L. V., and Khoroshkova, Y. V. (2021). Biochemical value of berries of promising edible honeysuckle varieties for the production of functional food products. *Iop Conf. Series. Earth Environ. Sci.* 845, 12097. doi: 10.1088/1755-1315/845/1/012097

Kucharska, A. Z., Sokol-Letowska, A., Oszmianski, J., Piorecki, N., and Fecka, I. (2017). Iridoids, phenolic compounds and antioxidant activity of edible honeysuckle berries (*Lonicera caerulea* var. *Kamtschatica* sevast.). *Molecules* 22, 405. doi: 10.3390/molecules22030405

La Rosa, G., Sozio, C., Pipicelli, L., Raia, M., Palmiero, A., Santillo, M., et al. (2023). Antioxidant, anti-inflammatory and pro-differentiative effects of chlorogenic acid on M03-13 human oligodendrocyte-like cells. *Int. J. Mol. Sci.* 24, 16731. doi: 10.3390/ijms242316731

Li, F., Tanaka, K., Watanabe, S., and Tezuka, Y. (2016). Dipasperoside B, a new trisiridoid glucoside from dipsacus asper. *Nat. Prod. Commun.* 11, 891–894. doi: 10.1177/1934578X1601100706

Li, G., Liu, J., Zhang, H., Jia, L., Liu, Y., Li, J., et al. (2023). Volatile metabolome and floral transcriptome analyses reveal the volatile components of strongly fragrant progeny of *Malus* × *robusta*. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1065219

Li, J., Jia, G., Wang, J., Li, J., and Yang, L. (2013). Weight and index ingredients components content comparison of different flowering phase in *Lonicera japonica.* var. *Chinensis* (wats.) Bak. *J. Henan Agric. Univ.* 47, 534–537. doi: 10.3969/j.issn.1000-2340.2013.05.005

Li, J., Ye, C., Lian, X., Wang, L., and Niu, Y. (2019). Study on breeding system and pollination biology of *Lonicera japonica* var. *Chinensis*. *J. Henan Agric. Univ.* 53, 581–590. doi: 10.16445/j.cnki.1000-2340.2019.04.013

Liu, Z., Cheng, Y., and Chao, Z. (2023). A comprehensive quality analysis of different colors of medicinal and edible honeysuckle. *Foods* 12, 3126. doi: 10.3390/foods12163126

Liu, Z., Fang, Y., Wu, C., Hai, X., Xu, B., Li, Z., et al. (2022). The difference of volatile compounds in female and male buds of herpetospermum pedunculosum based on HS-SPME-GC-MS and multivariate statistical analysis. *Molecules* 27, 1288. doi: 10.3390/molecules27041288

Luo, J., He, W., Li, X., Ji, X., and Liu, J. (2021). Anti-acne vulgaris effects of chlorogenic acid by anti-inflammatory activity and lipogenesis inhibition. *Exp. Dermatol.* 30, 865–871. doi: 10.1111/exd.14277

Lv, Q., Xing, Y., Liu, J., Dong, D., Liu, Y., Qiao, H., et al. (2021). Lonicerin targets EZH2 to alleviate ulcerative colitis by autophagy-mediated NLRP3 inflammasome inactivation. *Acta Pharm. Sin. B.* 11, 2880–2899. doi: 10.1016/j.apsb.2021.03.011

Ma, A., Zou, F., Zhang, R., and Zhao, X. (2022). The effects and underlying mechanisms of medicine and food homologous flowers on the prevention and treatment of related diseases. *J. Food Biochem.* 46, e14430. doi: 10.1111/jfbc.14430

Markkinen, N., Pariyani, R., Jokioja, J., Kortesniemi, M., Laaksonen, O., Yang, B., et al. (2022). NMR-based metabolomics approach on optimization of malolactic fermentation of sea buckthorn juice with lactiplantibacillus plantarum. *Food Chem.* 366, 130630. doi: 10.1016/j.foodchem.2021.130630

Mu, W., Hu, N., Zhang, L., Jiang, W., Yan, T., Zhang, T., et al. (2022). *Lonicerae japonicae* flos ameliorates radiotherapy-induced mesenteric artery endothelial dysfunction through GTPCH$_1$/BH$_4$/eNOS pathway. *Phytomedicine* 102, 154146. doi: 10.1016/j.phymed.2022.154146

Ramón-Sierra, J. M., Villanueva, M. A., Yam-Puc, A., Rodríguez-Mendiola, M., Arias-Castro, C., Ortiz-Vázquez, E., et al. (2022). Antimicrobial and antioxidant activity of proteins isolated from *Melipona beecheii* honey. *Food Chemistry: X* 13, 100177. doi: 10.1016/j.fochx.2021.100177

Ren, Z., Raut, N. A., Lawal, T. O., Patel, S. R., Lee, S. M., and Mahady, G. B. (2021). Peonidin-3-O-glucoside and cyanidin increase osteoblast differentiation and reducerankl-induced bone resorption in transgenic medaka. *Phytother. Res.* 35, 6255–6269. doi: 10.1002/ptr.7271

Seo, S., and Kwon, B. (2023). Immune regulation through tryptophan metabolism. *Exp. Mol. Med.* 55, 1371-1379. doi: 10.1038/s12276-023-01028-7

Siebert, D. A., Paganelli, C. J., Queiroz, G. S., and Alberton, M. D. (2021). Anti-inflammatory activity of the epicuticular wax and its isolated compounds catechin and gallocatechin from *Eugenia brasiliensis* Lam. (Myrtaceae) leaves. *Nat. Prod. Res.* 35, 4720–4723. doi: 10.1080/14786419.2019.1710707

Su, X., Zhu, Z., Zhang, L., Wang, Q., Xu, M., Lu, C., et al. (2021). Anti-inflammatory property and functional substances of *Lonicerae japonicae* caulis. *J. Ethnopharmacol.* 267, 113502. doi: 10.1016/j.jep.2020.113502

Sun, J., Wang, Z., Chen, L., and Sun, G. (2021). Hypolipidemic effects and preliminary mechanism of chrysanthemum flavonoids, its main components luteolin and luteoloside in hyperlipidemia rats. *Antioxidants* 10, 1309. doi: 10.3390/antiox10081309

Tan, Z., Lu, D., Li, L., Yu, Y., Xu, L., Dong, W., et al. (2022). Loning and expression analysis of the anthocyanin synthase encoding gene RLJANS1 and its relationship with anthocyanin accumulation in honeysuckle. *Jiangsu Agric. Sci.* 50, 69–75. doi: 10.15889/j.issn.1002-1302.2022.21.010

Tang, R., Lin, L., Liu, Y., and Li, H. (2024). Bibliometric and visual analysis of global publications on kaempferol. *Front. Nutr.* 11, 1442574. doi: 10.3389/fnut.2024.1442574

Wang, H., Hua, J., Yu, Q., Li, J., Wang, J., Dong, W., et al. (2021). Widely targeted metabolomic analysis reveals dynamic changes in non-volatile and volatile metabolites during green tea processing. *Food Chem.* 363, 130131. doi: 10.1016/j.foodchem.2021.130131

Wang, M., Huang, H., Wang, L., Yang, H., He, S., Liu, F., et al. (2021). Herbal extract mixture modulates intestinal antioxidative capacity and microbiota in weaning piglets. *Front. Microbiol.* 12. doi: 10.3389/fmicb.2021.706758

Xiao, Y., Ren, Q., and Wu, L. (2022). The pharmacokinetic property and pharmacological activity of acteoside: A review. *Biomed. Pharmacother.* 153, 113296. doi: 10.1016/j.biopha.2022.113296

Yang, W. J., Liu, C., Gu, Z. Y., Zhang, X. Y., Cheng, B., Mao, Y., et al. (2014). Protective effects of acacetin isolated from Ziziphora clinopodioides Lam. (Xintahua) on neonatal rat cardiomyocytes. *Chin. Med.* 9, 28. doi: 10.1186/s13020-014-0028-3

Yeh, Y., Doan, L. H., Huang, Z., Chu, L., Shi, T., Lee, Y., et al. (2022). Honeysuckle (*Lonicera japonica*) and huangqi (*Astragalus membranaceus*) suppress SARS-CoV-2 entry and COVID-19 related cytokine storm in *vitro*. *Front. Pharmacol.* 12. doi: 10.3389/fphar.2021.765553

Yu, X. (2013). Preliminary research of anthocyanidins in lonicera japonica *var.* Chinensis and its formation mechanism (Wuhan: Wuhan Polytechnic University).

Yuan, Y., Yang, J., Yu, X., Huang, L., and Lin, S. (2014). Anthocyanins from buds of *Lonicera japonica* Thunb. Var. *Chinensis* (wats.) Bak. *Food Res. Int.* 62, 812–818. doi: 10.1016/j.foodres.2014.03.026

Zhao, X., Li, H., Wang, J., Guo, Y., Liu, B., Deng, X., et al. (2016). Verbascoside alleviates pneumococcal pneumonia by reducing pneumolysin oligomers. *Mol. Pharmacol.* 89, 376–387. doi: 10.1124/mol.115.100610

Zhao, Y., Jiang, C., Lu, J., Sun, Y., and Cui, Y. (2023). Research progress of proanthocyanidins and anthocyanidins. *Phytother. Res.* 37, 2552–2577. doi: 10.1002/ptr.7850

Zhou, Y., and Tang, R. C. (2018). Facile and eco-friendly fabrication of AgNPs coated silk for antibacterial and antioxidant textiles using honeysuckle extract. *J. Photochem. Photobiol. B.* 178, 463–471. doi: 10.1016/j.jphotobiol.2017.12.003

Zou, S., Wu, J., Shahid, M. Q., He, Y., Lin, S., Liu, Z., et al. (2020). Identification of key taste components in loquat using widely targeted metabolomics. *Food Chem.* 323, 126822. doi: 10.1016/j.foodchem.2020.126822

Frontiers in Plant Science

TYPE Opinion
PUBLISHED 26 November 2024
DOI 10.3389/fpls.2024.1506569

OPEN ACCESS

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Mihir Halder,
Barasat Government College, India

*CORRESPONDENCE
Mengquan Yang
mengquan_yang@outlook.com

RECEIVED 05 October 2024
ACCEPTED 13 November 2024
PUBLISHED 26 November 2024

CITATION
Yin Q and Yang M (2024) Chemoproteomics approach to elucidating biosynthetic pathway of plant natural products.
Front. Plant Sci. 15:1506569.
doi: 10.3389/fpls.2024.1506569

Chemoproteomics approach to elucidating biosynthetic pathway of plant natural products

Quanyu Yin and Mengquan Yang*

National Tobacco Cultivation, Physiology and Biochemistry Research Center, Flavors and Fragrance Engineering and Technology Research Center of Henan Province, College of Tobacco Science, Henan Agricultural University, Zhengzhou, Henan, China

KEYWORDS

chemoproteomics, biosynthesis, natural product, affinity probe, medicinal plant

1 Introduction

Natural products derived from medicinal plants are a class of compounds with extensive biological activities, playing a crucial role in the pharmaceutical, food, and cosmetics industries. Due to their excellent physiological functions, increasing attention is being paid to the biosynthesis pathways of plant natural products (PNPs) (Kawatra et al., 2022; Halder and Jha, 2023). However, as market demand continues to grow, traditional harvesting and plant extraction methods exert immense pressure on the environment (Singh, 2023). In recent years, the rapid advancement of synthetic biology has offered new approaches for producing structurally complex bioactive small-molecule compounds using biotechnology (Hesami et al., 2023). Nevertheless, the lack of knowledge about biosynthetic pathways significantly impedes the large-scale biomanufacturing of natural products from medicinal plants. Unlike microorganisms, the biosynthetic genes for plant natural products are relatively dispersed across chromosomes, and medicinal plants often lack efficient genetic manipulation systems, which hinders the elucidation of their biosynthetic pathways. Recently, chemoproteomics based on activity probes has demonstrated great potential in elucidating plant natural product biosynthesis (steviol glycosides, camptothecin, chalcomoracin, etc.), as it enables the rapid identification of functional proteins interacting with substrates, thereby accelerating the discovery of biosynthetic pathways (Li et al., 2018; Zhou et al., 2018; Gao et al., 2020; Wong et al., 2020; Zhang et al., 2024) (Figure 1).

Traditional approaches have played a crucial role in advancing our understanding of plant natural product biosynthetic pathways, laying a foundation for emerging technologies like chemoproteomics. Gene knockout and RNA interference (RNAi) methods, for example, have been widely used to identify genes involved in biosynthetic pathways by observing phenotypic changes in metabolite production when specific genes are silenced (Zhao et al., 2016). Additionally, multi-omics approaches, such as transcriptomics, offer insights by gene coexpression, though these methods can be limited by the need for extensive data analysis and do not directly identify enzyme activities (Liu et al., 2024a; Swamidatta and Lichman,

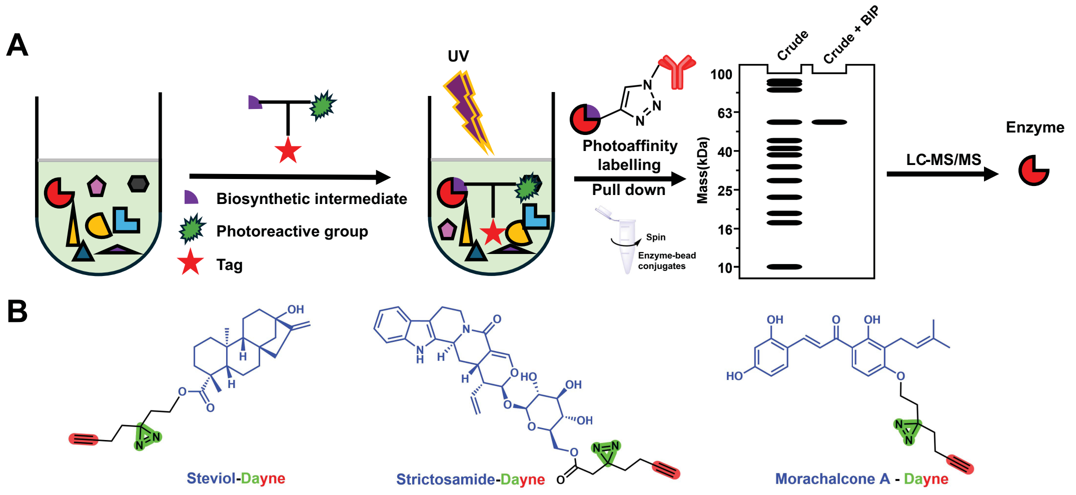


FIGURE 1
(A) Chemoproteomics approach to characterize the enzymes involved in the biosynthesis of plant natural products. **(B)** Reported photoaffinity-based probes for enzyme discovery.

2024). Heterologous gene expression, often in microbial and plant systems, has enabled functional analysis of individual genes or gene clusters by recreating biosynthetic pathways outside the native plant context (Lau and Sattely, 2015; Hong et al., 2022; Yang et al., 2024). However, traditional biochemical assays used to verify enzyme function may require large amounts of purified protein, a time-intensive process (Tatsis et al., 2017). These methods, while foundational, often fall short in dissecting complex pathways directly within plants, which is where chemoproteomics, with its activity-based probes and functional annotation capabilities, offers distinct advantages. Consequently, applying chemoproteomics technology to comprehensively analyze the biosynthesis of plant natural products not only has practical value for the rapid identification of functional genes involved in biosynthesis but also holds strategic significance for achieving large-scale production of medicinal plant natural products through synthetic biology (Zhang et al., 2022; Gao et al., 2023; Liu et al., 2023; Zhang et al., 2023; Golubova et al., 2024; Jiang et al., 2024; Liu et al., 2024b).

2 Workflow of affinity probes

Affinity probes are specialized chemical tools used in chemoproteomics to isolate and identify active enzymes within complex biological samples, particularly those involved in plant natural product biosynthesis. These probes typically consist of a binding moiety that targets the enzyme's active site, a reactive tag that enables enzyme capture through covalent attachment after activation and a reporter tag for detection. Effective affinity probe design requires specificity to mimic natural substrates, stability in biological conditions, and controlled reactivity to ensure selective and durable binding (Parker and Pratt, 2020; Fang et al., 2021) (Figure 2). The primary advantage of affinity probes lies in their ability to selectively target and capture active enzymes within native proteomes, bypassing the need for extensive purification or genetic manipulation. However, challenges such as non-specific binding and the complexity of probe design can limit their effectiveness (Tabana et al., 2023). Affinity probes have proven invaluable for mapping biosynthetic pathways in plants, as shown in studies on enzymes synthesizing steviol glycosides and other complex natural products.

3 UDP-glycosyltransferases in steviol glycosides biosynthesis

Steviol glycosides, the sweetening agents derived from *Stevia rebaudiana*, have been widely studied due to their potential as non-caloric sweeteners (Masand et al., 2024). Recent research utilizing a chemoproteomics-based strategy successfully identified the UDP-glycosyltransferases (UGTs) *Sr*UGT73E1, *At*UGT73C1 and *At*UGT73C5, which play a pivotal role in catalyzing the glycosylation of steviol to form steviol glycosides (Li et al., 2018; Zhou et al., 2018; Wong et al., 2020). The use of a photoaffinity probe specific to steviol, combined with mass spectrometry, allowed researchers to selectively profile the UGTs responsible for the final glycosylation steps. This discovery not only advances our understanding of the biosynthetic pathway of steviol glycosides but also offers a novel platform for the rapid identification of other enzymes involved in glycosylation, enabling synthetic biology approaches for scalable production.

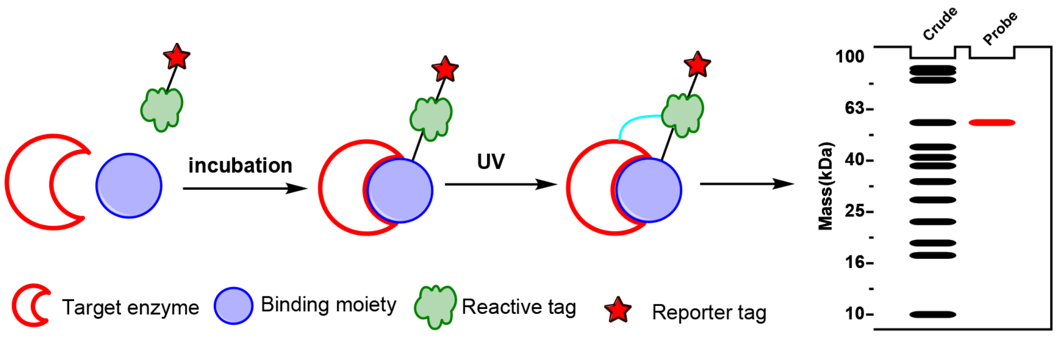


FIGURE 2
The workflow of affinity probes.

4 Chalcomoracin biosynthesis through FAD-dependent cycloaddition

Chalcomoracin, a bioactive flavonoid isolated from mulberry (*Morus alba*), is synthesized through a highly unique flavin adenine dinucleotide (FAD)-dependent intermolecular Diels-Alder reaction. For years, the enzyme responsible for this cycloaddition reaction was unknown, despite its importance in the formation of chalcomoracin's characteristic cyclohexene ring. Recent studies have identified a novel enzyme, *Morus alba* Diels–Alderase (*Ma*DA), through a biosynthetic intermediate probe (BIP)-based chemoproteomics strategy (Gao et al., 2020). *Ma*DA catalyzes the [4 + 2] cycloaddition with high specificity and enantioselectivity, marking the first discovery of a stand-alone intermolecular Diels-Alderase in plants (Gao et al., 2020, 2024). The use of chemoproteomics in this context allowed the functional characterization of this enzyme, which had remained inaccessible through traditional genomics or transcriptomics due to the lack of gene clustering in plant biosynthetic pathways.

5 Camptothecin biosynthesis and the role of OpCYP716E111

Camptothecin, an alkaloid with potent anti-cancer properties, is derived from *Camptotheca acuminata* and *Ophiorrhiza pumila* (Yang et al., 2021). The biosynthesis of camptothecin has long been a subject of study, with a significant gap in understanding the steps following strictosamide formation. A breakthrough came with the discovery of *Op*CYP716E111, an epoxidase responsible for catalyzing the conversion of strictosamide to strictosamide epoxide. Using a chemoproteomic approach, researchers designed a diazirine-based probe specific to strictosamide, which enabled the selective identification of *Op*CYP716E111 in the proteome of *Ophiorrhiza pumila* (Zhang et al., 2024). This discovery fills a critical gap in the camptothecin biosynthesis pathway and underscores the power of chemoproteomics to uncover previously unknown enzymes involved in complex plant metabolic processes.

6 The broader impact of chemoproteomics in plant biosynthesis

The discoveries surrounding steviol glycosides, chalcomoracin, and camptothecin highlight the broad applicability of chemoproteomics in the field of plant natural product biosynthesis (Li et al., 2018; Zhou et al., 2018; Gao et al., 2020; Wong et al., 2020; Zhang et al., 2024). Traditional methods such as transcriptomics and gene knockout studies often fall short in plants due to the dispersed nature of biosynthetic genes, making it difficult to pinpoint the enzymes responsible for each step. Chemoproteomics circumvents this issue by directly targeting enzyme activity through small molecule probes, allowing for rapid functional annotation of enzymes even in non-model plants. This approach is particularly advantageous in plants where secondary metabolism genes are not organized into clusters, a feature common in microbial systems but rare in plants. Chemoproteomics offers unique advantages for studying biosynthetic pathways, such as high sensitivity for detecting low-abundance enzymes without needing gene cloning or protein expression steps. It can also distinguish between closely related isoforms and profile multiple enzymes simultaneously, enabling a comprehensive view of metabolic networks in plant systems. These features make chemoproteomics particularly valuable for advancing research in natural product biosynthesis. Furthermore, the integration of chemoproteomics with synthetic biology holds the promise of sustainable production of these valuable compounds. By identifying and characterizing the enzymes involved in natural product biosynthesis, researchers can reconstitute these pathways in microbial hosts, enabling the scalable and controlled production of complex plant-derived compounds (Zhang et al., 2022; Gao et al., 2023).

Despite its advantages, chemoproteomics faces challenges in studying complex plant biosynthetic pathways. These include non-specific binding of probes, the need for extensive optimization in non-model plants, and reliance on high-quality mass spectrometry data for accurate enzyme identification. Additionally, challenges arise from probe design limitations that may affect binding efficiency and specificity. Future advancements, such as improved probe selectivity and integration with other omics technologies, hold promise for overcoming these obstacles and enhancing the approach's utility in biosynthetic research.

7 Conclusion

Chemoproteomics has proven to be an indispensable tool in elucidating the biosynthetic pathways of complex natural products like steviol glycosides, chalcomoracin, and camptothecin. The ability to directly profile active enzymes involved in these pathways offers a new frontier in plant natural product research, accelerating the discovery of key biosynthetic genes and facilitating their application in synthetic biology. As this field continues to grow, chemoproteomics will likely play a central role in unlocking the full potential of plant-derived natural products for pharmaceutical and industrial applications.

Author contributions

QY: Conceptualization, Writing – review & editing. MY: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was financially supported by the National Natural Science Foundation of China (32400218), the Key Technology R&D Program of Henan Province (242102110240, 232102110053), the Special Support Fund for High-level Talents and skills improvement of Henan Agricultural University (30501474).

Conflict of interest

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

Generative AI statement

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

Publisher's note

References

Fang, H., Peng, B., Ong, S. Y., Wu, Q., Li, L., and Yao, S. Q. (2021). Recent advances in activity-based probes (ABPs) and affinity-based probes (AfBPs) for profiling of enzymes. *Chem. Sci.* 12, 8288–8310. doi: 10.1039/D1SC01359A

Gao, L., Ding, Q., and Lei, X. (2024). Hunting for the intermolecular diels-alderase. *Acc. Chem. Res.* 57, 2166–2183. doi: 10.1021/acs.accounts.4c00315

Gao, L., Su, C., Du, X., Wang, R., Chen, S., Zhou, Y., et al. (2020). FAD-dependent enzyme-catalysed intermolecular [4 + 2] cycloaddition in natural product biosynthesis. *Nat. Chem.* 12, 620–628. doi: 10.1038/s41557-020-0467-7

Gao, J., Zuo, Y., Xiao, F., Wang, Y., Li, D., Xu, J., et al. (2023). Biosynthesis of catharanthine in engineered Pichia pastoris. *Nat. Synthesis*. 2, 231–242. doi: 10.1038/s44160-022-00205-2

Golubova, D., Tansley, C., Su, H., and Patron, N. J. (2024). Engineering Nicotiana benthamiana as a platform for natural product biosynthesis. *Curr. Opin. Plant Biol.* 81, 102611. doi: 10.1016/j.pbi.2024.102611

Halder, M., and Jha, S. (2023). "Medicinal plants and bioactive phytochemical diversity: A fountainhead of potential drugs against human diseases," in *Medicinal Plants: Biodiversity, Biotechnology and Conservation*. Eds. S. Jha and M. Halder (Springer Nature Singapore, Singapore), 39–93.

Hesami, M., Pepe, M., Baiton, A., and Jones, A. M. P. (2023). Current status and future prospects in cannabinoid production through in *vitro* culture and synthetic biology. *Biotechnol. Adv.* 62, 108074. doi: 10.1016/j.biotechadv.2022.108074

Hong, B., Grzech, D., Caputi, L., Sonawane, P., Lopez, C. E. R., Kamileen, M. O., et al. (2022). Biosynthesis of strychnine. *Nature* 607, 617–622. doi: 10.1038/s41586-022-04950-4

Jiang, B., Gao, L., Wang, H., Sun, Y., Zhang, X., Ke, H., et al. (2024). Characterization and heterologous reconstitution of Taxus biosynthetic enzymes leading to baccatin III. *Science* 383, 622–629. doi: 10.1126/science.adj3484

Kawatra, A., Gupta, S., Dhankhar, R., Singh, P., and Gulati, P. (2022). "Application of phytochemicals in therapeutic, food, flavor, and cosmetic industries," in *Phytochemical Genomics: Plant Metabolomics and Medicinal Plant Genomics*. Eds. M. K. Swamy and A. Kumar (Springer Nature Singapore, Singapore), 85–108.

Lau, W., and Sattely, E. S. (2015). Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science* 349, 1224–1228. doi: 10.1126/science.aac7202

Li, W., Zhou, Y., You, W., Yang, M., Ma, Y., Wang, M., et al. (2018). Development of photoaffinity probe for the discovery of steviol glycosides biosynthesis pathway in stevia rebuadiana and rapid substrate screening. *ACS Chem. Biol.* 13, 1944–1949. doi: 10.1021/acschembio.8b00285

Liu, F., Fernie, A. R., and Zhang, Y. (2024a). Plant gene co-expression defines the biosynthetic pathway of neuroactive alkaloids. *Mol. Plant*. 17, 372–374. doi: 10.1016/j.molp.2024.02.002

Liu, J. C., de la Pena, R., Tocol, C., and Sattely, E. S. (2024b). Reconstitution of early paclitaxel biosynthetic network. *Nat. Commun.* 15, 1419. doi: 10.1038/s41467-024-45574-8

Liu, X., Zhang, P., Zhao, Q., and Huang, A. C. (2023). Making small molecules in plants: A chassis for synthetic biology-based production of plant natural products. *J. Integr. Plant Biol.* 65, 417–443. doi: 10.1111/jipb.13330

Masand, M., Sharma, S., Kumari, S., Pal, P., Majeed, A., Singh, G., et al. (2024). High-quality haplotype-resolved chromosome assembly provides evolutionary insights and targeted steviol glycosides (SGs) biosynthesis in Stevia rebaudiana Bertoni. *Plant Biotechnol. J.* doi: 10.1111/pbi.14446

Parker, C. G., and Pratt, M. R. (2020). Click chemistry in proteomic investigations. *Cell* 180, 605–632. doi: 10.1016/j.cell.2020.01.025

Singh, H. (2023). "Enhancement of plant secondary metabolites by genetic manipulation," in *Genetic Manipulation of Secondary Metabolites in Medicinal Plant*. Eds. R. Singh and N. Kumar (Springer Nature Singapore, Singapore), 59–90.

Swamidatta, S. H., and Lichman, B. R. (2024). Beyond co-expression: pathway discovery for plant pharmaceuticals. *Curr. Opin. Biotechnol.* 88, 103147. doi: 10.1016/j.copbio.2024.103147

Tabana, Y., Babu, D., Fahlman, R., Siraki, A. G., and Barakat, K. (2023). Target identification of small molecules: an overview of the current applications in drug discovery. *BMC Biotechnol.* 23, 44. doi: 10.1186/s12896-023-00815-4

Tatsis, E. C., Carqueijeiro, I., Dugé De Bernonville, T., Franke, J., Dang, T. T. T., Oudin, A., et al. (2017). A three enzyme system to generate the Strychnos alkaloid scaffold from a central biosynthetic intermediate. *Nat. Commun.* 8, 316. doi: 10.1038/s41467-017-00154-x

Wong, N. K., Zhong, S., Li, W., Zhou, F., Deng, Z., and Zhou, Y. (2020). Selective profiling of steviol-catalyzing UDP-glycosyltransferases with a metabolically synthesized probe. *Chem. Commun. (Camb)* 56, 12387–12390. doi: 10.1039/D0CC04948D

Yang, M., Wang, Q., Liu, Y., Hao, X., Wang, C., Liang, Y., et al. (2021). Divergent camptothecin biosynthetic pathway in Ophiorrhiza pumila. *BMC Biol.* 19, 122. doi: 10.1186/s12915-021-01051-y

Yang, C., Wang, Y., Su, Z., Xiong, L., Wang, P., Lei, W., et al. (2024). Biosynthesis of the highly oxygenated tetracyclic core skeleton of Taxol. *Nat. Commun.* 15, 2339. doi: 10.1038/s41467-024-46583-3

Zhang, J., Hansen, L. G., Gudich, O., Viehrig, K., Lassen, L. M. M., Schrubbers, L., et al. (2022). A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature.* 609, 341–347. doi: 10.1038/s41586-022-05157-3

Zhang, T., Wang, Y., Wu, S., Tian, E., Yang, C., Zhou, Z., et al. (2024). Chemoproteomics reveals the epoxidase enzyme for the biosynthesis of camptothecin in Ophiorrhiza pumila. *J. Integr. Plant Biol.* 66, 1044–1047. doi: 10.1111/jipb.13594

Zhang, Y., Wiese, L., Fang, H., Alseekh, S., Perez De Souza, L., Scossa, F., et al. (2023). Synthetic biology identifies the minimal gene set required for paclitaxel biosynthesis in a plant chassis. *Mol. Plant* 16, 1951–1961. doi: 10.1016/j.molp.2023.10.016

Zhao, Q., Zhang, Y., Wang, G., Hill, L., Weng, J.-K., Chen, X.-Y., et al. (2016). A specialized flavone biosynthetic pathway has evolved in the medicinal plant, Scutellaria baicalensis. *Sci. Adv.* 2, e1501780. doi: 10.1126/sciadv.1501780

Zhou, Y., Li, W., You, W., Di, Z., Wang, M., Zhou, H., et al. (2018). Discovery of: Arabidopsis UGT73C1 as a steviol-catalyzing UDP-glycosyltransferase with chemical probes. *Chem. Commun.* 54, 7179–7182. doi: 10.1039/C7CC09951G

Frontiers in Plant Science

TYPE Original Research
PUBLISHED 07 January 2025
DOI 10.3389/fpls.2024.1470307

OPEN ACCESS

EDITED BY
Xin Fang,
Chinese Academy of Sciences (CAS), China

REVIEWED BY
Alfredo J. Ibáñez,
Pontificia Universidad Católica del Perú, Peru
Hui Liu,
Chinese Academy of Sciences (CAS), China
Gang Li,
Qingdao University, China

*CORRESPONDENCE
Alain Muselli
✉ muselli_a@univ-corse.fr

RECEIVED 25 July 2024
ACCEPTED 26 November 2024
PUBLISHED 07 January 2025

CITATION
Pannequin A, Muselli A, Marcourt L, Ferreira Queiroz E, Quiros-Guerrero L-M, Asakawa Y, Dounoue-Kubo M and Wolfender J-L (2025) Comprehensive comparative metabolome study of a large collection of Corsican bryophytes. *Front. Plant Sci.* 15:1470307. doi: 10.3389/fpls.2024.1470307

Comprehensive comparative metabolome study of a large collection of Corsican bryophytes

Anaïs Pannequin[1,2], Alain Muselli[1]*, Laurence Marcourt[2,3], Emerson Ferreira Queiroz[2,3], Luis-Manuel Quiros-Guerrero[2,3], Yoshinori Asakawa[4], Miwa Dounoue-Kubo[4] and Jean-Luc Wolfender[2,3]

[1]Université de Corse, Unité Mixte de Recherche du Centre national de la recherche scientifique (UMR CNRS) SPE 6134, Laboratoire Chimie des Produits Naturels, Corte, France, [2]Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, Geneva, Switzerland, [3]School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland, [4]Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Tokushima, Japan

Introduction: Bryophytes are non-vascular plants that appeared on Earth before vascular plants. More than 24,000 species are reported worldwide, and only a small proportion have been studied. However, part of their biosynthetic potential has been unveiled and more than 1,600 terpenoids have been detected and identified. The study of bryophytes faces challenges due to their small size, and sociology, making it difficult to collect large amounts of uncontaminated samples. Additionally, their chemical specificity and the scarcity of chemical data specific to this branch further complicate their study. Traditionally, research on bryophytes has focused only on specific species or classes of compounds.

Methods: In contrast, our work proposes the first untargeted metabolite profiling investigation of a large collection of bryophytes (63 species) mainly issued from Corsican biodiversity. Metabolite profiling was performed by UHPLC-HRMS/MS and the data was extensively annotated using computational tools and molecular networking. This allowed us to describe in detail the chemical space covered by our collection and to establish comparisons between all the moss and liverwort species available. To validate some of the structural annotations, 3 liverworts (*Frullania tamarisci, Pellia epiphylla, Plagiochila porelloides*) and 2 mosses (*Antitrichia curtipendula* and *Dicranum scoparium*), available in larger quantities were fractionated using high-resolution semi-preparative HPLC, yielding 20 pure compounds. Five of them were newly discovered.

Results and discussion: This study highlights the main compositional differences between mosses and liverworts at the chemical class level. By analyzing given molecular network clusters, specific biosynthetic features or compounds that are characteristic of certain species are highlighted and discussed in detail.

KEYWORDS

bryophytes, liverworts, mosses, metabolomics, natural products, molecular networking, Corsican biodiversity

1 Introduction

Bryophytes are divided into three phyla: mosses (*Bryophyta*, 12,700 species), liverworts (*Marchantiophyta*, 9,000 species), and hornworts (*Anthocerotophyta*, 225 species). Taxonomically placed between green algae and vascular plants, bryophytes are the first land plants (Christenhusz and Byng, 2016). Despite having 24,000 species distributed worldwide, the number of studies on the chemistry of bryophytes is still limited. However, 2,200 compounds have been identified in such organisms, mainly volatile compounds from liverworts. A great proportion are new compounds, and many of them have proved to be biologically active (Asakawa et al., 2013a; Horn et al., 2021; Novaković et al., 2021).

Sampling represents one of the biggest challenges when working with bryophytes. They are small organisms that rarely grow in large quantities. They grow in complex communities. As a result, it is rare to find "clean" samples in the wild, i.e., free of other bryophytes, insects, and debris from vascular plants, etc. The quantity available per species is highly variable, and samples must be carefully cleaned (Asakawa and Ludwiczuk, 2013).

Among the bryophytes, the lipophilic terpenoids of liverworts have been the most extensively studied due to the existence of intracellular structures containing oil droplets, known as *oil bodies*, which are characteristic of the phylum. These specialized organelles contain a range of lipophilic molecules, mostly mono- and sesquiterpenoids. Liverworts have therefore been the subject of a great deal of research into their volatile terpenes in essential oils and lipophilic extracts. Liverworts are described as an abundant source of new natural sesquiterpenoids with a wide variety of carbon skeletons (Asakawa, 2012; Asakawa et al., 2013b).

Mosses, in contrast, have not been broadly investigated, despite being an abundant class of bryophytes. They are known to contain high levels of fatty acids, making the detection/isolation of secondary metabolites challenging. Flavonoids and some di- and triterpenoids are the most common other components detected in mosses (Asakawa et al., 2013b; Lu et al., 2019).

Bis-bibenzyls and bibenzyls are listed as remarkable and specific constituents of bryophytes because of their important therapeutic role. They are frequently found in dimeric or polymeric forms in polar solvent extracts of liverworts, such as marchantin and riccardin (Asakawa et al., 2021). Finally, nitrogenated compounds are very rare in bryophytes. Only nine compounds have been described in liverworts and none in mosses: skatole, two prenyl indole derivatives, isotachin A and B, two coriandrins, and two methyl tridentatol (von Reuß and König, 2005). Noteworthy, such data were summarized in three extensive reviews by Asakawa et al (Asakawa, 1982, 1995; Asakawa et al., 2013b).

Modern analytical approaches enable the high-throughput analysis of a large number of plant extracts from small quantities (Wolfender et al., 2019). These approaches mainly rely on the metabolite profiling of a crude extract with ultra-high-performance liquid chromatography coupled with high-resolution mass spectrometry (UHPLC–HRMS). On such a platform, data-dependent analysis enables high-resolution tandem mass spectrometry (HRMS/MS) fragmentation spectra to be recorded on most detected compounds with unprecedented sensitivity. Annotation is then carried out by spectral comparison with public libraries [Global Natural Products Social Molecular Networking (GNPS)] (Nothias et al., 2020) and/or by computational tools [SIRIUS, Taxonomically Informed Metabolite Annotation (TIMA), etc.] (Dührkop et al., 2019; Rutz et al., 2019), generating precise information on the chemical composition of the extracts. This makes it possible to study a large number of species simultaneously to define compositional traits and prioritize extracts for targeted isolation and full characterization of selected metabolites (Allard et al., 2023). Such metabolomic approaches are being used more often to study the metabolome of vascular plants (Shen et al., 2023). To date, only a few studies have been published using this kind of approach on a small collection to establish chemotaxonomic correlations (Peters et al., 2019, 2021).

The present study focuses mainly on the UHPLC–HRMS/MS metabolite profiling of a representative set of bryophytes from the Corsican biodiversity. In Corsica, over 574 species have been reported (Sotiaux et al., 2007, 2008). Only four species were chemically investigated, and this was limited to the profiling of the essential oils (Pannequin et al., 2017, 2020, 2023; Pannequin, 2019). For this, 60 species including liverworts and mosses were extracted with solvents of increasing polarity. All extracts were systemically profiled by UHPLC–HRMS/MS under the same conditions, and all the data were gathered in a massive molecular network. As only a few reference MS/MS spectra were available, targeted isolation was carried out on abundant species for the unambiguous identification of numerous metabolites, which were used to increase confidence in the annotation.

2 Results

2.1 Design and validation of molecular network

2.1.1 Generation of an extract collection of bryophytes

A total of 60 Corsican species, including 14 liverworts and 46 mosses, were collected. Additionally, three Japanese liverworts, as part of a collaboration with the Tokushima Bunri University (Japan), were added to the collection. Supplementary Table S1 summarizes the data related to the extract collection, including taxonomical information and yield (Supplementary Figure S1; Supplementary Table S1).

Bryophytes, in particular mosses, are known to produce many fatty acids. These primary metabolites are common to many vascular plants, and in this study, they were not profiled (Lu et al., 2019). They were depleted to obtain an enrichment of secondary metabolites, which are known to be more species-specific. To reduce their amounts and improve the detection of metabolites of medium polarity, the dry sample material was extracted successively with hexane (HEX), methylene chloride (DCM), and methanol (MeOH). Only the DCM and MeOH extracts were further analyzed. The means of extraction yields of liverworts and mosses are similar, respectively: 0.21% and 0.25% for HEX extracts, 0.58% and 0.53% for DCM extracts, and 1.93% and 2.01% for MeOH extracts. However, boxplots of extraction yields (Supplementary Figure S2) show high variability between species independently of the phylum. Several species appear as outliers: three liverworts (*Frullania tamarisci*, *Porella arboris-vitae*, and *Porella obtusata*) with extraction yields of 1.7%, 1.5%, and 3.4%, respectively, to DCM extracts and 2.9%, 4.7%, and 4.4%, respectively, to MeOH extracts and two mosses (*Bartramia pomiformis* and *Orthotrichum rupestre*) with 6.3% and 6.0%, respectively, to MeOH extract yields, three times more than the average.

The composition of the HEX extracts was evaluated by ^{1}H NMR profiling. This revealed that they contained mainly fatty acids. The DCM and MeOH extracts were submitted to Solid Phase extraction (SPE) to reduce the amount of residual fatty acids and highly polar metabolites (this process is referred to as "sample clean-up" and continued in Supplementary Figure S3). After sample clean-up, on average, DCM moss extracts lost 83% of their mass, while liverwort extracts lost 60%, which highlights the presence of lipids even after hexane defatting of the dried plant material. MeOH moss extracts lost 61% of their mass after cleaning, while liverwort extracts lost only 45%. These last losses may be attributed to the presence of polar residues (mainly sugars based on ^{1}H NMR profiling).

The comprehensive approach to secondary metabolite enrichment indicates that natural products of medium polarity (DCM + MeOH) comprise 12.9 mg per gram of dry liverworts and 7.8 mg per gram of dry mosses. This shows that mosses produce about half the amount of secondary metabolites compared with liverworts. In addition to the ends in the mean values, extraction yields were found to be very species-specific; for example, *P. arboris-vitae* yielded 35.9 mg per g of the dry plant (MeOH extract). For most details, refer to the boxplot figure in Supplementary Figure S2.

2.1.2 Comprehensive UHPLC–HRMS/MS metabolite profiling of bryophyte extracts

All enriched DCM and MeOH extracts were systematically profiled by reversed-phase UHPLC–HRMS/MS in positive (PI) and negative ionization (NI) modes using a generic 9-min linear gradient. This yielded 252 chromatographic profiles: 126 in PI and 126 in NI.

Despite the sample clean-up procedure, evaluation of the metabolite profile of an initial set of species still revealed residual features associated with fatty acids after 6 min of elution and sugars between 0 and 1 min of elution. To focus on secondary metabolites, only features detected within a chromatographic window of 1 to 6 min were retained for further data exploration.

Both PI and NI data were gathered in two separate feature-based molecular networks (MNs): PI, 8,843 nodes including 5,252 grouped in 370 clusters; and NI, 4,572 nodes including 2,660 nodes grouped in 190 clusters (Figure 1). For ease of discussion, clusters have been numbered from largest to smallest (PI clusters are preceded by P and NI by N).

2.1.3 Annotation of detected metabolite

All HRMS/MS spectra of the detected features were annotated using the SIRIUS software (Dührkop et al., 2015, 2021; Ludwig et al., 2020; Hoffmann et al., 2022) and, in particular, the CANOPUS module, which provides confident chemical class information using the deep neural network-based structural classification tool NPClassifier that provides pathways, superclasses, and class information (Dührkop et al., 2021).

To establish a list of individual candidate structures for all features detected with an MS^2 spectrum, a workflow that initially performed a spectral matching against an *in-silico* MS/MS database of NP spectra was used. For this, a modified cosine score was calculated between experimental and simulated spectra (Allard et al., 2016). Candidate structures were then reweighted based on the taxonomy of the source organism and in the light of previously reported occurrences gathered by the LOTUS initiative (Rutz et al., 2022). This automated workflow is known as TIMA (Rutz et al., 2019). All these annotations are available in the MASSIVE repository of the data (see *Materials and methods*).

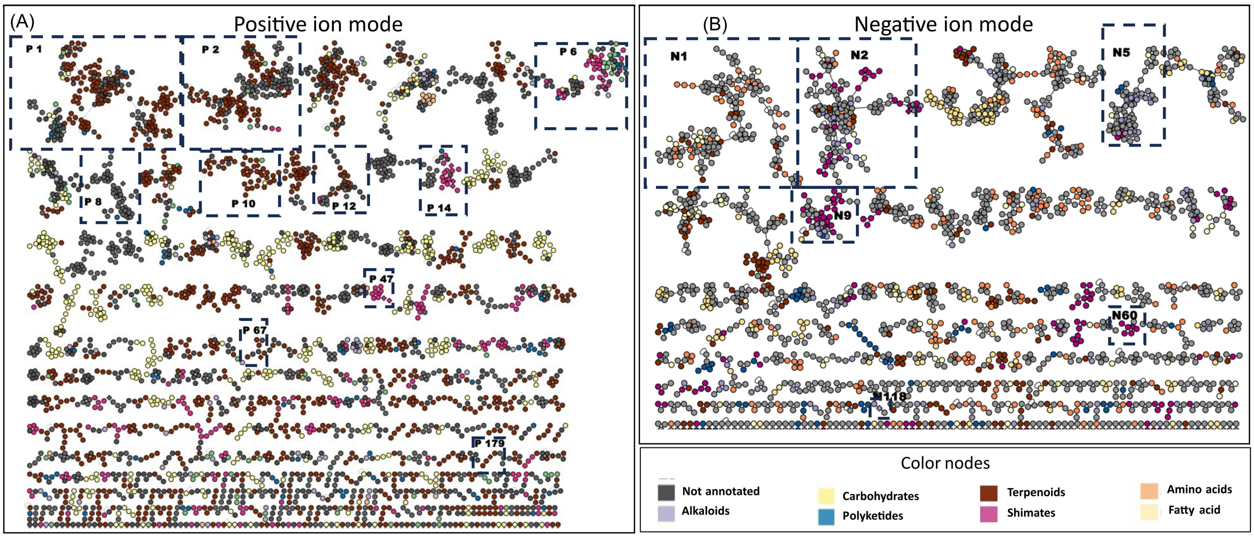


FIGURE 1
Ion identity molecular network of the DCM and MeOH extracts of bryophyte species **(A)** in PI and **(B)** in NI including clusters larger than two nodes. Nodes are colored based on the confident NPClassifier chemical pathway. Validated clusters are numbered and framed. DCM, methylene chloride; MeOH, methanol; PI, positive ionization; NI, negative ionization.

For all features of interest displayed in the figures, only the best structural candidates proposed by TIMA matching the NPClassifier superclass from SIRIUS were considered valid. In cases where TIMA was not giving coherent annotation proposals, the GNPS annotations automatically obtained during the molecular network construction were considered.

This information was mapped on the MN using color-coded nodes according to the SIRIUS pathway annotation, giving a quick overview of the chemical diversity of the extracts (Figure 1).

It is generally accepted that the annotations of pathways and superclass via SIRIUS have a high level of confidence (expressed as probability hereafter) (Dührkop et al., 2021). In this study, only pathway and superclass annotations with a probability greater than 0.8 in PI and 0.7 in NI were exploited. Using this protocol, 52% of all detected PI features could be attributed to a given NP pathway with good confidence (4,467 features in the dataset). Among them, 51% were annotated with structures derived from TIMA.

Over the whole dataset, significantly more PI than NI features were detected. While the PI mode provided detection of most compound types, for the fatty acids, shikimate, and phenylpropanoids (flavonoids, stilbenoids, etc.), better ionization and detection were obtained in NI. Thus, because of this complementary, both modes were kept in the data meaning to obtain a comprehensive overview of the metabolite repartition.

To summarize most of the annotation results obtained and to facilitate comparison across all samples, the number of features annotated in each superclass was gathered by species.

The 3,120 PI and 1,049 NI features were associated with 28 and 14 superclasses, respectively (all the 14 NI superclasses were also annotated in PI). For each sample, the feature intensities of each superclass were summed, assessing the occurrence of the superclass in each organism. To obtain a clearer understanding of the results, the proportion of chemical superclass annotations per species was displayed as a heatmap in the PI mode that was supplemented specifically for fatty acids, shikimates, and phenylpropanoids in the NI mode. For ease of reading, only superclasses whose sum area is greater than 5E5 were kept in the heatmap (Figure 2). In addition, a bar plot representing the number of annotated *vs.* unannotated features per species is displayed, indicating the richness of metabolites in each species and their annotation ratio.

The interpretation of these heatmap results and comparison across mosses and liverworts together with a detailed discussion of specific clusters of the network are provided in section 2.2.

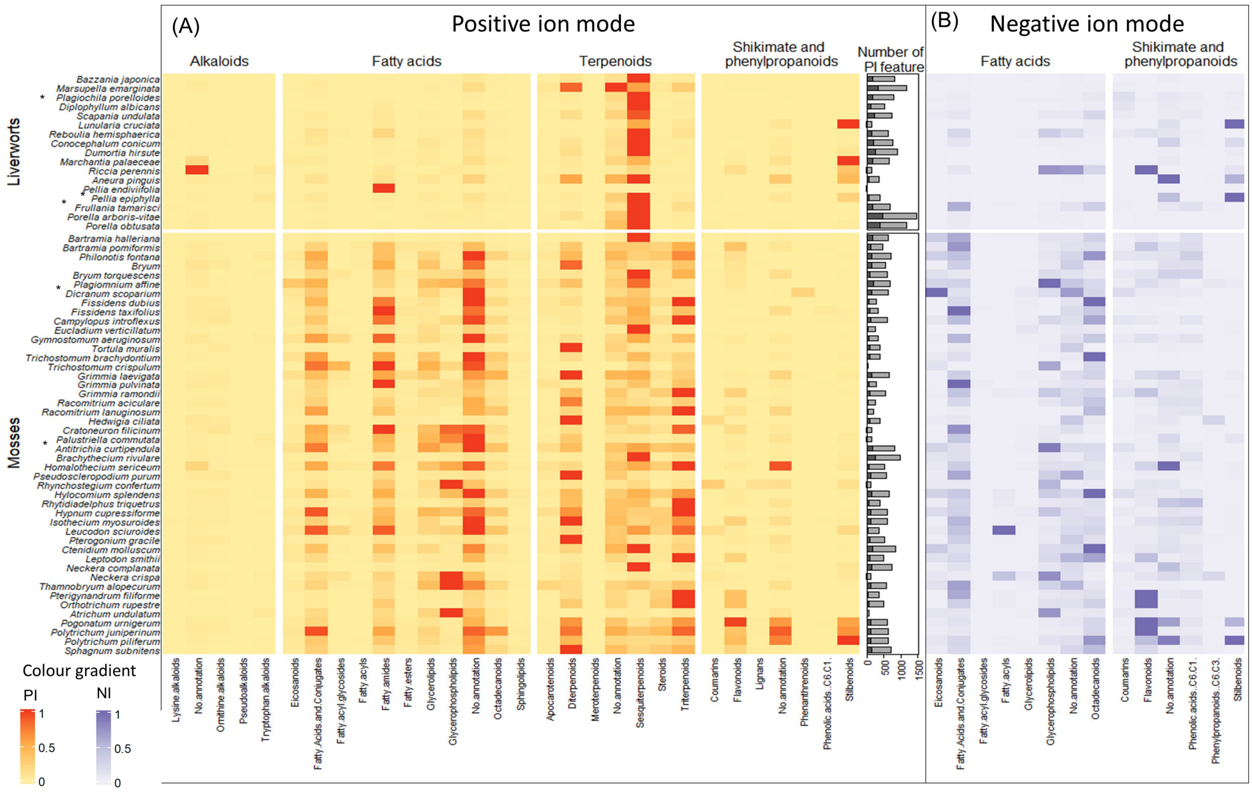


FIGURE 2
Heatmap illustrating the richness of chemical entities in the compound superclasses of bryophyte species in **(A)** PI and **(B)** NI. The color gradient corresponds to the proportion of annotated chemical superclasses in each species. Only superclasses that are the sum of the areas greater than 5E5 are represented. **(A)** The heatmap shows the richness of compound superclasses in PI. Only annotations with cosine score greater than 0.8 are shown. Compound classes with the highest abundance in the species are shown in red, and those with the lowest abundance are shown in yellow. The bar plot shows the number of features divided into annotated features in light gray and unannotated features in dark gray. **(B)** The heatmap shows the abundance of fatty acids, and shikimates and phenylpropanoids in the bryophyte species in NI. Only annotations with probability greater than 0.7 are shown. Dark blue represents the classes of compounds that are most abundant in the species and light blue those that are least abundant. PI, positive ionization; NI, negative ionization.

2.1.4 Isolation of metabolites used to validate annotations

Most of the results obtained are based on spectral comparisons using computational approaches. Before entering a more detailed discussion of the MN and the specific annotation of metabolites, targeted isolation of compounds was performed to obtain a maximum number of standards for unambiguous identification of features in the molecular networking (nodes).

The nodes corresponding to isolated standards were used as anchor points to check the consistency of annotations within MN clusters and for annotation propagation purposes (Nothias et al., 2020).

To this end, five species of interest were selected based on their MN, heatmap, and the availability of plant material, and therefore extract, available in the collection. Specifically, two liverworts, *F. tamarisci* and *Plagiochila porelloides*, known to produce terpenoids such as eudesmanolides (Pannequin et al., 2017) and plagiochilines, respectively (Pannequin et al., 2023), which are strongly represented in clusters P1 and P2 (Figures 1, 2), were chosen for the targeted isolation of the sesquiterpenoids. *Pellia epiphylla* is another liverwort producing specific stilbenoids and was selected for the isolation of perrottetin-type bis-bibenzyls (Cullmann et al., 1997).

Finally, as mosses have been little studied in the literature, two mosses, *Antitrichia curtipendula* and *Dicranum scoparium*, were selected to complete the selection. The metabolite profiling results indicated that they produce flavonoids and phenanthrenes.

To isolate the main constituents of these five bryophyte species, a generic isolation approach based on a one-step fractionation of the extracts using high-resolution semi-prep HR-HPLC was developed.

Before separation and to maximize separation efficiency, the extracts of the five selected species were subjected to a protocol allowing significant secondary metabolite enrichment. This “sample enrichment” protocol combined a two-step liquid–liquid extraction followed by SPE enrichment of the medium polarity partition (see Supplementary Figure S3). This protocol enables sugars and fatty acids to be removed efficiently on a gram scale from the crude extract.

For example, the “sample enrichment” protocol applied to 1 g of crude MeOH extract of *D. scoparium* (moss) yielded 40 mg of enriched extract (i.e., a loss of 96% of the initial mass) (Supplementary Figure S3). The enriched secondary metabolites represented less than 0.4 mg per g of dry plant material (0.04%) (for all extracts, see Supplementary Table S1).

The chromatographic conditions were first optimized on a reversed-phase C18 column on an HPLC–photodiode array (PDA)–evaporative light scattering detector (ELSD) instrument at the analytical scale. This optimized gradient was geometrically transferred to the semi-preparative HPLC scale (Guillarme et al., 2008). The enriched extracts were injected using a dry loading method to maintain high chromatographic resolution (Queiroz et al., 2019).

For ease of understanding, the code of each molecule consists of the acronym of the origins (Ac, *A. curtipendula*; Ds, *D. scoparium*; Ft, *F. tamarisci*; Pe, *P. epiphylla*; Pp, *P. porelloides*; and St, standard origin) followed by a number.

In the case of *D. scoparium*, four compounds were isolated and fully characterized by HRMS, NMR, and UV. Among them, three flavanones previously reported in *D. scoparium* were identified: apigenin 7-*O*-[2,4-di-*O*-(α-L-rhamnopyranosyl)]-β-D-glucopyranoside (Becker et al., 1986), Ds_1; 7-[(*O*-6-Deoxy-α-L-mannopyranosyl-(1→2)-*O*-[6-deoxy-α-L-mannopyranosyl-(1→4)]-β-D-glucopyranosyl) oxy]-5-hydroxy-2-(3-hydroxy-4-methoxyphenyl)-4*H*-1-benzopyran-4-one (Osterdahl, 1978), Ds_2; and kaempferol-3-β-D-(6-*O*-*trans*-*p*-coumaroyl)glucopyranoside (Tsukamoto et al., 2004), Ds_3. A new phenanthrene was also isolated (Ds_4), and the ^{1}H NMR spectrum showed the presence of three aromatic groups. A 1,3,4-trisubstituted benzene was identified from the aromatic protons at δ_H 6.93 (1H, dd, J = 9.2, 2.7 Hz, H-6), 7.12 (1H, d, J = 2.7 Hz, H-8), and 9.17 (1H, d, J = 9.2 Hz, H-5). A tetrasubstituted one was characterized by the two *meta*-coupled protons at δ_H 6.87 (1H, d, J = 2.6 Hz, H-3) and 7.29 (1H, d, J = 2.6 Hz, H-1). A singlet at δ_H 7.06 (1H, s, H-9) belongs to the third cycle. The Rotating-frame Overhauser Effect (ROE) correlations from the methoxy group at δ_H 3.89 to H-1 and H-3 positioned it in C-2; the one at δ_H 4.04 was placed in C-4 thanks to its correlation to H-3, while the third one at δ_H 4.02 was in C-10 due to its correlation with H-1 and H-9. The ROE spectroscopy (ROESY) from the hydroxyl at δ_H 9.54 to H-6 and H-8 located it at C-7. The heteronuclear multiple-bond correlations (HMBC) confirmed the identification of Ds_4 as 7-hydroxy-2,4,10-trimethoxyphenanthrene. The structure of Ds_4 with carbon numbering is shown in Figure 4, which illustrates all new structures.

The same method was applied to the other species. From *A. curtipendula*, the procedure enabled the isolation and identification of a previously undescribed phenanthrene. The ^{1}H NMR spectrum showed the presence of a 1,3,4-trisubstituted benzene characterized by the three aromatic protons at δ_H 7.20 (1H, dd, J = 9.0, 2.8 Hz, H-3), 7.33 (1H, d, J = 2.8 Hz, H-1), and 8.33 (1H, d, J = 9.0 Hz, H-4); two *ortho*-coupled protons (J = 8.8 Hz) at δ_H 7.49 and 7.59, which belong to the second ring; and two singlets at δ_H 7.18 and 7.89 positioned therefore in *para* form the third ring. The HMBCs from the two singlets, H-5 and H-8, to the carbons at δ_C 145.9 (C-7) and 147.0 (C-6) indicated the presence of hydroxy groups in these positions. The correlations from H-4 and the methoxy group at δ_H 3.88 to the carbon at δ_C 156.9 (C-2) placed the methoxyl in C-2. The ROESY correlations from H-1 to H-10, H-9 to H-8, and H-5 to H4 and from the methoxyl to H-1 and H-3 confirmed that Ac_1 was a 6,7-dihydroxy-2-methoxyphenanthrene.

Six compounds have been isolated from the liverwort *F. tamarisci*, five of which correspond to eudesmanolide sesquiterpenes and one to the known triterpene ursolic acid Ft_6. Three eudesmanolide derivatives were known in the *Frullania* genus: oxy-frullanolide Ft_2 was previously isolated in *Frullania dilatata* (Asakawa et al., 1981), and frullanolide Ft_3 and γ-cyclocostunolide Ft_4 were commonly reported as *F. tamarisci* components (Pannequin et al., 2017).

A new *cis*-fused isomer of oxo-cyclocostunolide Ft_1 was characterized. The *cis* configuration of the lactone was in agreement with the $^3J_{H6\text{-}H7}$ coupling constant of 6 Hz, whereas it was 10 Hz in the *trans*-fused lactone-like oxo-cyclocostunolide (Nadgouda et al., 1978). The ^{13}C chemical shift values of CH-6 and CH-7 were also a good indicative of the *cis* or *trans* configuration: they were reported in *trans*-eudesmanolides like arbusculin A (Fan et al., 2016) at δ_C 82.2 and 51.4, respectively, and in *cis* series like frullanolide (Ft_3) at δ_C 76.1 and 41.4, respectively. In Ft_1, the CH-6 and CH-7 carbons were observed at δ_C 75.6 and 40.6, respectively, and confirmed the identification as oxo-frullanolide.

Interestingly, a new eudesmanolide dimer Ft_5 with close similarities to muscicolide A previously described in *Frullania musciola* (Kraut et al., 1994) was also identified. In HRMS, the spectrum of Ft_5 did not show $[M+H]^+$, but the ions at *m/z* 465.3061 and *m/z* 500.3391 corresponded to $[M–H_2O+H]^+$ and $[M+NH_4]^+$, respectively. An ion at *m/z* 233.1542 appears in the HRMS spectrum and MS/MS spectra. This can be corresponded to monomeric fragments of Ft_5. Due to the high overlap of NMR signals of each monomer, the relative configuration of Ft_5 was difficult to assign. However, the NOE correlations from H-6 at δ_H 4.05 to H_3-14 at δ_H 1.09 and H_3-15 at δ_H 1.30 indicated the axial position of these protons and that they were on the same side. For the same reasons that explained those previously for oxo-frullanolide Ft_1, the ^{13}C chemical shift values of CH-6 and CH-7 (and CH-6′ and CH-7′) at δ_C 82.2 and 49.6 (and 82.2 and 51.4), respectively, indicated that H-6 and H-7 (and H-6′ and H-7′) were in a *trans* configuration. The ^{13}C NMR chemical shift values of CH-5 and CH_3-14 were also a good indicative of their relative configuration since they were observed in muscicolide A at δ_C 55.2 and 19.9 for CH-5 and CH_3-14, respectively (H-5 and H_3-14 being *trans*), and 49.3 and 21.2 for CH-5′ and CH_3-14′, respectively (H-5′ and H_3-14′ being *trans*), and in muscicolide B at δ_C 55.4 and 18.3 for CH-5 and CH_3-14, respectively (H-5 and H_3-14 being *trans*), and 45.5 and 32.0 for CH-5′ and CH_3-14′, respectively (H-5′ and H_3-14′ being *cis*) (Kraut et al., 1994). In Ft_5, the CH-5 and CH_3-14 (CH-5′ and CH_3-14′) were observed at δ_C 59.8 and 19.2 (53.4 and 21.6), respectively, indicating a *trans* configuration of H-5 and H_3-14 as well as H-5′ and H_3-14′. The ^{13}C chemical shift values of CH_3-15 (δ_C 24.7) and CH_3-15′ (δ_C 24.1) indicated that they were both in an axial configuration. Indeed, in 4-(6-hydroxy-12-oxo-11(13)-eudesmen-4-yloxy)-11(13)-eudesmen-12,6-olide, a 4-*O*-4′ dimer between 4-*epi*-arbusculin A and its open form at the lactone, isolated from the liverwort *F. tamarisci* (Toyota et al., 1998), the axial methyl CH_3-15 was observed at δ_C 22.7 and the equatorial CH_3-15′ at δ_C 31.7. Finally, the multiplicity of H-1 (d, J = 10.4 Hz) has determined its axial position. Ft_5 was a new eudesmanolide dimer that was consequently named tamariscolide A.

This targeted isolation procedure was applied to the liverwort *P. porelloides* to yield various sesquiterpenes. Plagiochiline derivatives were isolated, which are characteristic compounds of the genus. Among them, plagiochiline D Pp_1 (Nagashima et al., 1994), plagiochiline *R*-15-yl octanoate Pp_2, and plagiochiline *R*-15-yl dec-4-enoate Pp_3 (Toyota et al., 1994) were previously described in this species, and plagiochiline *R*-15-yl hexanoate Pp_4 has not yet been described. Complete structural elucidation of Pp_4 was carried out by extensive NMR analysis; it should be noted that the NMR data of Pp_4 were similar to those of Pp_2 except for the lateral carbon chain. Complementarily, the eudesmanolide, diplophyllin Pp_5, previously described in the genus *Plagiochila* (Spörle et al., 1991), was also isolated.

Finally, the same chromatographic procedure applied to the liverwort, *P. epiphylla*, allowed the isolation of four perrottetin-type bis-bibenzyls already known in the species: 10′-hydroxyperrottetin E Pe_1, perrottetin E Pe_2, 10′-hydroxy-11-methoxy-perrottetin E Pe_3, and 11-methoxy-perrottetin E Pe_4 (Cullmann et al., 1997).

These 20 isolated fully characterized compounds, as well as 12 commercial standards and three compounds previously isolated from Japanese liverworts, were analyzed under the same conditions used for extract metabolomic profiling.

These 35 standards (25 terpenoids, including 18 sesquiterpenoids, three diterpenoids, and four triterpenoids, as well as five bis-bibenzyl stilbenoids, three flavonoids, and two new phenanthrenes) enabled a formal identification of the corresponding nodes and propagation of the annotation in 11 clusters in PI and three in NI. These structures, related to MS and NMR data, are summarized in Supplementary Table S2 and Supplementary S6.

2.2 Chemical diversity in bryophytes

2.2.1 Overview of distribution of chemical classes in mosses and liverworts

To obtain a good overview of the distribution of chemical superclasses in the different samples, the data were summarized in the form of a heatmap. The heatmap displays the sum intensities of all features corresponding to a given superclass (Figure 2, see Section 2.1.3).

The chemical superclasses were divided according to their pathway (alkaloids, fatty acids, terpenoids and shikimates, and phenylpropanoids). They are represented vertically and the species concerned horizontally according to their taxonomic proximity within two sections: liverworts in the upper part and mosses in the bottom part. A color gradient, from red to yellow in PI and dark blue to light blue in NI, represents the superclass according to their abundance.

The number of features detected by species, which is a measure of the extent of the chemodiversity, varies greatly from one species to another, independently of their phylum. These variations for the PI mode are represented in a bar plot in Figure 2. The species with the most features detected are the liverworts *P. arboris-vitae*, *Marsupella emarginata*, and *P. obtusata* with more than 1,461, 1,182, and 1,167 features, respectively, and they are followed by one moss, *Brachythecium rivulare*, with 990 features. Species with fewer features detected are *Pellia endiviifolia* and *Trichostomum crispulum* with only 18 and 43 features, respectively. This variation in detected features indicates an important variability among species and no real trend between phyla. Overall, SIRIUS was able to annotate more than 80% of the features (in light gray) with at least one chemical superclass.

The heatmap in the PI mode (Figure 2) differentiates between liverworts (upper panel) and mosses (lower panel). Liverworts are mainly characterized by the presence of abundant sesquiterpenoids. Within liverwort, however, most of the thallus liverworts displayed stilbenoids principally detected in the NI mode and, in general, fewer sesquiterpenoids.

The mosses are marked by the presence of fatty acids, which were also evidenced by the extraction yield during the cleaning process. Concerning terpenoids, and in comparison with liverworts, di- and tri-terpenoids were often detected. Sesquiterpenoids were also present but were, in general, less abundant. Mosses were also found to contain flavonoids mainly highlighted in the NI mode.

In general, compounds from the shikimate and phenylpropanoid pathway are poorly represented, with the exception of stilbenoids. Those are found in liverworts but were also surprisingly detected in three moss species of the family Polytrichaceae.

Remarkably, of all the samples, only one bryophyte, the liverwort *Riccia perrenis*, produced alkaloids and, at the same time, biosynthesized only a few terpenes.

More detailed information could be retrieved from the MN. In PI, the MN generated 370 clusters of more than two nodes and 4,368 single nodes (Figure 1A). Particular attention was paid to the main cluster, which generally groups together compounds of the same structural and indicates the main trend in chemical composition. The SIRIUS annotation shows that the first 16 clusters (first two lines of the MN) correspond to terpenoids (eight clusters: P1, P2, P3, P5, P9, P10, P11, and P12), fatty acids (three clusters: P4, P7, and P15), and stilbenoids (two clusters: P6 and P13). As suggested by the heatmap, the MN analysis shows that the most important clusters common to mosses and liverworts are annotated by SIRIUS as terpenoids (e.g., P1 and P9) and fatty acids (e.g., P4). In NI, the MN is made up of 4,604 nodes grouped into 190 clusters of more than two nodes and 2,245 single nodes (Figure 1B). In this mode, the annotations were less reliable, except for stilbenoids and flavonoids, which ionize better. MN reveals a few homogeneous clusters from the shikimate and phenylpropanoid pathway (e.g., N60) and some flavonoids (e.g., N9). The annotations for the other clusters were more heterogeneous and therefore difficult to use at this level.

2.2.2 Investigation of selective chemical classes and taxonomical relationships

To obtain a more detailed view of the metabolite composition of the bryophyte species in the collection, the annotations will be analyzed by compound class, starting with those with the greatest number of features.These data will be presented in relation to the taxonomic relationships that exist between species, particularly between the two phyla, and will be discussed in greater detail in light of the published data in the Discussion section.

2.2.2.1 Terpenoids

In the MN, analysis in the PI mode made it possible to associate 3,631 nodes with terpenoids (1,561 sesquiterpenoids, 643 diterpenoids, and 604 triterpenoids). Sesquiterpenoids are the largest group in liverworts, while di- and triterpenoids are the most represented groups in mosses. This can be seen on the heatmap in terms of intensity but also in terms of frequency of annotation in the MN.

Volatile compounds such as monoterpenoids are not discussed in this work.

2.2.2.1.1 Sesquiterpenoids

Sesquiterpenes were associated with 1,561 nodes, which were mainly detected in liverworts (with 503 specific nodes), while only 70 were specific to mosses.

As shown in Figure 3A, the MN cluster P1 is the largest (211 nodes) and is mainly specific to the liverwort *P. porelloides*. The confidence of annotations in P1 was improved by an unambiguous identification of all nodes corresponding to the plagiochiline derivatives isolated from *P. porelloides*: plagiochiline D Pp_1, plagiochiline *R*-15-yl octanoate Pp_2, plagiochiline *R*-15-yl(4*Z*)-dec-4-enoate Pp_3 and a new derivative with a hexanoyl chain in C-15, plagiochiline *R*-15-yl hexanoate Pp_4.

While most nodes were only found in *P. porelloides*, two distant branches of P1 were shared by several species of moss and liverwort. It should be noted that some of these nodes are specific to the species *Thamnobryum alopecurum*, a moss for which no phytochemical studies are available.

Cluster P2 is composed of 168 nodes that were also mostly annotated as sesquiterpenoids, with the eudesmane, caryophyllene, and guaiane skeletons (Figure 3B). The annotations of this cluster were consolidated by HPLC–MS/MS data obtained on eight isolated components or standards: parthenolide St_4, costunolide St_5, dehydrocostus lactone St_6, diplophyllin Pp_5, telekin St_3, a new dimeric eudesmane sesquiterpene lactone (an isomer of muscicolide), and tamariscolide Ft_5 as well as frullanolide Ft_3, alantolactone St_1, and isoalantolactone St_2. These last three isomeric compounds however were coeluted and grouped under one node. This cluster was associated with sesquiterpenes of the γ-lactone type existing in the form of monomer or dimer as is the case for Ft_5, a dimer of γ-cyclocostunolide Ft_3. Among P2, 122 nodes are common to both phyla, and 40 are specific to liverworts while six to mosses. It is noticeable that the P8 cluster, which includes 51 nodes, brings together ions resulting from the in-source fragmentations of sesquiterpenes such as alantolactone St_1, frullanolide Ft_3, γ-cyclocostunolide Ft_4, and dehydrocostus lactone St_6. Among the sesquiterpene γ-lactones, the oxy-frullanolide Ft_2 was associated with cluster P12, holding 33 nodes annotated as sesquiterpenoids. It is noticeable that two other eudesmanolides, oxo-frullanolide Ft_1 and γ-cyclocostunolide Ft_4, were not linked to any cluster and appeared as single nodes when the precursor ion of the MS/MS was the $[M+H]^+$. As mentioned, however, their in-source fragments are clustered in P8 (Figure 3D).

Cinnamolide St_8 was found in the P67 cluster, which holds several drimane-type sesquiterpene lactones. These compounds were identified in *P. arboris-vitae* and *P. obtusata*. However, cinnamolide St_8 and its derivatives were detected for the first time in mosses of the genus: *Dicranale*, *Hedwigiale*, *Bryale*, *Polytrichale*, *Hypnale*, *Grimmiale Bartramiale* and *Polytrichale* and in liverworts of order: *Marchantiale* and *Jungermanniale* (Figure 3I).

2.2.2.1.2 Di- and triterpenoids

Several clusters (P9, P33, P37, etc.) were associated with diterpenoids. Only P179, a small cluster including only four nodes that were shared by both mosses and liverworts, was confirmed by a standard clerod-3,13(16),14-trien-17-oic acid St_9 (Figure 3J).

Despite the high similarity of betulin St_12 and its acid St_13, two lupane derivatives, only betulin St_12 was associated with cluster P10. This cluster comprises 47 nodes mainly annotated as lupane-type triterpenoids (Figure 3E). In PI, St_13 appears in a

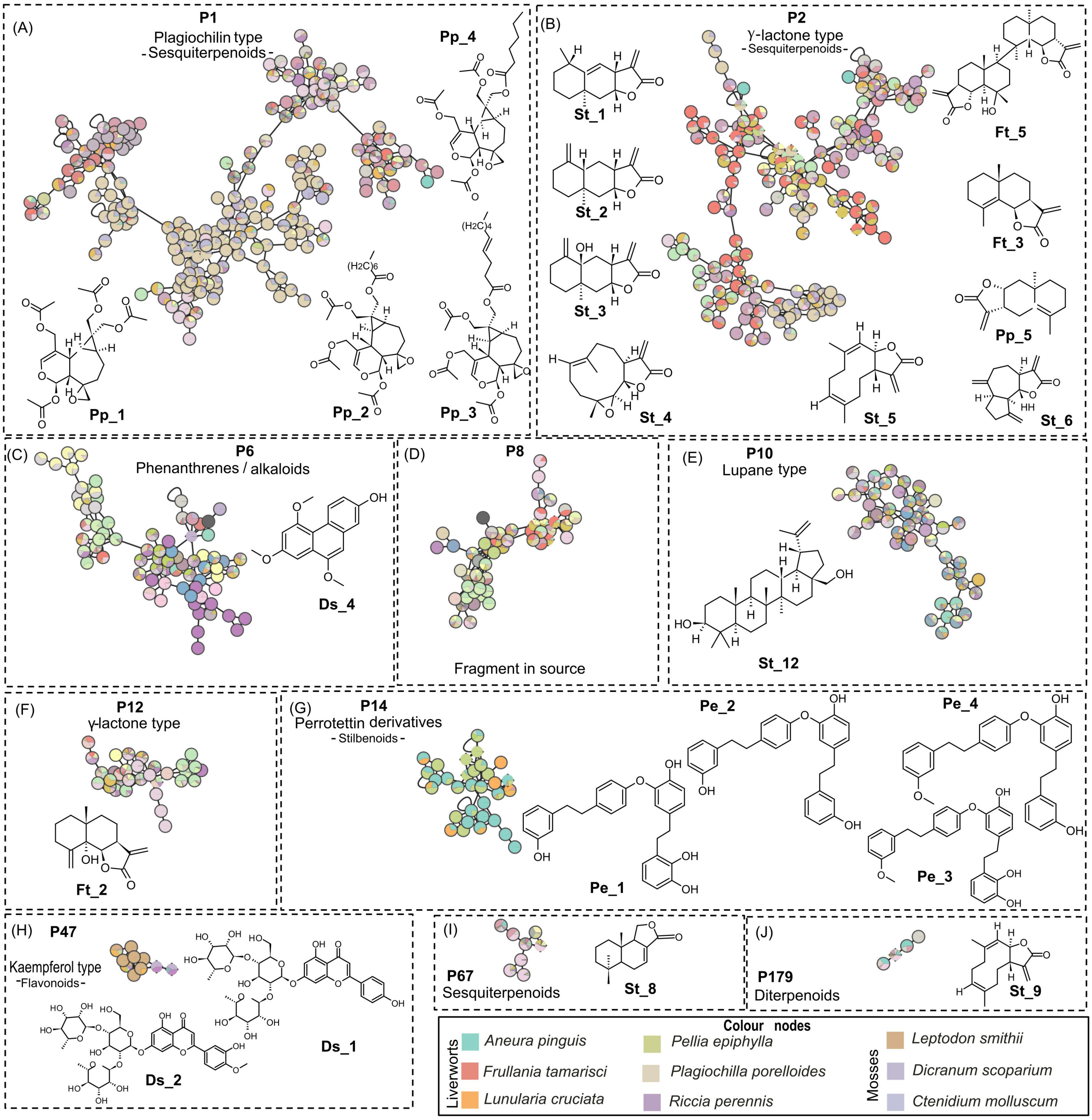


FIGURE 3
Selected PI clusters from Figure 1, validated by standard or isolated compounds: **(A)** P1, **(B)** P2, **(C)** P6, **(D)** P8, **(E)** P10, **(F)** P12, **(G)** P14, **(H)** P47, **(I)** P67, and **(J)** P179. The pie charts within node indicate the distribution of MS intensities of features across the species. Nodes with single color are characteristic of metabolites found only in a specific species. Nodes with multiple colors denote shared metabolites. Clusters shared by multiple species highlight either specific shared features and/or shared chemical classes. Species-specific clusters correspond to one color node. Color node codes are available in the Supplementary Material. PI, positive ionization.

single node while in a very small cluster (two nodes) in NI with ursolic acid Ft_6, another triterpenoid. The MN showed that P10 was common with mosses and liverworts, while betulin has been previously described only in the liverwort, *P. epiphylla*, and betulinic acid in two mosses, *Heteroscyphus coalitus* and *Ptilidium pulcherrimum*.

Tamariscolide Ft_5 isolated from *F. tamarisci* is mislabeled as a triterpenoid on the basis of its crude formula and was finally identified as a dimer of γ-cyclocostunolide, a lactone sesquiterpene. Ft_5 shares the P2 cluster associated with monomeric or dimeric sesquiterpene lactones.

2.2.2.2 Shikimate and phenylpropanoids

2.2.2.2.1 Stilbenoids

In PI, the MN shows 800 nodes annotated to shikimates and phenylpropanoids, and among them, 113 were stilbenoids. The annotations of cluster P14 were ensured by four bis(bibenzyl)-type perrottetin derivatives (Pe_1, Pe_2, Pe_3, and Pe_4) isolated from *P. epiphylla* (Figures 3G, 5A). This cluster included mainly nodes with molecular weights between 427.19 and 476.20 corresponding to perrottetin derivatives but also other nodes with molecular weights of 800 or higher, suggesting perrottetin dimers. It is noticeable that cluster P14 was mainly shared by four thalloid liverworts: *P. epiphylla*, *Lunularia cruciata*, *Marchantia paleacea*, and *Aneura pinguis*. The perrottetin derivatives have been previously described in these species, except for *A. pinguis*, for which no data were available in the literature.

The MS signal in NI of stilbenoids was better than in PI, as clearly seen on the heatmaps and MN (Figures 1B, 2B). A large N2 cluster grouped perrottetin derivatives from *P. epiphylla* (Pe_1, Pe_2, Pe_3, and Pe_4) and riccardin G St_15, which appears as a single node in PI (Figure 5B). Suggesting marchantin-type compounds, such as marchantin A, C, G, H, and M, marchantinquinone, and riccardin C and F, the TIMA annotations were highly relevant in this case. In addition, analysis of the UV spectra of these compounds showed absorbance between 220 and 307 nm, confirming the presence of molecules with similar frameworks.

Cluster N5 was annotated as alkaloids, and shikimates and phenylpropanoids, with annotation probabilities between 0.8 and 0.5. N5 displayed 63 nodes mainly specific to *Bazzania japonica*. The SIRIUS annotations were not reliable, while TIMA proposed several chlorinated macrocyclic bis-bibenzyls belonging to bazzanin with a good final score. TIMA annotations appear to be reliable because these compounds were specific to the rare and little-studied *Bazzania* genus (Martini et al., 1998; Scher et al., 2003). In addition, analysis of the isotopic pattern confirms the presence of one or more chlorines, depending on the molecule.

2.2.2.2.2 Flavonoids

The isolation procedure carried out in the present study enabled us to purify three flavanones: Ds_1, Ds_2, and Ds_3. In positive MN, Ds_1 and Ds_2 belong to a small P47 cluster (10 nodes). The TIMA tools provide seven annotations, including five glycosylated kaempferols (Figure 6A). The two triglycosylated apigenins, Ds_1 and Ds_2, were the most intense in *D. scoparium*, while the other nodes were detected in the moss *Leptodon smithii* and mainly annotated to kaempferol derivatives. As noted, the chemical composition of the moss *L. smithii* has never been studied. Ds_3, a kaempferol derivative, has a very weak signal in PI but belongs to a small N60 cluster (seven nodes) in NI (Figure 6B). For N60, GNPS offers annotations corresponding to three kaempferol derivatives, which were ensured by absorptions at 254–280-nm and 340–360-nm characteristics to flavones and flavonols in the UV spectra. Among these nodes, two are specific to *T. alopecurum*, one to *Plagiomnium affine*, and one to *M. emarginata*. For these three species, flavonoids have never been reported in the literature, and, in particular, no data are available for *T. alopecurum*. However, as reported in the literature (Asakawa et al., 2013a), triterpenoid compounds from *P. affine* and sesquiterpenoids from *M. emarginata* were detected in our study, but only in weak ions.

2.2.2.2.3 Phenanthrenes

In PI, the large P6 cluster is partly annotated as phenanthrenes. One of these nodes corresponds to Ds_4, a new phenanthrene derivative isolated from *D. scoparium* (Figures 3C, 7A). In negative mode, Ds_4 validates the little N118 cluster (three nodes), which includes three specific nodes to the mosses *T. alopecurum*, *P. affine*, and *D. scoparium* and one to the liverwort *M. emarginata*. The undescribed phenanthrene Ac_1 isolated from *A. curtipendula* displayed cluster N9 (43 nodes), which indiscriminately represents liverworts and mosses. Within cluster N9, most nodes were assigned to several species, but a few nodes were specific to *T. alopecurum*, *M. paleacea*, and *B. japonica*. For the chemical class phenanthrenes, the SIRIUS and TIMA annotations proved to be unreliable (Figure 7B).

2.2.2.3 Alkaloids

Nitrogenated metabolites are rare in bryophytes. Our metabolomic approach shows that 2.5% of 8,443 nodes (217) were annotated as alkaloids with a higher probability (confidence score of 0.8). Among them, 59 and 99 were specific to mosses and liverworts, respectively. Alkaloids shared the P6 cluster with phenanthrenes, so they do not form a specific cluster; however, with 13 nodes under 77, they represented a homogenous part of P6 (Figure 7A). Among those, nine nodes are specific to the liverwort *Riccia perennis*; they correspond to the major peaks observed in Charged Aerosol Detector (CAD). This trend was confirmed in the heatmap (Figure 3). It is noteworthy that five nodes of the P6 cluster displayed a molecular formula containing one nitrogen and one or two sulfurs like the compounds identified from the liverwort *Corsinia coriandrina* (von Reuß and König, 2004, 2005). Inspection of the isotopic pattern of these nodes confirms the presence of one or more sulfur atoms, depending on the case. These compounds also show a base peak at *m/z* 206.06 reported in the *O*-methyltridentatols of *C. coriandrina*. However, further analysis is required to confirm compound annotations of the cluster.

3 Discussion

The study is the first to investigate a large collection of bryophyte extracts under the same analytical conditions. These samples represent 60 species, 15 orders, and 41 families, i.e., 1/3 of French bryophyte orders and 10% of all Corsican bryophyte species. The systematic study of the metabolomes of this representative collection has made it possible, for the first time, to determine the full range of chemotaxonomic relationships within the mosses and liverworts and between these two phyla.

Mosses are understudied in the literature because they are mainly composed of fatty acids (Asakawa et al., 2013a).

The secondary metabolite enrichment procedure developed in this study proved to be an effective solution for overcoming this problem and improving the detection of target compounds.

The loss of mass of raw moss extracts is 72% on average, whereas it is only 52.5% for liverwort extracts.

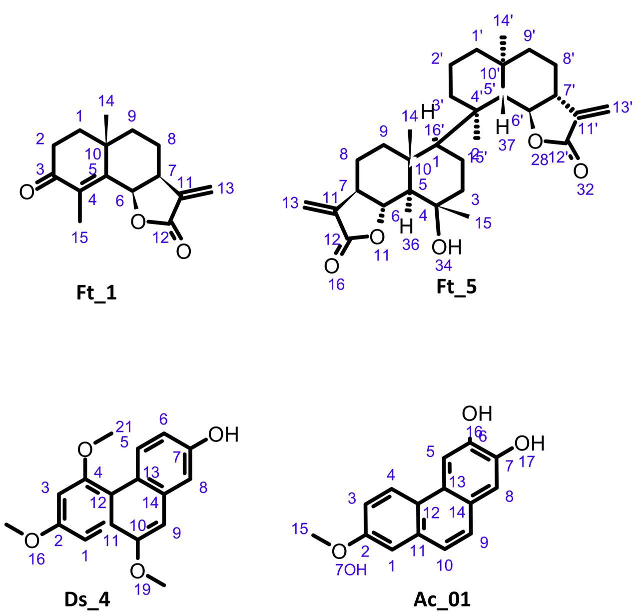


FIGURE 4
Structure of undescribed compounds: oxo-frullanolide Ft_1, tamariscolide Ft_5, plagiochiline *R*-15-yl hexanoate Pp_4, 7-hydroxy-2,4,10-trimethoxyphenanthrene Ds_4, and 6,7-dihydroxy-2-methoxyphenanthrene Ac_01.

Although fatty acids remained dominant in the moss extracts, di- and triterpenoids were detected as an important chemical superclass, except for three mosses belonging to the family Polytrichaceae (*Pogonatum urnigerum*, *Polytrichum juniperium*, and *Polytrichum piliferum*), which produce flavonoids and stilbenoids. In the literature, mosses produce small quantities of flavonoids and very few stilbenoids, and the presence of large quantities of flavonoids and stilbenoids in this family makes it unique.

According to the literature (Asakawa et al., 2013a), liverworts produce many sesquiterpenoids. Our data confirm the high proportion of sesquiterpenes in this phylum. They distinguish thallus liverworts, characterized by a high proportion of stilbenoids, from leaf liverworts, which tend to produce more sesquiterpenoids. The sesquiterpenoids have formed homogeneous and specific clusters in the MN.

Among them, a large P1 cluster was found to be almost unique to the hepatic *P. porelloides* and encompasses numerous plagiochiline derivatives for which four (Pp_1, Pp_2, Pp_3, and Pp_4) were unambiguously identified after isolation.

Plagiochiline derivatives have recently been reported for their cytotoxic activities (Vergoten and Bailly, 2023). The richness of P1 indicates a possible discovery of additional bioactive plagiochilines. As more than 1,600 liverworts of the genus have been reported, they can be considered a good source for such compounds.

A large homogeneous cluster P2 was assigned to the γ-sesquiterpene lactones, which included many species of liverworts and mosses. Sesquiterpene lactones are known for their diverse biological activities, as potent herbicides, insecticides, anticancer agents, and antifoulants, and for their wide applicability in the food and perfume industries (Sartori et al., 2021). The literature describes the presence of active γ-lactones preferentially in liverworts. Our MN analysis showed clusters of γ-lactones associated with both phyla, opening new avenues of investigation.

In the thallus liverworts, the main chemical superclass was stilbenoid, in good agreement with the literature (Asakawa et al., 2021). In our data, clusters P14 and N2 were validated with five isolated compounds corresponding to the bis-bibenzyl stilbenoids, which are known to have significant biological activities. They were both homogenous to the thallus liverworts. Among the species, the occurrence of bis-bibenzyl stilbenoids in *L. cruciata* and *M. paleacea* is well known; however, *A. pinguis* appears to be a high potential species for finding novel bis-bibenzyl stilbenoids (numerous specific nodes).

Finally, the liverwort *R. perennis* appears as a very original species because of the presence of alkaloids. This result is remarkable given the scarcity of nitrogen compounds described in bryophytes. Only nine alkaloids were reported in bryophytes. Their presence in *R. perennis* is described here for the first time. Several of our annotations corresponded to alkaloids that also contain sulfur, in which the proposed molecular formula does not correspond to any nitrogen–sulfur compounds already described in bryophytes. Sulfur alkaloids are common in marine organisms and have shown a wide variety of biological activities: cytotoxicity, antiviral, anti-inflammatory, and antioxidant (Zhang et al., 2023). In the bryophytes, only isotachin A, isotachin B, coriandrins, and methyl tridentatols, four sulfur-containing alkaloids, are known in very rare thallus liverworts (von Reuß and König, 2004, 2005). Consequently, the *Riccia* genus, which includes 19 species in Corsica, appears as a promising reserve of sulfur-containing alkaloids.

The present article shows an initial chemical mapping of the library of the bryoflora extracts, supported by important clusters validated with pure compounds and focusing on species with high interest. Our metabolomics innovative approach has produced a

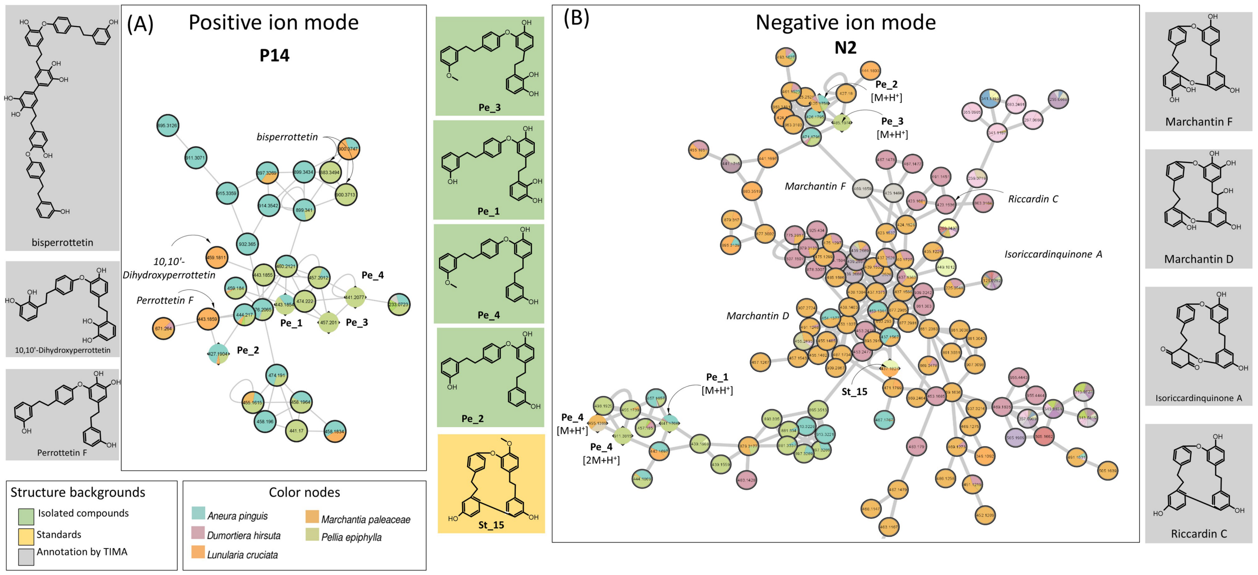


FIGURE 5
Confirmed cluster of perrottetin derivatives **(A)** in PI (P14) and **(B)** NI (N2). Confirmed annotations appear on a green background (isolated compounds) or an orange background (standard compounds), and selected annotations from TIMA tools appear on gray background. PI, positive ionization; NI, negative ionization; TIMA, Taxonomically Informed Metabolite Annotation.

large dataset that could be the subject of more detailed work, as well as an approach targeting biological active ingredients.

A great deal of information is still unexploited and will be the subject of future work. All these data are available on the GNPS platform so that everyone can approach the chemistry of bryophytes in a new way and advance the knowledge of these plants whose full potential is not yet known.

4 Materials and methods

4.1 Plant materials

4.1.1 Harvesting

Samples of the 60 bryophytes were collected at random locations in Corsica. Sampling was performed in 2021. Botanical determination was performed according to the botanical determination keys summarized in Bryophyte Flora (Augier, 1966) by Achille Pioli, a specialist in mosses, and voucher specimens were deposited in the herbarium of the University of Corsica, Corte (France). Three Japanese liverworts collected as part of a collaboration with the Tokushima Bunri University, Japan, were added to the collection in Tokushima. The sample numbers, the geographical origin of the different samples, and the voucher codes for each specimen analyzed are listed in Supplementary Table S1. All samples were harvested in the wet season of 2021 (January–April and September–December 2021).

4.1.2 Extractions

After 15 days of drying at room temperature, plant material was powdered by cryo-grinding and successively extracted with hexane, methylene chloride, and methanol for 24 h each. Solvent volume corresponded to v = mass of sample × 100. The extracts were dried using a rotary evaporator or lyophilization. The masses and the yields obtained are summarized in Supplementary Table S1.

4.2 Cleaning of extracts

To reduce the amount of apolar compounds, such as fatty acids, and polar compounds such as saccharides, the extracts were cleaned by sample clean-up before HRMS[electrospray ionization (ESI)]–MS analysis and sample enrichment before the isolation step. This protocol corresponds to Supplementary Figure S2.

4.2.1 Sample clean-up

The solid-phase extraction procedure used C18 cartridges (50 µm, 12 mL, 1,000 mg; Finisterre, Teknokroma, Spain) in conjunction with a Teknokroma extraction manifold system (12-position manifold with a 13 × 75 mm test tube rack). The vacuum pressure on the manifold was maintained at ≤5 inches (12.7 mmHg) throughout the duration of the SPE protocol. The cartridges were conditioned using 10 mL of methanol–water (50:50) and 10 mL of pure methanol.

A sample with a mass of 50 mg was dissolved in 5 mL of mixture MeOH:H_2O 95:5. Analytes were eluted with 10 mL methanol–water (95:5). Eluates were evaporated using Genevac. The samples were reconstituted in methanol at 5 mg·mL^{-1}, and they were transferred into a microplate for LC–MS/MS analysis.

4.2.2 Sample enrichment

The selected extracts were dissolved in a mixture of MeOH/H_2O (7:3) at 5 mg·mL^{-1}. They were extracted by liquid–liquid extraction (LLE) with hexane to equal volume. The two phases were separated,

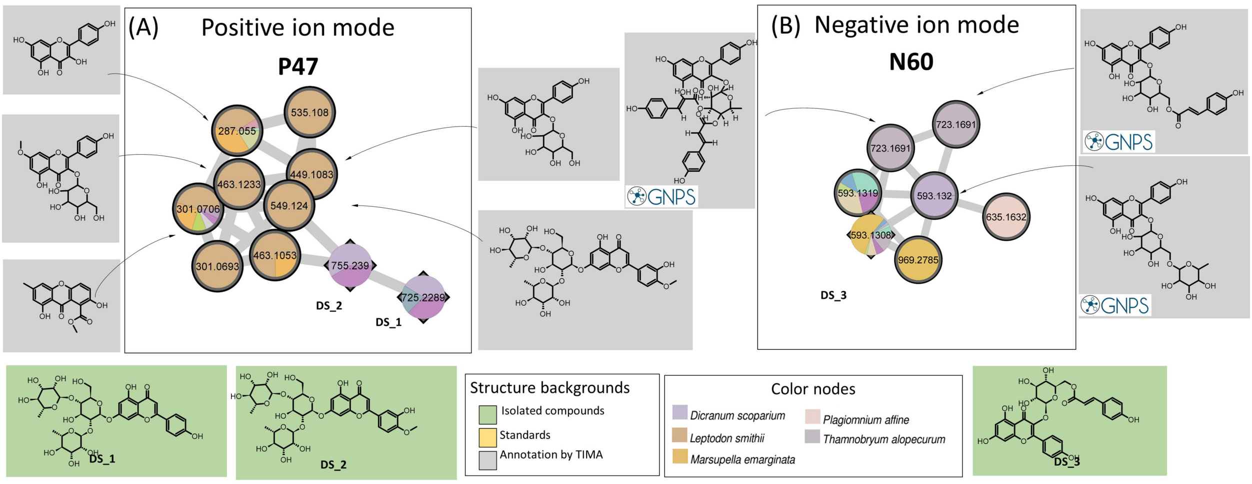


FIGURE 6
Confirmed cluster of flavonoid derivatives **(A)** in PI (P47) and **(B)** NI (N60). Confirmed annotations appear on a green background (isolated compounds) or an orange background (standard compounds), and selected annotations from TIMA tools appear on gray background. PI, positive ionization; NI, negative ionization; TIMA, Taxonomically Informed Metabolite Annotation.

dried, and evaporated. The extract obtained from the organic phase (MeOH/H_2O) was dissolved in butanol at 5 mg·mL^{-1} and extracted by LLE with water to equal volume. After drying and evaporation, the extract obtained from the organic phase was cleaned by SPE.

SPE C18-cartridges (50 µm, 12 mL, 1,000 mg; Finisterre, Teknokroma, Spain) were conditioned with a 10-mL mixture of methanol–water (50:50) and 10 mL water pure. Samples were mixed with deactivated silica and disposed of on the head of the SPE cartridge after conditioning. Samples passed into the cartridges at a flow rate of approximately 5 mL/min under vacuum. Analytes were eluted successively with 10 mL of methanol–water (90:10), 10 mL of water, and 10 mL of ethyl acetate. Eluates were evaporated using Genevac.

4.3 Mass spectrometry analysis

4.3.1 UHPLC–PDA–ELSD–(Q)MS metabolite profiling

Analysis and chromatographic data were obtained on an ultra-high-performance liquid chromatography system equipped with a photodiode array, an evaporative light-scattering detector, and a single quadrupole detector using heated electrospray ionization (UHPLC–PDA–ELSD–QDA) (Waters, Milford, MA, USA). The ESI parameters were as follows: capillary voltage 800 V, cone voltage 15 V, source temperature 120°C, and probe temperature 600°C. The acquisition was performed in positive ionization mode

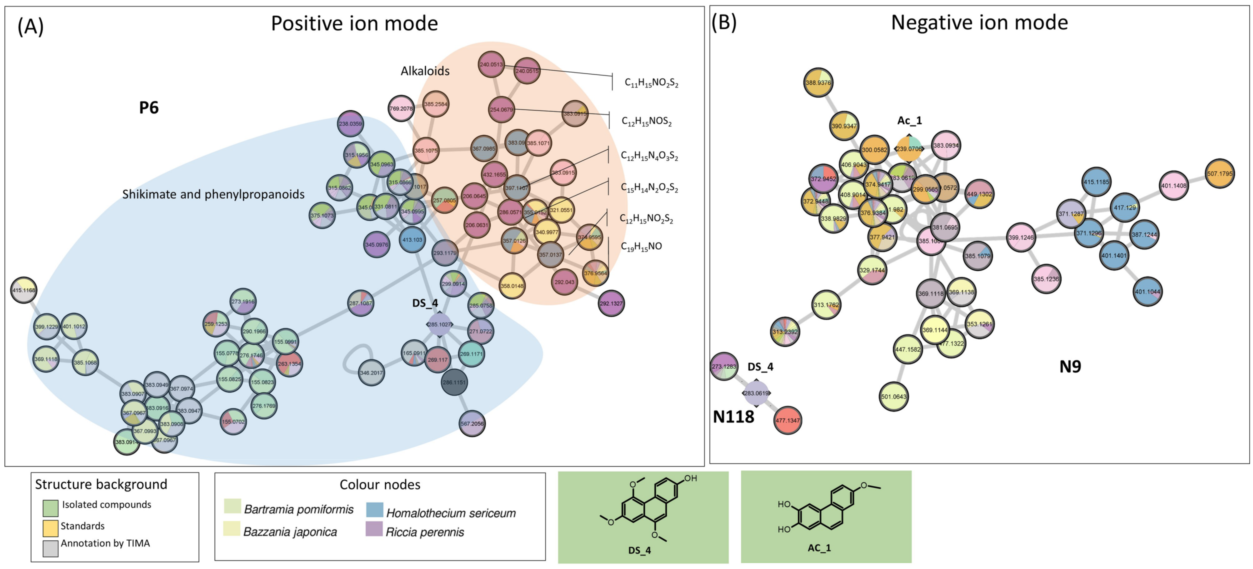


FIGURE 7
(A) Cluster sharing of phenanthrene and alkaloids in PI (P6). Nodes annotated with shikimates and phenylpropanoids by NPClassifier on blue background and with alkaloids on orange background. **(B)** Validated cluster of phenanthrene derivatives in NI (N118 and N9). Confirmed annotations appear on a green background (isolated compounds). PI, positive ionization; NI, negative ionization.

with an m/z range of 150–1,000 Da. The chromatographic separation was performed on an Acquity UPLC BEH C18 column (50 × 2.1 mm i.d., 1.7 µm; Waters) at 0.6 mL/min, 40°C with H_2O (A) and MeCN (B), both containing 0.1% formic acid as solvents. The gradient was carried out as follows: 5%–100% B in 7 min, 1 min at 100% B, and a re-equilibration step at 5% B for 2 min. The ELSD temperature was fixed at 45°C, with a gain of 9. The PDA data were acquired from 190 to 500 nm, with a resolution of 1.2 nm. The sampling rate was set at 20 points/s.

4.3.2 UHPLC–HRMS/MS metabolite profiling

Analysis, data processing, and feature-based molecular network generation chromatographic data with high-resolution MS were obtained on a Waters Acquity UHPLC system equipped with a Q-Exactive Focus mass spectrometer (Thermo Scientific, Bremen, Germany), using heated electrospray ionization source (HESI-II). The chromatographic separation was carried out on an Acquity UPLC BEH C18 column (50 × 2.1 mm i.d., 1.7 µm; Waters) at 0.6 mL/min, 40°C with H_2O (A) and MeCN (B), both containing 0.1% formic acid as solvents. The gradient was carried out as follows: 5%–100% B in 7 min, 1 min at 100% B, and a re-equilibration step at 5% B in 2 min. The ionization parameters were the same as those used in Rutz et al. (2019).

4.3.3 UHPLC–HRMS/MS data processing

The raw UHPLC–HRMS/MS files were converted into mzXML files using the MSConvert software. The mzXML files were then processed using the open software MZmine (3.4.16) (Schmid et al., 2023). The mass detection step was performed using a centroid mass detector with a noise level set at $1E^6$ for MS1 and $1E^4$ for MS^2 in PI and $1E^4$ for MS1 and $1E^4$ in MS^2 in NI. The ADAP chromatogram builder was employed with a minimum group size of scans of 4, a minimum group intensity threshold of $1E^6$, a minimum highest intensity of $1E^6$ in PI and $5E^4$ in NI, and an m/z tolerance of 10 ppm. The deconvolution was carried out with the ADAP (Wavelets) algorithm, using a signal-to-noise threshold of 50, a minimum feature height of $1E^6$, a coefficient/area threshold of 100, a peak duration range of 0.01–0.9 min, and a wavelet range between 0.01 and 0.08 min. The m/z and retention time (RT) for MS^2 scan pairing were, respectively, set to 0.005 Da and 0.1 min. The isotopes were grouped using the isotope peak grouper algorithm with an m/z tolerance of 3 ppm, an RT tolerance of 0.05 min, and a maximum charge of 2, using the most intense isotope as the representative one. The alignment was carried out with the join aligner with an m/z tolerance of 15 ppm, an RT tolerance of 0.1 min, and a weight tolerance for m/z and RT of 10 each. Ion identity networking parameters were set to m/z tolerance, 0.002 m/z or 5 ppm; check, one feature; min height, $1E^3$ with ion identity library parameters set to MS mode, positive; maximum charge, 2; maximum molecules/cluster, 2; adducts, M+H, M+Na, M+K; modifications, M–H_2O, M–NH_3.

4.3.4 Molecular network generation

The MZmine aligned table was exported in MGF format for the processing of the Feature-based Molecular Networking (FBMN). The spectral data were uploaded on the GNPS platform (Nothias et al., 2020). A network was generated with a minimum cosine score of 0.85 and a minimum of five matching peaks. The experimental spectra were searched against GNPS's spectral libraries. The obtained network was visualized in the software Cytoscape (3.9.1, Institute for Systems Biology, Seattle, WA, USA) (Smoot et al., 2011).

The mass spectrometry data were deposited on the MassIVE public repository nos. MSV000093186 (PI) and MSV000093188 (NI) with different GNPS job parameters, and resulting data are available at the following addresses:

PI network: ID=895c9f23e6df4c30a42d8774af4121c1.
NI network: ID=5e783c91e5e74ed9aee07cc59fc12830.

4.3.5 Class annotation for the dataset

Filtered features detected in PI and NI in the general dataset were annotated using a computational approach integrating SIRIUS (molecular formula) (Ludwig et al., 2020), CSI:fingerID (probabilistic molecular fingerprint by machine learning substructure prediction and *in silico* annotation) (Dührkop et al., 2015), and CANOPUS (systematic class annotation) (Dührkop et al., 2019, 2021). In the second step, the spectral file and attribute metadata obtained after the MN step were annotated by matching the MS1 and MS^2 spectra data with the LOTUS-ISDB [in-house database containing the *in silico* fragmentation spectra of all the compounds present in the Dictionary of Natural Products (DNP) and LOTUS databases] complemented with structure–organism pairs coming from the DNP (Allard et al., 2016; Rutz et al., 2019). The following parameters were used: spectral match parameters: parent mass tolerance 0.01 Da, MS/MS tolerance 0.01 Da, minimum cosine score 0.2, and minimum peaks 6. Spectral match of MS/MS spectra against the database provided a list of 50 chemical structure candidates for every feature. The candidates were re-ranked by taxonomic reweighting after ponderation of their spectral score inversely proportional to the taxonomic distance between the biological source of the candidate and that of the one-off analyzed sample(s) in which the feature is detected.

4.3.6 Chromatographic optimization and semi-preparative HPLC–UV isolation

The separation conditions of the DCM or MeOH extracts were optimized on an HP 1260 Agilent high-performance liquid chromatography equipped with a photodiode array detector and an ELSD detector (HPLC–PDA–ELSD) (Agilent Technologies, Santa Clara, CA, USA). The chromatographic separation was performed on an XBridge C18 column (250 × 4.6 mm i.d., 5 µm; Waters) equipped with a C18 pre-column at 1 mL/min, with H_2O (A) and MeCN (B), both containing 0.1% formic acid as solvents. The UV absorbance was measured at 280 and 360 nm, and UV–Vis spectra were recorded between 190 and 600 nm (step 2 nm). The optimized gradient used for the *F. tamarisci* DCM extracts and *P. porelloides* MeOH extracts was as follows: 5 mn at 30% B, 30%–100% B in 45 min, and 10 min at 100% B. The optimized gradient *P. epiphylla* DCM extract used for the was as follows: 5 mn at 35% B, 35%–100% B in 45 min, and 10 min at 100% B. The optimized gradient *A. curtipendula* MeOH extract used for the was as follows: 5 mn at 15% B, 15%–75% B in

45 min, and 10 min at 100% B. The optimized gradient *D. scoparium* MeOH extract used for the was as follows: 5 mn at 15% B, 15%–75% B in 45 min, and 10 min at 100% B.

These chromatographic methods were geometrically transferred (Guillarme et al., 2008) to the semi-preparative scale on a Shimadzu system equipped with an LC-20, module pumps, an SPD-20 A UV/VIS, a 7725I Rheodyne® valve, and an FRC-40 fraction collector (Shimadzu, Kyoto, Japan). The separation was performed on an XBridge C18 column (250 mm × 19 mm i.d., 5 µm; Waters) equipped with a C18 pre-column cartridge holder (10 mm × 19 mm i.d., 5 µm; Waters) at 17 mL/min, with H_2O (A) and MeCN (B) both containing 0.1% formic acid as solvents. The UV detection was set at 280 and 360 nm. The mixtures were injected into the semi-preparative HPLC column using a dry-load methodology developed in our laboratory (Queiroz et al., 2019). Masses injected and fractionation details are summarized in Supplementary Figure S5.

4.3.7 NMR conditions

A Bruker Avance Neo 600 MHz NMR spectrometer equipped with a QCI 5 mm Cryoprobe and a SampleJet automated sample changer (Bruker BioSpin, Rheinstetten, Germany) was employed for 1D-NMR (^{1}H and ^{13}C-NMR) and 2D-NMR [correlation spectroscopy (COSY), multiplicity editing heteronuclear single-quantum correlation (edited-HSQC), HMBC, ROESY, and total correlation spectroscopy (TOCSY)] spectroscopy. Chemical shifts are reported in parts per million (δ) using the residual solvent signal at δ_H 7.26; δ_C 77.2 for $CDCl_3$, δ_H 3.31; δ_C 49.0 for CD_3OD and δ_H 2.50; and δ_C 39.5 for DMSO-d_6. Chemical shifts (*J*) are reported in Hz.

4.3.8 Standard compounds

Metabolite standards were purchased from Biopurify (Chengdu, China): alantolactone St_1, isoalantolactone St_2, parthenolide St_4, costunolide St_5, dehydrocostus lactone St_6, artemisinin St_7, cinnamolide St_8, tanshinone I St_10 and IV St_11, betulin St_12, betulinic acid St_13, and celastrol St_14. Telekin St_3, clerod-3,13(16),14-trien-17-oic acid St_9, and riccardin G St_15 were isolated by Prof. Nagashima and Prof. Asakawa from Japan liverworts.

4.3.9 Description of the isolated compounds

The NMR descriptions of known compounds (Ft_2, Ft_3, Ft_4, Ft_6, Pp_1, Pp_2, Pp_3, Pp_5, Pe_1, Pe_2, Pe_3, Pe_5, Ds_1, Ds_2, and Ds_3) are summarized in the Supplementary Material. All spectra are according to the literature. The NMR spectra of unknown compounds (Ft_1, Ft_5, Pp_4, Ds_4, and Ac_1) are summarized in the Supplementary Material (Supplementary Figures S7-S33).

Six compounds were obtained from *F. tamarisci* (the yields for each compound are the sum of the compound isolated from the two extracts) in DCM extract from 20 mg [Ft_1 (1 mg)] and in MeOH Extracts Ft_2 (2.8 mg), Ft_3 (6.1 mg), Ft_4 (6.2 mg), Ft_5 (3.9 mg), and Ft_6 (3 mg)].

Ft_1: oxo-frullanolide ^{1}H NMR ($CDCl_3$, 600 MHz) δ 1.28 (3H, s, H_3-14), 1.42 (1H, overlapped, H-9ax), 1.60 (1H, overlapped, H-9eq), 1.68 (1H, overlapped, H-8ax), 1.70 (1H, overlapped, H-1eq), 1.80 (1H, overlapped, H-8eq), 1.90 (3H, s, H-15), 1.97 (1H, td, *J* = 14.2, 13.7, 5.1 Hz, H-1ax), 2.52 (1H, dt, *J* = 12.7, 5.1 Hz, H-2eq), 2.73 (1H, m, H-2ax), 3.17 (1H, td, *J* = 7.2, 6.4 Hz, H-7), 5.37 (1H, d, *J* = 6.4 Hz, H-6), 5.71 (1H, s, H-13″), 6.30 (1H, s, H-13′); ^{13}C NMR ($CDCl_3$, 151 MHz) δ 11.1 (CH_3-15), 24.7 (CH_3-14), 25.4 (CH_2-8), 34.2 (CH_2-2), 34.5 (C-10), 36.5 (CH_2-9), 37.5 (CH_2-1), 40.6 (CH-7), 75.6 (CH-6), 122.6 (CH_2-13), 136.8 (C-4), 140.5 (C-11), 151.7 (C-5), 170.3 (C-12), 198.8 (C-3); HR-ESI/MS, see Supplementary Table S2.

Ft_2: oxy-frullanolide (Sangsopha et al., 2016) HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Ft_3: frullanolide (Chou and Liao, 2013) HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Ft_4: γ-cyclocostunolide (Kraut et al., 1994) HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Ft_5:tamariscolide ^{1}H NMR ($CDCl_3$, 600 MHz) δ 1.09 (3H, s, H_3-14), 1.10 (3H, s, H_3-14′), 1.12 (1H, overlapped, H-1′b), 1.30 (3H, s, H_3-15), 1.31 (3H, s, H_3-15′), 1.37 (1H, m, H1′-a), 1.41 (2H, m, H_2-9′), 1.47 (4H, m, H-2′b, H-2b, H-3′a, H-8b), 1.53 (1H, m, H-3′b), 1.59 (1H, m, H-2′a), 1.63 (2H, m, H-3b, H-8′b), 1.70 (1H, m, H-9b), 1.77 (1H, m, H-2a), 1.80 (1H, m, H-3a), 1.94 (1H, d, *J* = 10.5 Hz, H-5′), 1.98 (1H, m, H-8′a), 2.00 (1H, m, H-5), 2.05 (1H, d, *J* = 13.4 Hz, H-8a), 2.12 (1H, d, *J* = 10.4 Hz, H-1), 2.27 (1H, d, *J* = 12.2 Hz, H-9a), 2.64 (2H, brs, H-7, H-7′), 4.05 (2H, m, H-6, H-6′), 5.38 (1H, s, H-13′b), 5.43 (1H, s, H-13b), 6.08 (1H, s, H-13′a), 6.10 (1H, s, H-13a); ^{13}C NMR ($CDCl_3$, 151 MHz) δ 18.3 (CH_2-2′), 19.2 (CH_3-14), 21.6 (CH_3-14′), 22.5 (CH_2-8′), 22.7 (CH_2-8), 23.1 (CH_2-2), 24.1 (CH_3-15′), 24.7 (CH_3-15), 35.0 (CH_2-3′), 38.3 (C-10′), 40.8 (CH_2-3), 41.0 (CH_2-1′), 42.0 (CH_2-9), 42.5 (C-4′), 45.2 (CH_2-9′), 45.4 (C-10), 49.6 (CH-7), 51.4 (CH-7′), 53.4 (CH-5′), 55.8 (CH-1), 59.8 (CH-5), 71.8 (C-4), 82.2 (CH-6′), 82.4 (CH-6), 117.3 (CH_2-13′), 118.1 (CH_2-13), 138.7 (C-11), 139.7 (C-11′), 169.8 (C-12), 170.3 (C-12′); HR-ESI/MS, see Supplementary Table S2.

Ft_6:ursolic acid (Acebey-Castellon et al., 2011): HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Five compounds were obtained from *P. porelloides*: (the yields for each compound are the sum of the compound isolated from the two extracts) in MeOH extract from 100 mg: Pp_1 (9.2 mg), Pp_2 (3.4 mg), Pp_3 (3.6 mg), Pp_4 (3.6 mg), and Pp_5 (1.3 mg).

Pp_1: plagiochiline D (Asakawa et al., 1979): HR-ESI/MS see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Pp_2: plagiochiline *R*-15-yl octanoate (Toyota et al., 1994; Ramírez et al., 2017): HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Pp_3: plagiochiline *R*-15-yl dec-4-enoate (Toyota et al., 1994; Ramírez et al., 2017): HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Pp_4: plagiochiline *R*-15-yl hexanoate ^{1}H NMR (DMSO-d_6, 600 MHz) δ 0.85 (3H, t, *J* = 7.1 Hz, H_3-6′), 0.99 (1H, t, *J* = 10.0 Hz, H-6), 1.07 (1H, m, H-9α), 1.20 (1H, m, H-8α), 1.25 (4H, m, H_2-5′, H_2-4′), 1.30 (1H, m, H-7), 1.51 (2H, p, $J = 7.4$ Hz, H_2-3′), 1.62 (1H, dd, $J = 10.0$, 3.2 Hz, H-1), 1.99 (1H, m, H-9β), 2.00 (3H, s, H_3-14b), 2.00 (1H, overlapped, H-8β), 2.03 (3H, s, H_3-12b), 2.11 (3H, s, H_3-2b), 2.27 (1H, dd, $J = 9.7$, 3.6 Hz, H-5), 2.28 (2H, td, $J = 7.4$, 2.2 Hz, H_2-2′), 2.40 (2H, AB, H-11), 3.74 (1H, d, $J = 11.4$ Hz, H-15″), 4.05 (1H, d, $J = 11.4$ Hz, H-15′), 4.21 (2H, AB, H_2-14), 4.42 (1H, d, $J = 12.4$ Hz, H-12″), 4.56 (1H, dd, $J = 12.4$, 1.4 Hz, H-12′), 6.46 (1H, s, H-3), 6.67 (1H, d, $J = 10.2$ Hz, H-2); ^{13}C NMR (DMSO-d_6, 151 MHz) δ 13.8 (CH_3-6′), 20.6 (CH_3-14b), 20.7 (CH_2-8), 20.8 (CH_3-2b, CH_3-12b), 21.7 (CH_2-5′), 24.1 (CH_2-3′), 20.9 (CH_3-2b), 24.3 (CH-7), 25.7 (C-13), 27.5 (CH-6), 29.7 (CH-5), 30.5 (CH_2-4′), 33.4 (CH_2-2′), 33.5 (CH_2-9), 49.1 (CH-1), 51.2 (CH_2-11), 59.2 (C-10), 60.9 (CH_2-14), 62.3 (CH_2-12), 68.6 (CH_2-15), 91.1 (CH-2), 116.0 (C-4), 139.6 (CH-3), 169.2 (C-2a), 170.5 (C-12a), 170.6 (C-14a), 172.8 (C-1′); HR-ESI/MS, see Supplementary Table S2.

Pp_5: diplophyllin (Ohta et al., 1977): HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Four compounds were obtained from *P. epiphylla* (the yields for each compound are the sum of the compound isolated from the two extracts) in DCM extract from 60 mg: Pe_1 (15.2 mg), Pe_2 (4.8 mg), Pe_3 (3.3 mg), and Pp_4 (2.3 mg).

Pe_1: 10-hydroxyperrottetin E (Cullmann et al., 1997), HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Pe_2: perrottetin E (Cullmann et al., 1997), HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Pe_3: 10-hydroxy-11-methoxy-perrottetin E, HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Pe_4: 11-methoxy-perrottetin E, HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Four compounds were obtained from *D. scoparium* (the yields for each compound are the sum of the compound isolated from the two extracts) in DCM extract from 40 mg: Ds_1 (0.5 mg), Ds_2 (0.7 mg), Ds_3 (0.5 mg), and Ds_4 (0.5 mg).

Ds_1: apigenin 7-*O*-[2,4-di-*O*-(α-ʟ-rhamnopyranosyl)]-β-ᴅ-glucopyranoside (Becker et al., 1986), HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Ds_2: 7-[(*O*-6-deoxy-α-ʟ-mannopyranosyl-(1→2)-*O*-[6-deoxy-α-ʟ-mannopyranosyl-(1→4)]-β-ᴅ-glucopyranosyl)oxy]-5-hydroxy-2-(3-hydroxy-4-methoxyphenyl)-4*H*-1-benzopyran-4-one (Osterdahl, 1978), HR-ESI/MS, see Supplementary Table S2; ^{1}H and ^{13}C NMR data, see Supplementary Material S1.

Ds_3: tiliroside = kaempferol-3-β-ᴅ-(6-*O*-*trans*-*p*-coumaroyl) glucopyranoside (Tsukamoto et al., 2004), HR-ESI/MS: see Supplementary Table S2, ^{1}H and ^{13}C NMR data, see Supplementary Table S2.

Ds_4: 7-hydroxy-2,4,10-trimethoxyphenanthrene ^{1}H NMR (DMSO-d_6, 600 MHz) δ 3.89 (3H, s, $2OCH_3$), 4.02 (3H, s, $10OCH_3$), 4.04 (3H, s, $4OCH_3$), 6.87 (1H, d, $J = 2.6$ Hz, H-3), 6.93 (1H, dd, $J = 9.2$, 2.7 Hz, H-6), 7.06 (1H, s, H-9), 7.12 (1H, d, $J = 2.7$ Hz, H-8), 7.29 (1H, d, $J = 2.6$ Hz, H-1), 9.17 (1H, d, $J = 9.2$ Hz, H-5), 9.54 (1H, s, 7OH); ^{13}C NMR (DMSO-d_6, 151 MHz) δ 55.1 ($2OCH_3$), 55.5 ($10OCH_3$), 55.8 ($4OCH_3$), 95.0 (CH-1), 99.8 (CH-3), 103.4 (CH-9), 110.5 (CH-8), 114.3 (CH-6), 116.0 (C-12), 118.9 (C-13), 127.5 (C-11), 128.6 (CH-5), 133.7 (C-14), 152.2 (C-10), 154.9 (C-7), 157.1 (C-2), 158.7 (C-4); HR-ESI/MS, see Supplementary Table S2.

One compound was obtained from *A. curtipendula* (the yields for each compound are the sum of the compound isolated from the two extracts) in MeOH extract from 100 mg: Ac_01 (0.2 mg).

Ac_01: 6,7-dihydroxy-2-methoxyphenanthrene ^{1}H NMR (DMSO-d_6, 600 MHz) δ 3.88 (3H, s, $2OCH_3$), 7.18 (1H, s, H-8), 7.20 (1H, dd, $J = 9.0$, 2.8 Hz, H-3), 7.33 (1H, d, $J = 2.8$ Hz, H-1), 7.49 (1H, d, $J = 8.8$ Hz, H-10), 7.55 (1H, d, $J = 8.8$ Hz, H-9), 7.89 (1H, s, H-5), 8.33 (1H, d, $J = 9.0$ Hz, H-4); ^{1}H NMR (DMSO-d_6, 151 MHz) δ 55.1 ($2OCH_3$), 106.6 (CH-5), 108.3 (CH-1), 112.1 (CH-8), 116.5 (CH-3), 123.3 (CH-10), 123.7 (CH-4), 124.3 (C-13), 125.2 (C-14), 126.5 (CH-9), 131.9 (C-11), 145.9 (C-7), 147.0 (C-6), 156.9 (C-2); HR-ESI/MS, see Supplementary Table S2.

4.4 Statistical and diversity analyses

Statistical analyses were carried out using R 4.0.2 and the following additional packages: dplyr, ComplexHeatmap, and circlize. The statistical analyses, peak, and classification tables were normalized by species. Missing values were imputed with zeros. The data obtained are summarized in the PI and NI heatmaps (Figure 2). The color gradient from red to yellow in PI and dark blue to light blue in NI is expressed horizontally and corresponds to the abundance of compound classes by species. Species are divided into two phyla (liverworts and mosses) and ordered according to their taxonomic proximity. Species marked with an asterisk have been selected and isolated to a validated cluster in MN. The barplot shows the number of compounds annotated (in light gray) versus the number of compounds not annotated (in dark gray).

Data availability statement

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

Author contributions

AP: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. AM: Conceptualization, Supervision, Writing – review & editing. LM: Formal analysis, Writing – review & editing. EF: Writing – review & editing. L-MQ-G: Writing – review & editing. YA: Investigation, Writing – review & editing. MD-K: Writing – review & editing. J-LW: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The authors acknowledge the Collectivity of Corsica for the funding.

Acknowledgments

The authors thank Mr. Achille Pioli for his contribution to the collection and identification of plants and Mme El Ali Manal for the plant collection and English language revision. The authors are grateful to Prof. Fumihiro Nagashima for the Japanese samples and the isolation of the standards.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1470307/full#supplementary-material

References

Acebey-Castellon, I. L., Voutquenne-Nazabadioko, L., Doan Thi Mai, H., Roseau, N., Bouthagane, N., Muhammad, D., et al. (2011). Triterpenoid saponins from *symplocos lancifolia*. *J. Nat. Prod.* 74, 163–168. doi: 10.1021/np100502y

Allard, P.-M., Gaudry, A., Quirós-Guerrero, L.-M., Rutz, A., Dounoue-Kubo, M., Walker, T. W. N., et al. (2023). Open and reusable annotated mass spectrometry dataset of a chemodiverse collection of 1,600 plant extracts. *GigaScience* 12, giac124. doi: 10.1093/gigascience/giac124

Allard, P.-M., Péresse, T., Bisson, J., Gindro, K., Marcourt, L., Pham, V. C., et al. (2016). Integration of molecular networking and in-Silico MS/MS Fragmentation for natural products dereplication. *Anal. Chem.* 88, 3317–3323. doi: 10.1021/acs.analchem.5b04804

Asakawa, Y. (1982). "Chemical constituents of the hepaticae," in *Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products* (Springer, Vienna), 1–285. doi: 10.1007/978-3-7091-8677-0_1

Asakawa, Y. (1995). "Chemical constituents of the bryophytes," in *Progress in the Chemistry of Organic Natural Products* (Springer, Vienna), 1–562. doi: 10.1007/978-3-7091-6896-7_1

Asakawa, Y. (2012). Bio- and chemical diversity of bryophytes: chemical structures and bioactivity of scents, and related compounds. *Aroma Res.* 13, 270–278.

Asakawa, Y., and Ludwiczuk, A. (2013). "Bryophytes: Liverworts, Mosses, and Hornworts: extraction and isolation procedures," in *Metabolomics Tools for Natural Product Discovery* (Humana Press, Totowa, NJ), 1–20. doi: 10.1007/978-1-62703-577-4_1

Asakawa, Y., Ludwiczuk, A., and Nagashima, F. (2013a). Phytochemical and biological studies of bryophytes. *Phytochemistry* 91, 52–80. doi: 10.1016/j.phytochem.2012.04.012

Asakawa, Y., Ludwiczuk, D. A., and Nagashima, D. F. (2013b). "Chemical diversity of bryophytes," in *Chemical Constituents of Bryophytes* (Springer, Vienna), 21–24. doi: 10.1007/978-3-7091-1084-3_3

Asakawa, Y., Ludwiczuk, A., Novakovic, M., Bukvicki, D., and Anchang, K. Y. (2021). Bis-bibenzyls, bibenzyls, and terpenoids in 33 genera of the marchantiophyta (Liverworts): structures, synthesis, and bioactivity. *J. Nat. Prod.* 85 (3), 729–762. doi: 10.1021/acs.jnatprod.1c00302

Asakawa, Y., Matsuda, R., Toyota, M., Hattori, S., and Ourisson, G. (1981). Terpenoids and bibenzyls of 25 liverwort *Frullania* species. *Phytochemistry* 20, 2187–2194. doi: 10.1016/0031-9422(81)80111-2

Asakawa, Y., Toyota, M., Takemoto, T., and Suire, C. (1979). lagiochilines C, D, E and F, four novel secoaromadendrane-type sesquiterpene hemiacetals from *Plagiochila asplenioides* and *Plagiochila semidecurrens*. *Phytochemistry* 18, 1355–1357. doi: 10.1016/0031-9422(79)83021-6

Augier, J. (1966). *Flore des bryophytes* (Paris: P. Lechevalier).

Becker, R., Mues, R., Zinsmeister, H. D., Herzog, F., and Geiger, H. (1986). A new biflavone and further flavonoids from the moss *Hylocomium splendens*. *Z. für Naturforschung C* 41, 507–510. doi: 10.1515/znc-1986-5-602

Chou, Y., and Liao, C. (2013). First asymmetric total syntheses and determination of absolute configurations of (+)-Eudesmadiene-12,6-olide and (+)-Frullanolide. *Organic letter* 15, 1584–1587. doi: 10.1021/ol4003724

Christenhusz, M. J. M., and Byng, J. W. (2016). The number of known plants species in the world and its annual increase. *Phytotaxa* 261, 201–217. doi: 10.11646/phytotaxa.261.3.1

Cullmann, F., Becker, H., Pandolfi, E., Roeckner, E., and Eicher, T. (1997). Bibenzyl derivatives from *Pellia epiphylla*. *Phytochemistry* 45, 1235–1247. doi: 10.1016/S0031-9422(97)00118-0

Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., et al. (2019). SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* 16, 299–302. doi: 10.1038/s41592-019-0344-8

Dührkop, K., Nothias, L.-F., Fleischauer, M., Reher, R., Ludwig, M., Hoffmann, M. A., et al. (2021). Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nat. Biotechnol.* 39, 462–471. doi: 10.1038/s41587-020-0740-8

Dührkop, K., Shen, H., Meusel, M., Rousu, J., and Böcker, S. (2015). Searching molecular structure databases with tandem mass spectra using CSI : FingerID. *Proc. Natl. Acad. Sci.* 112, 12580–12585. doi: 10.1073/pnas.1509788112

Fan, G.-X., Dong, L.-L., Li, H.-H., Li, Z.-Y., Zhang, Z.-X., and Fei, D.-Q. (2016). Sesquiterpenoids and other chemical components from the roots of *dolomiaea souliei*. *Chem. Nat. Compd* 52, 754–757. doi: 10.1007/s10600-016-1766-5

Guillarme, D., Nguyen, D. T. T., Rudaz, S., and Veuthey, J.-L. (2008). Method transfer for fast liquid chromatography in pharmaceutical analysis: Application to short columns packed with small particle. Part II: Gradient experiments. *Eur. J. Pharmaceutics Biopharmaceutics* 68, 430–440. doi: 10.1016/j.ejpb.2007.06.018

Hoffmann, M. A., Nothias, L.-F., Ludwig, M., Fleischauer, M., Gentry, E. C., Witting, M., et al. (2022). High-confidence structural annotation of metabolites absent from spectral libraries. *Nat. Biotechnol.* 40, 411–421. doi: 10.1038/s41587-021-01045-9

Horn, A., Pascal, A., Lončarević, I., Volpatto Marques, R., Lu, Y., Miguel, S., et al. (2021). Natural products from bryophytes: from basic biology to biotechnological applications. *Crit. Rev. Plant Sci.* 40, 191–217. doi: 10.1080/07352689.2021.1911034

Kraut, L., Mues, R., and Sim-Sim, M. (1994). Sesquiterpene lactones and 3-benzylphthalides from *Frullania muscicola*. *Phytochemistry* 37, 1337–1346. doi: 10.1016/S0031-9422(00)90409-6

Lu, Y., Eiriksson, F. F., Thorsteinsdottir, M., and Simonsen, H. T. (2019). Valuable fatty acids in bryophytes—production, biosynthesis, analysis and applications. *Plants* 8, 524. doi: 10.3390/plants8110524

Ludwig, M., Nothias, L.-F., Dührkop, K., Koester, I., Fleischauer, M., Hoffmann, M. A., et al. (2020). Database-independent molecular formula annotation using Gibbs sampling through ZODIAC. *Nat. Mach. Intell.* 2, 629–641. doi: 10.1038/s42256-020-00234-6

Martini, U., Zapp, J., and Becker, H. (1998). Chlorinated macrocyclic bisbibenzyls from the liverwort *Bazzania trilobata*. *Phytochemistry* 47, 89–96. doi: 10.1016/S0031-9422(97)00495-0

Nadgouda, S. A., Trivedi, G. K., and Bhattacharyya, S. C. (1978). Sensitized photo oxygenation of alpha cyclo costunolide and di hydro alpha cyclo costunolide a biogenetic type transformation of costunolide to santonin. *Indian J. Chem. Sect. B* 16, 16–19.

Nagashima, F., Tanaka, H., Toyota, M., Hashimoto, T., Kan, Y., Takaoka, S., et al. (1994). Sesqui- and diterpenoids from *Plagiochila* species. *Phytochemistry* 36, 1425–1430. doi: 10.1016/S0031-9422(00)89735-6

Nothias, L.-F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., et al. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* 17, 905–908. doi: 10.1038/s41592-020-0933-6

Novaković, M., Ludwiczuk, A., Bukvički, D., and Asakawa, Y. (2021). Phytochemicals from bryophytes: Structures and biological activity. *J. Serbian Chem. Soc.* 86, 1139–1175. doi: 10.2298/JSC211027100N

Ohta, Y., Andersen, N. H., and Liu, C.-B. (1977). Sesquiterpene constituents of two liverworts of genus *diplophyllum*: Novel eudesmanolides and cytotoxicity studies for enantiomeric methylene lactones. *Tetrahedron* 33, 617–628. doi: 10.1016/0040-4020(77)80301-3

Osterdahl, B. (1978). Chemical studies on bryophytes.20. New branched flavonoid-omicron-triglycoside from *dicranum-scoparium*. *Acta Chemica Scandinavica Ser. B-Organic Chem. Biochem.* 32, 714–716. doi: 10.3891/acta.chem.scand.32b-0714

Pannequin, A. (2019). *Caractérisation chimique des bryophytes de Corse et propriétés biologiques* (Université Pascal Paoli). Available online at: https://tel.archives-ouvertes.fr/tel-03482569 (Accessed March 10, 2022).

Pannequin, A., Laurini, E., Giordano, L., Muselli, A., Pricl, S., and Tintaru, A. (2020). Caution: chemical instability of natural biomolecules during routine analysis. *Molecules* 25, 3292. doi: 10.3390/molecules25143292

Pannequin, A., Quetin-Leclercq, J., Costa, J., Tintaru, A., and Muselli, A. (2023). First phytochemical profiling and *in-vitro* antiprotozoal activity of essential oil and extract of *plagiochila porelloides*. *Molecules* 28, 616. doi: 10.3390/molecules28020616

Pannequin, A., Tintaru, A., Desjobert, J.-M., Costa, J., and Muselli, A. (2017). New advances in the volatile metabolites of *Frullania tamarisci*. *Flavour Fragr J.* 32:409–418. doi: 10.1002/ffj.3407

Peters, K., Balcke, G., Kleinenkuhnen, N., Treutler, H., and Neumann, S. (2021). Untargeted in silico compound classification—A novel metabolomics method to assess the chemodiversity in bryophytes. *Int. J. Mol. Sci.* 22, 3251. doi: 10.3390/ijms22063251

Peters, K., Treutler, H., Döll, S., Kindt, A. S. D., Hankemeier, T., and Neumann, S. (2019). Chemical diversity and classification of secondary metabolites in nine bryophyte species. *Metabolites* 9, 222. doi: 10.3390/metabo9100222

Queiroz, E. F., Alfattani, A., Afzan, A., Marcourt, L., Guillarme, D., and Wolfender, J.-L. (2019). Utility of dry load injection for an efficient natural products isolation at the semi-preparative chromatographic scale. *J. Chromatogr. A* 1598, 85–91. doi: 10.1016/j.chroma.2019.03.042

Ramírez, M., Kamiya, N., Popich, S., Asakawa, Y., and Bardón, A. (2017). Constituents of the Argentine Liverwort *Plagiochila diversifolia* and their insecticidal activities. *Chem. Biodiversity* 14, e1700229. doi: 10.1002/cbdv.201700229

Rutz, A., Dounoue-Kubo, M., Ollivier, S., Bisson, J., Bagheri, M., Saesong, T., et al. (2019). Taxonomically informed scoring enhances confidence in natural products annotation. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01329

Rutz, A., Sorokina, M., Galgonek, J., Mietchen, D., Willighagen, E., Gaudry, A., et al. (2022). The LOTUS initiative for open knowledge management in natural products research. *eLife* 11, e70780. doi: 10.7554/eLife.70780

Sangsopha, W., Lekphrom, R., Kanokmedhakul, S., and Kanokmedhakul, K. (2016). Cytotoxic and antimalarial constituents from aerial parts of *Sphaeranthus indicus*. *Phytochem. Lett.* 17, 278–281. doi: 10.1016/j.phytol.2016.08.001

Sartori, S. K., Diaz, M. A. N., and Diaz-Muñoz, G. (2021). Lactones: Classification, synthesis, biological activities, and industrial applications. *Tetrahedron* 84, 132001. doi: 10.1016/j.tet.2021.132001

Scher, J. M., Zapp, J., Schmidt, A., and Becker, H. (2003). Bazzanins L–R, chlorinated macrocyclic bisbibenzyls from the liverwort *Lepidozia incurvata*. *Phytochemistry* 64, 791–796. doi: 10.1016/S0031-9422(03)00382-0

Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrlund, T. S., et al. (2023). Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* 41, 447–449. doi: 10.1038/s41587-023-01690-2

Shen, S., Zhan, C., Yang, C., Fernie, A. R., and Luo, J. (2023). Metabolomics-centered mining of plant metabolic diversity and function: Past decade and future perspectives. *Mol. Plant* 16, 43–63. doi: 10.1016/j.molp.2022.09.007

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432. doi: 10.1093/bioinformatics/btq675

Sotiaux, A., Pioli, A., Royaud, A., Schumacker, R., and Vanderpoorten, A. (2007). A checklist of the bryophytes of Corsica (France): new records and a review of the literature. *J. Bryology* 29, 41–53. doi: 10.1179/174328207X171872

Sotiaux, A., Sotiaux, O., and Vanderpoorten, A. (2008). Additions to the bryophyte flora of Corsica. *Cryoptogamie Bryologie* 29, 267–274.

Spörle, J., Becker, H., Allen, N. S., and Gupta, M. P. (1991). Spiroterpenoids from plagiochila moritziana. *Phytochemistry* 30, 3043–3047. doi: 10.1016/S0031-9422(00)98249-9

Toyota, M., Nakamura, I., Huneck, S., and Asakawa, Y. (1994). Sesquiterpene esters from the liverwort *Plagiochila porelloides*. *Phytochemistry* 37, 1091–1093. doi: 10.1016/S0031-9422(00)89535-7

Toyota, M., Nishimoto, C., and Asakawa, Y. (1998). Eudesmane-type sesquiterpenoids from Japanese liverwort. *Frullania tamarisci subsp. obscura.* 46, 542–544.

Tsukamoto, S., Tomise, K., Aburatani, M., Onuki, H., Hirorta, H., Ishiharajima, E., et al. (2004). Isolation of cytochrome P450 inhibitors from strawberry fruit, *Fragaria ananassa*. *J. Nat. Prod* 67, 1839–1841. doi: 10.1021/np0400104

Vergoten, G., and Bailly, C. (2023). The Plagiochilins from *Plagiochila* Liverworts: binding to α-Tubulin and drug design perspectives. *AppliedChem* 3, 217–228. doi: 10.3390/appliedchem3020014

von Reuß, S. H., and König, W. A. (2004). Corsifurans A–C, 2-arylbenzofurans of presumed stilbenoid origin from *Corsinia coriandrina* (Hepaticae). *Phytochemistry* 65, 3113–3118. doi: 10.1016/j.phytochem.2004.10.002

von Reuß, S. H., and König, W. A. (2005). Olefinic isothiocyanates and iminodithiocarbonates from the liverwort corsinia coriandrina. *Eur. J. Organic Chem.* 2005, 1184–1188. doi: 10.1002/ejoc.200400586

Wolfender, J.-L., Nuzillard, J.-M., van der Hooft, J. J. J., Renault, J.-H., and Bertrand, S. (2019). Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography–high-resolution tandem mass spectrometry and NMR profiling, in silico databases, and chemometrics. *Anal. Chem.* 91, 704–742. doi: 10.1021/acs.analchem.8b05112

Zhang, Z., Li, Y., Sun, Y., Wang, W., Song, X., and Zhang, D. (2023). Chemical diversity and biological activities of marine-derived sulphur containing alkaloids: A comprehensive update. *Arabian J. Chem.* 16, 105011. doi: 10.1016/j.arabjc.2023.105011

Frontiers in Plant Science

TYPE Original Research
PUBLISHED 21 January 2025
DOI 10.3389/fpls.2024.1514614

Check for updates

OPEN ACCESS

EDITED BY
Lei Zhang,
Jiangsu Normal University, China

REVIEWED BY
Simon Scofield,
Cardiff University, United Kingdom
Zhengqing Xie,
Zhengzhou University, China

*CORRESPONDENCE
Zongxia Yu
yuzx@lsbg.cn
Lei Yang
leiyang@cemps.ac.cn

RECEIVED 21 October 2024
ACCEPTED 24 December 2024
PUBLISHED 21 January 2025

CITATION
Yu Z, Lv R, Hong B and Yang L (2025) Integrating cotyledon-based virus-induced gene silencing with visual marker promises a rapid, highly effective validation of gene functions in *Nepeta cataria*.
Front. Plant Sci. 15:1514614.
doi: 10.3389/fpls.2024.1514614

Integrating cotyledon-based virus-induced gene silencing with visual marker promises a rapid, highly effective validation of gene functions in *Nepeta cataria*

Zongxia Yu[1,2]*, Ruo Lv[1], Bo Hong[1] and Lei Yang[2]*

[1]Jiangxi Key Laboratory for Sustainable Utilization of Chinese Materia Medica Resources, Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang, Jiangxi, China, [2]Shanghai Key Laboratory of Plant Functional Genomics and Resources, Shanghai Chenshan Botanical Garden, Shanghai, China

Nepeta spp. generate volatile nepetalactone iridoids that have cat-attractant and insect-repellent activities. They differ from typical mint family (Lamiaceae) iridoids, which are non-volatile glucosides, and also vary from other species in the Nepetoideae sub-family, which do not generate iridoids. The chemistry and evolution of *Nepeta* make it suitable for further investigation. However, the lack of transgenic technology hampers the molecular and genetic investigations in *Nepeta*. Virus-induced gene silencing (VIGS) is a powerful tool to detect gene functions *in vivo*. Here, we constructed a modified VIGS method in *Nepeta cataria*, using cotyledon infiltration, with the gene silencing effect spreading to the first two pairs of true leaves. The VIGS efficiency reached as high as 84.4%, and the procedure takes only 3 weeks. We employed this method to validate the role of geraniol 8-hydroxylase in nepetalactone biosynthesis with *ChlH* as a visual marker in *N. cataria*. The method is also applicable to *Nepeta mussinii*. Thus, we developed an easy and effective VIGS approach, which will be advantageous for endogenous gene studies in two *Nepeta* species and holds the potential for application in other plants.

KEYWORDS

VIGS, cotyledon infiltration, visual marker, G8H, nepetalactone, catmint

Introduction

Virus-induced gene silencing (VIGS) is an easy and quick tool for genetic and functional validations of genes in non-model plants. VIGS makes use of the plant's natural antiviral immune system to knock down endogenous gene levels by posttranscriptional gene silencing (PTGS) or transcriptional gene silencing (TGS). When plant viruses invade plants, they replicate and transcribe viral RNAs, which are processed into short interfering RNA (siRNA) by plant DICER-LIKE (DCL) and HUA ENHANCER 1 (HEN1) proteins. Then, siRNAs are equipped with ARGONAUTE (AGO) nucleases and other particles of RNA-induced silencing complex (RISC) to degrade the complementary gene transcripts or methylate the target gene (Matzke and Mosher, 2014; Carbonell and Carrington, 2015; Taochy et al., 2017). Beyond gene silencing, the derived toolbox of VIGS has broadly expanded to encompass virus-induced overexpression (VOX), virus-induced genome editing (VIGE), and host-induced gene silencing (HIGS) (Cheuk and Houde, 2019; Jiang et al., 2021; Lei et al., 2021), expanding the applications from clarifying gene functions to revealing molecular mechanisms, protecting crops from biological and abiotic stresses (Wu et al., 2020; Yao et al., 2020).

The viruses used for VIGS are mostly single-stranded DNA viruses like the African cassava mosaic virus (ACMV) or RNA viruses like the foxtail mosaic virus for monocots and tobacco rattle virus (TRV) for dicots (Wu et al., 2011; Liu et al., 2016; Rössner et al., 2022). TRV is commonly used, as it has a wide host range and the ability to infect meristematic tissue (Li et al., 2018). The genome of TRV has been disassembled and constructed into two vectors, named TRV1 and TRV2. TRV1 comprises replicases, a transmission protein, and a cysteine-rich protein, while TRV2 contains a coat protein and a multiple cloning site for the insertion of the target gene's fragment (Shi et al., 2021). The fragment is selected from the coding sequence of the target gene with a length of 200–400 bp. The viral genome harboring the target gene's fragment is then delivered to the plants mediated by *Agrobacterium tumefaciens* in diverse ways such as syringe injection, vacuum infiltration, spraying, and rubbing.

Nepeta spp. generate iridoids, a type of monoterpenoid. These are widespread in the Lamiaceae family but exclusive to the *Nepeta* genus within the Nepetoideae sub-family. This characteristic renders *Nepeta* a model genus for investigating iridoid evolution within the Lamiaceae family (Boachon et al., 2018). Nepetalactone represents the major iridoid of *Nepeta* spp. and is well-known for attracting cats; this is why *Nepeta* spp. are also called catmint or catnip (Uenoyama et al., 2021). The biosynthesis of nepetalactone starts with geranyl pyrophosphate (GPP), which is converted into geraniol, 8-hydroxygeraniol, and 8-oxogeranial under the catalysis of geraniol synthase (GES), geraniol 8-hydroxylase (G8H), and 8-hydroxygeraniol oxidoreductase (HGO), respectively. Then, iridoid synthase (ISY) reduces 8-oxogeranial to form nepetalactol, the precursor to iridoids in plants. Nepetalactol-related short-chain dehydrogenases (NEPSs) and a major latex protein-like (MLPL) enzyme can contribute to the stereoselectivity of cyclization during the conversion of 8-oxogeranial to nepetalactol (Lichman et al., 2019; Lichman et al., 2020). Then, NEPSs convert nepetalactol isomers to nepetalactone (Lichman et al., 2020; Smit and Lichman, 2022). However, the lack of a mutant pool and transgenic tools in *Nepeta* spp. impedes the verification of the gene function *in planta*.

VIGS is an alternative for elucidating gene function in non-model plants recalcitrant to stable transformation including the Solanaceae, Cruciferae, Malvaceae, and Gramineae families (Shi et al., 2021). Within Lamiaceae, it has been successfully utilized to resolve pentacyclic triterpene biosynthesis in *Ocimum basilicum* (Misra et al., 2017). Palmer et al. successfully applied VIGS in *Nepeta cataria* by inoculating the nodes in *N. cataria* cuttings together with visual markers and thereby validated the functions of *GES*, *ISY*, and *MLPL* (Palmer et al., 2022). Herein, we introduce a more rapid and effective VIGS method applicable to *N. cataria* and *Nepeta mussinii* using cotyledon infiltration alongside a visual marker. Through this approach, we provide *in vivo* validation of *G8H* participating in nepetalactone biosynthesis.

Methods

Plant growth conditions

The seeds of *N. cataria* and *N. mussinii* were bought from SeedCorner Kings Seeds (https://www.kingsseedsdirect.com/). The seeds were sowed in 1-cm depth under the compost by toothpicks and cultured under 16/8-h light and 25°C/22°C day/night regime. For material collection, the chlorotic area of the infected leaves from three plants was cut and collected as one sample and immediately frozen in liquid nitrogen (LN2); at least three biological samples were prepared for each vector. The sample was ground into powder by TissueLyser II (QIAGEN, Valencia, CA, USA), separated, and weighed for qRT-PCR and gas chromatography–mass spectrometry (GC–MS) analysis.

Vector construction

*ChlH*s from *N. cataria* and *N. mussinii* were picked by alignment with reported *ChlH*s from other plants (Figure 1). The amino acid sequences of *Vitis vinifera* (VvChlH, NP_001268078.1), *Glycine max* (GmChlH, AXB26707.1), *Prunus persica* (PpChlH, ACO57443.1), and *Fragaria* × *ananassa* (FaChlH, AEN74910.1) were used as queries to search the homologous genes from the database of catmint (https://datadryad.org/stash/dataset/doi:10.5061/dryad.88tj450). The *ChlH* (329 bp) fragment was designed from the conserved region of *NcChlH* and *NmChlH* in order to work across both species with one vector, as the identities of *ChlH*s were high. *GES* (366 bp) and *G8H* (288 bp) were designed from the specific region of the coding sequences (CDSs) of the genes (Supplementary Data Sheet S1). All fragments were searched throughout the genomes of *N. cataria* and *N. mussinii* to avoid off-target effects and then amplified with overhangs homologous to the TRV2 vector using Phusion Plus DNA Polymerase (Thermo Fisher, Waltham, MA, USA; F630S). The *ChlH* fragment was inserted into TRV2 between *Eco*RI and *Bam*HI, and *GES* and

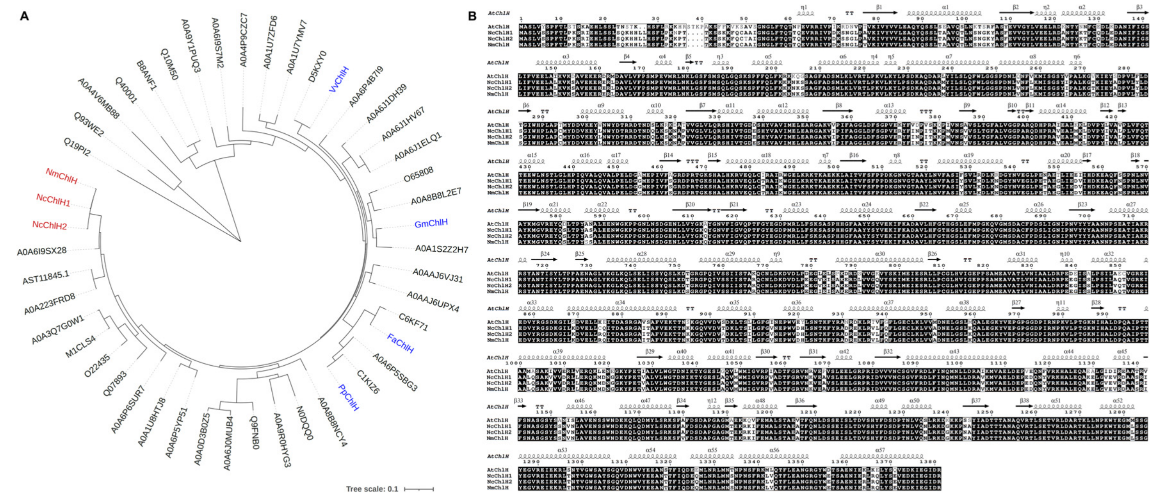


FIGURE 1

Phylogenetic analysis of *Nepeta* ChlHs. **(A)** Phylogenetic tree of *Nepeta* ChlHs with their homologs. The homologs (Supplementary Data Sheet S3) were identified from the UniProt database (https://www.uniprot.org/) using VvChlH, GmChlH, PpChlH, and FaChlH protein sequences as the Basic Local Alignment Search Tool-protein (BLASTP) queries, which were marked in blue color. The protein sequences of NcChlH1, NcChlH2, and NmChlH1 were directly obtained from *Nepeta cataria* and *Nepeta mussinii* transcriptome databases and were marked in red color. The tree was constructed using the maximum likelihood algorithm in TBtools (Chen et al., 2023). Branch point bootstrap values were calculated with 1,000 replicates. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. **(B)** Alignment of the protein sequences of *Nepeta* ChlHs with AtChlH. *Nepeta* ChlH sequences were aligned with the amino acid sequence and secondary structure of AtChlH (Q9FNB0) on the ESPript 3.0 website (https://espript.ibcp.fr/ESPript/ESPript/).

G8H were cloned between *Xho*I and *Sma*I using homologous recombination (Vazyme, Nanjing, China; C112-00).

RNA extraction, reverse transcription, and qRT-PCR analysis

Powdered tissue (~100 mg) was used for RNA extraction according to the instructions of the plant RNA purification kit (Vazyme, RC401-01). Total RNA (~1 µg) was used for reverse transcription to generate cDNA following the instructions of M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA; 28025-013). The primers for qRT-PCR analysis were designed behind the VIGS fragment. The cDNAs were diluted 10 times before performing the qRT-PCR using SYBR Green qPCR Master Mix (Selleckchem, Houston, TX, USA; B21202) on Bio-Rad CFX96. The method of $2^{-\Delta\Delta Ct}$ was used to measure gene expression level with ubiquitin as the internal standard. The primers used for vector construction and qRT-PCR analysis are listed in Supplementary Data Sheet S2.

Cotyledon-based virus-induced gene silencing in *N. cataria* and *N. mussinii*

The TRV bipartite system was selected for VIGS, which includes two vectors. TRV1 ensures the replication and movement of viral functions, and TRV2 encodes the coat protein and the fragment of the target gene for VIGS. TRV1 and TRV2 were transformed into *A. tumefaciens* GV3101 using the freeze–thaw method separately. The positive clones were inoculated in 1 mL Luria-Bertani (LB) medium containing kanamycin and gentamycin (50 mg/L) antibiotics and cultured overnight at 28°C, shaking at 220 rpm; 200-µL overnight cultures were inoculated in 10 mL LB with antibiotics, 10 mM MES, and 20 µM acetosyringone and shaken overnight at 28°C and 220 rpm. The cultures were centrifuged at 3,500 ×*g* for 15 min and resuspended in 2 mL fresh prepared infiltration buffer (10 mM MES, pH 5.8, 10 mM $MgCl_2$, and 200 µM acetosyringone). The OD_{600} was adjusted to 2.5, and TRV1 with TRV2 were mixed at the volume ratio of 1:1. The mixture was shaken for 3 h at 28°C and 100 rpm in the dark before being infiltrated into the cotyledons of seedlings 7 days after sowing using a 1-mL syringe. Fifteen to 40 seedlings were infiltrated for each group. The chlorotic phenotype appeared as early as 6 days after infiltration. The chlorotic parts from the first and second true leaves were collected 2 weeks post infiltration, or the whole first and second true leaves were collected when visual marker was lacking in the control, flash frozen in LN2, and stored in −80°C freezer for further qRT-PCR and GC–MS analysis.

Metabolites extraction and GC–MS analysis

Tissue powder (~50 mg) was weighed and transferred into a 2.0-mL tube, and 600 µL MeOH was added, vortexed for 10 s, and shaken for 10 min; 600 µL of hexane containing 30 ng/µL nonyl acetate as internal standard was added, vortexed, and shaken for 20 min. The mix was centrifuged at room temperature and top speed (14,600 ×*g*) for 10 min; 400 µL of the upper hexane layer was transferred into a new 1.5-mL tube and centrifuged again, and the supernatant was collected for GC–MS analysis.

For GC–MS analysis, the samples were injected in split mode (2 µL, split ratio 5:1) at an inlet temperature of 220°C. The column of Zebron ZB-5HT-INFERNO (30 m × 250 µm × 0.1 µm) was used for separation. Helium was used as carrier gas with a flow rate of 1.2 mL/min. The program was as follows: hold at 80°C for 5 min, increase to 110°C at a rate of 2.5°C/min, enhance to 280°C at 50°C/min, and hold for 4 min. Nonyl acetate (30 ng/µL, Sigma-Aldrich Corp., St. Louis, MO, USA) was used as the internal standard, and the contents of metabolites were calculated by comparing them to the internal standard.

Results

Cotyledon-based VIGS with rapid process and high efficiency stands out from other infection approaches in *N. cataria*

Magnesium chelatase subunit H (ChlH) participating in chlorophyll biosynthesis was chosen as the visual marker, as it will cause the chlorotic phenotypes to appear when knocked down by VIGS (Palmer et al., 2022). Four reported ChlHs (VvChlH, GmChlH, PpChlH, and FaChlH) were employed as queries to identify the native ChlHs from the databases of both *N. cataria* and *N. mussinii*, which were designated as NcChlH1, NcChlH2, and NmChlH1. *Nepeta* ChlHs exhibit high amino acid sequence identities with their homologs and share the highest homology with *Sesamum indicum* (Figure 1), suggesting that ChlH genes are highly conserved throughout evolution. The transcription levels of these *Nepeta ChlHs* are the most abundant in immature leaves, followed by mature leaves and closed flower buds (Supplementary Figure S1), which is in accordance with their roles in chlorophyll biosynthesis.

Three convenient VIGS infection methods (independent of vacuum devices) were compared in *N. cataria*: 1) cotyledon-based, which involved the cotyledons of the 7-day-old seedlings being infiltrated with the VIGS infection solution using a syringe; 2) wounded nodes, which was the approach described by Palmer et al (Palmer et al., 2022), wherein the leaves under the shoot of 1-month-old plants were removed from each node, both the nodes and the shoot apical meristem were punctured by toothpick and soaked with infection solution; 3) true-leaf VIGS, wherein the abaxial side of the leaves of 2-week-old plants was infiltrated with a syringe as described in many other plants. The chlorotic phenotype manifested in both cotyledon-based and wounded nodes approaches and could be observed as early as 1 week after VIGS (WAV) in the newly emerging leaves instead of the infiltrated ones (Figure 2). However, true-leaf VIGS did not work in the trials in *N. cataria*, considering that the second approach needed approximately 1 month for plant growth to provide sufficient nodes (≥3 nodes) and performed lower VIGS efficiency in our trials. Consequently, the quicker cotyledon-based VIGS was selected for further investigation.

The chlorotic phenomenon spread only to the first and second pairs of true leaves in both *N. cataria* and *N. mussinii*

The spread extent of chlorosis varies significantly in different plant species, ranging from a single pair of leaves to the whole plant. The VIGS plants were constantly monitored to determine the optimal time for collecting VIGS materials. After infiltration of the cotyledons, the proliferation of chlorosis was only observed in the first and second pairs of true leaves in *N. cataria* (Nc), but not the subsequently emerging leaves during three to four WAVs (Figures 3A–D). A closely related species, *N. mussinii* (Nm), was also investigated to see whether cotyledon-based VIGS is applicable to other plant species. Similar results were gained: chlorosis occurred in the first pair of true leaves in one WAV and then extended to the second pair in two WAVs but ceased to expand in three to four WAVs (Figures 3E–H). These results illustrated that the VIGS-induced repression of gene expression could propagate from cotyledon to the first and second pairs of true leaves without any further expansion in *N. cataria* and *N. mussinii*. The entire process of cotyledon-based VIGS could be completed within 3 weeks from seed germination to sample collection. Additionally, the cotyledon-based VIGS functioned effectively in both *N. cataria* and *N. mussinii*, indicating its potential application in a diverse range of plant species.

The VIGS efficiency in *N. cataria* was higher than in *N. mussinii*

To detect the efficiency of the cotyledon-based VIGS, the experiments were conducted with different infiltration times in *N. cataria* and *N. mussinii*. In *N. cataria*, when the infiltration was performed in the cotyledons of seedlings 7 days after sowing (DAS), the mean VIGS efficiency was 84.4% by comparing the chlorotic ones divided by the total infiltration amount. However, when the infiltration was performed on nine DAS, the VIGS efficiency decreased. In comparison, the mean efficiency was 47.7% in seven DAS batches and 36.7% in nine DAS batches in *N. mussinii*, which was lower than that of *N. cataria* (Table 1). Therefore, the VIGS efficiency was diverse among different plant species, even when they shared a close phylogenetic relationship. Seven DAS was recommended as a better infiltration time since the VIGS efficiency showed a decreasing trend in nine DAS batches.

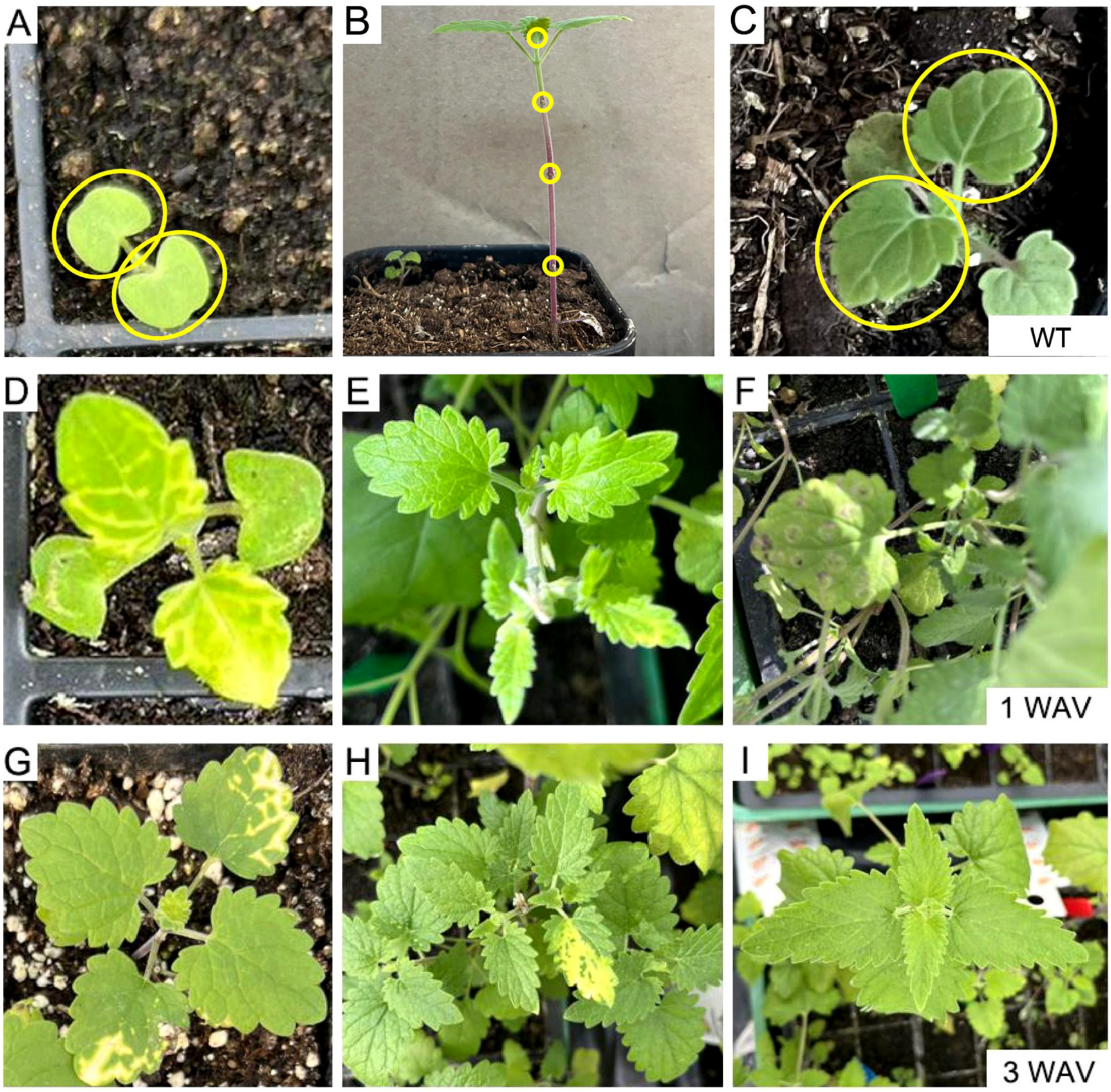


FIGURE 2
Comparison of different virus-induced gene silencing (VIGS) approaches in *Nepeta cataria*. **(A, D, G)** Cotyledon-based VIGS was used. **(B, E, H)** The wounded node method. **(C, F, I)** The true-leaf VIGS. **(A–C)** The wild type (WT) before treatment; the infiltration parts are circled in yellow. **(D–F)** One week after VIGS. **(G–I)** Three weeks later.

The visual marker *ChlH* aided in gene function validation

The area and time period of gene silencing of VIGS exhibit variability among species and methods. Take *ChlH* for example; chlorosis can appear in the different phyllotaxes or leaf parts (Figure 4B). A marker can be useful to guide sampling for the identification of gene function. Previously, *ChlH* was independently employed as a positive control to validate the VIGS function well, and sampling positions of target genes were roughly approximated based on the positive control (Figure 4A). In this study, fragments of the target genes and the visual marker *ChlH* were successively co-constructed into TRV2 (Figure 4B). The silencing of the target gene and *ChlH* would be expected to happen in the same area, thus enabling the use of chlorosis of *ChlH* as a guiding indicator. GES and G8H are two key enzymes in nepetalactone biosynthesis. The role of *GES* has been characterized both *in vitro* and *in vivo* and was chosen as a positive control.

To assess the impact of the visual marker on elucidating gene function, the entire leaves of the first and second true leaves were sampled following the method of a previous study (Figures 4A, 5A, C). In contrast, only the chlorotic regions were harvested in the visual-marker batch (Figures 4B, 5B, D). Although the expression level of *GES* and the content of *trans–cis* nepetalactone slightly decreased in comparison to the control, no significant difference was observed (Figures 5A, C). Conversely, with the assistance of *ChlH* during

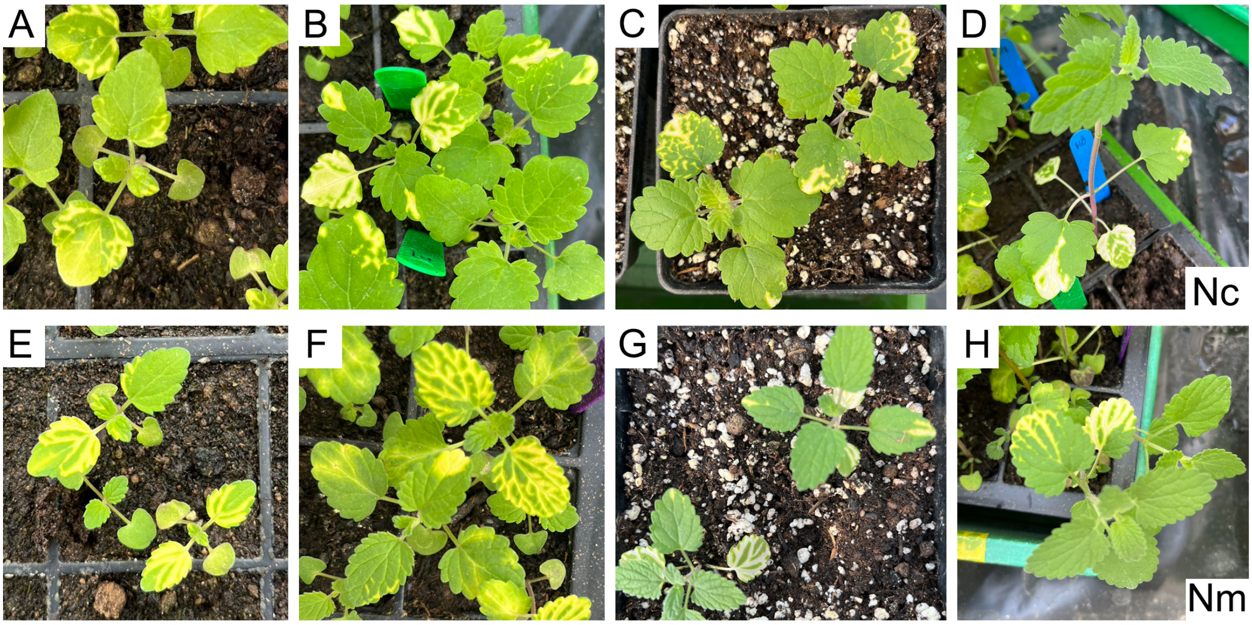


FIGURE 3
Evaluation of the spreading range and period. **(A–D)** *Nepeta cataria* (Nc). **(E–H)** *Nepeta mussinii* (Nm). **(A–H)** represent 1, 2, 3, and 4 weeks after Virus-induced gene silencing (VIGS), respectively.

sampling, both *GES* expression level and *trans–cis* nepetalactone amount were substantially reduced (Figures 5B, D). This indicates that the visual marker is a great "chaperone" in uncovering gene function.

G8H is proposed to mediate the conversion from geraniol to 8-hydroxygeraniol on the nepetalactone biosynthesis pathway; however, its endogenous role has not yet been validated. After interfering with the expression of *G8H* by VIGS, the content of *trans–cis* nepetalactone and the expression level of *G8H* were slightly diminished by 15% and 54%, respectively (Figure 6). Although the decrease in nepetalactone content was not large, the reduction in *G8H* expression and nepetalactone production was statistically significant ($p < 0.05$). These findings corroborated the involvement of *G8H* in nepetalactone biosynthesis in plants and demonstrated that cotyledon-based visual VIGS was valid in elucidating gene function in *N. cataria*.

TABLE 1 Assessment of cotyledon-based VIGS efficiency.

Species	Infiltration time	Chlorosis no.	Infiltration no.	VIGS efficiency %
Nepeta cataria (Nc)	7 DAS	13	15	84.4 ± 2.8
		23	27	
		18	22	
	9 DAS	15	19	75.1 ± 3.8[a]
		20	28	
		18	24	
Nepeta mussinii (Nm)	7 DAS	20	40	47.7 ± 2.9
		17	35	
		12	27	
	9 DAS	6	15	36.7 ± 3.4[a]
		7	19	
		7	21	

VIGS efficiency indicates mean ± SD.
VIGS, virus-induced gene silencing; DAS, days after sowing.
[a] Relative to the corresponding VIGS efficiency of 7 DAS and *p*-value <0.05.

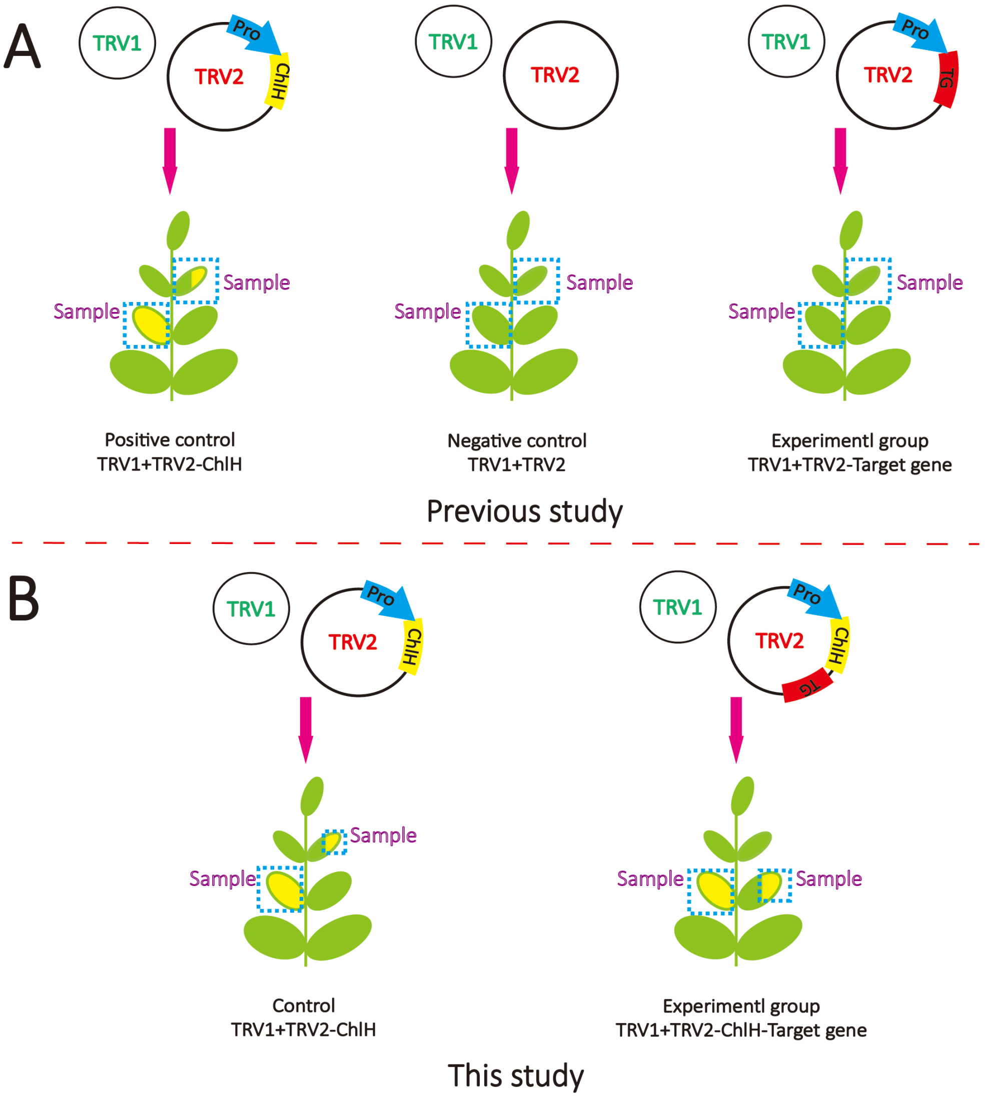


FIGURE 4
Schematic diagram of virus-induced gene silencing (VIGS) methods. **(A)** The VIGS method is commonly used in previous studies. TRV1 contains essential components for virus assembly, which is always co-transformed with TRV2, which contains the fragments of target genes. **(B)** The VIGS method was used in this study. The fragments of *ChlH* and target gene (TG) were inserted in TRV2 separately under the driving of one promoter (Pro).

Discussion

Virus-induced gene silencing is a remarkable tool for gene function studies in plants recalcitrant to genetic transformation. Numerous methods, such as leaf injection, spray inoculation, and vacuum infiltration, have been devised to conduct agrobacteria-mediated VIGS in diverse plant species (Ratcliff et al., 2001; Liu et al., 2002; Zhang et al., 2017), but limitations still exist, including the reliance on equipment in spray inoculation and vacuum infiltration and long waiting period for plant growth. Here, we presented a cotyledon-based visual VIGS system in *N. cataria* with the advantages of a short VIGS period, good VIGS efficiency, highly effective gene function validation through precise sampling with visual indicator, and adaptable potential in other plant species like *N. mussinii*.

The period of VIGS varies significantly according to the different methods. It can take some time for the plant to grow to be fit for infiltration. Leaf injection is usually executed on true leaves like the fully expanded leaves of 3-week-old *Nicotiana benthamiana* or *Solanum lycopersicum* (Ratcliff et al., 2001; Liu et al., 2002). The onset and persistence of chlorosis also differ. In tobacco, the chlorotic phenotype emerged 10 days post infiltration (dpi) and ended in the new leaves after 28 dpi (Ratcliff et al., 2001). Similarly, in tomato, the phenotype appeared at 10 dpi and lasted for

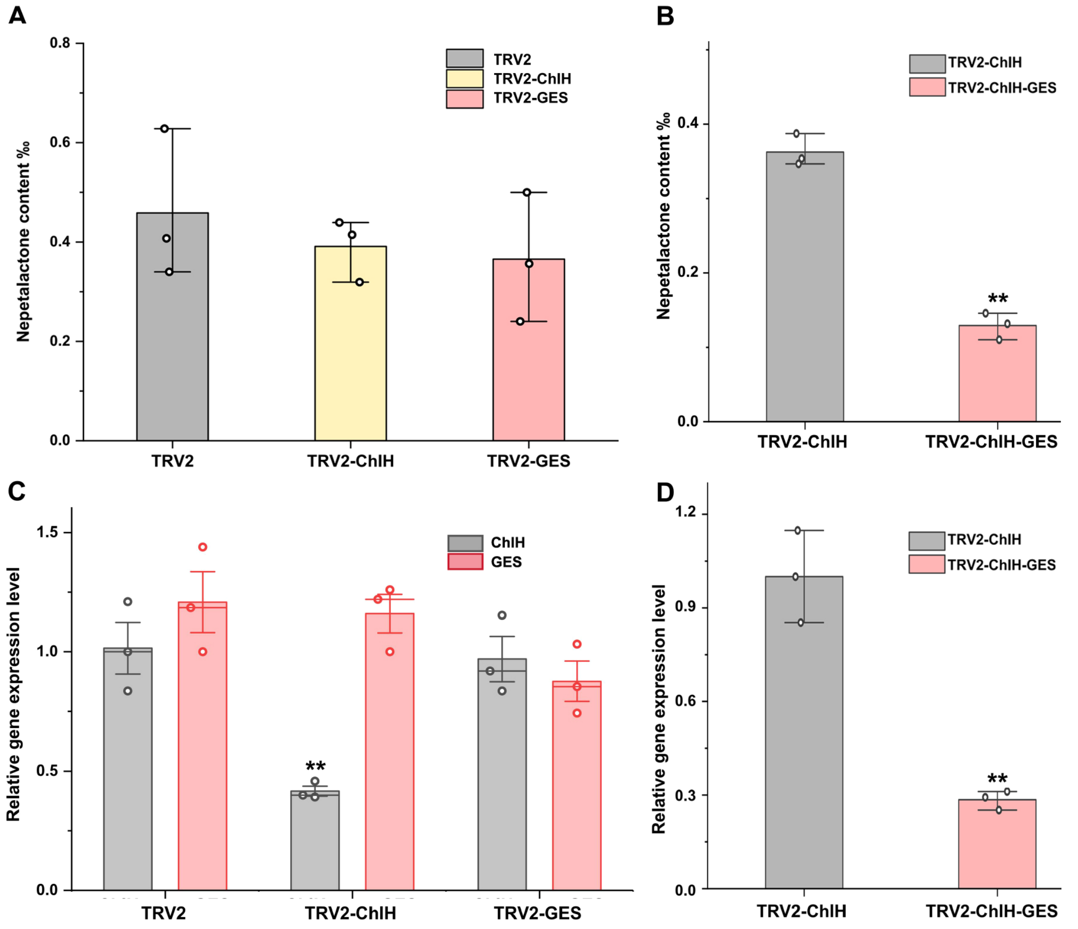


FIGURE 5
Assessment of the impact of visual marker on elucidating gene function in virus-induced gene silencing (VIGS). Panels **(A, C)** were without indicator; the content of *trans–cis* nepetalactone **(A)** and the expression level of *ChlH* and *GES* **(C)** in the first and second true leaves were measured. TRV2 empty vector was used as the negative control group. In panels **(B, D)**, the target gene's fragment was co-expressed with the visual marker *ChlH*. After *GES* gene interference by VIGS, the content of *trans–cis* nepetalactone **(B)** and *GES* expression level **(D)** in the chlorosis part were detected. TRV2-ChlH was used as the control group. Error bars were the standard error (n = 3 or 4). ** above the bracket indicates *p*-value <0.01 in *t*-test.

approximately 1 month (Liu et al., 2002). In *Euphorbia peplus*, chlorosis sustained from 5 to 40 dpi (Czechowski et al., 2022), while in cotton, the phenotype lasted from 2 to 3 weeks after infiltration to the flowering stage (Si et al., 2022). Palmer established the wounded node VIGS method in *N. cataria*, which encompassed steps like cutting branches, rooting, infection by wounding nodes, and sampling. The whole procedure was estimated to span 8–11 weeks including 5 weeks for plant growth (cutting and rooting) and 3–6 weeks for chlorosis appearance before sample collection. The approach here took only 1 week for germination and 2 weeks for chlorosis and sampling. Consequently, the entire cotyledon-based VIGS system in *N. cataria* was 3 weeks, which was considerably faster than the previous methods and independent of facilities.

The efficiency is critical to the applications of VIGS like gene function validation and pest defense. The concentrations of the infiltration solution contribute much to the VIGS efficiency. When the optical density at 600 nm (OD_{600}) was 1.2, efficiency was the highest in tea (*Camellia sinensis* [L.] O. Kuntze) (Li et al., 2023), whereas in *N. benthamiana*, the optimal efficiency was observed with an OD_{600} of 0.8 and the lowest with OD_{600} of 0.1 (Li et al., 2024). The VIGS efficiency varies much across different plant species even utilizing the same infiltration approach. As in our case, the efficiency was lower in *N. mussinii* compared to *N. cataria*. Similar results were reported in 19 soybean genotypes, where the silencing rate was 100% in PI 567301B, while it was 0% in 10 other genotypes such as Williams (Gedling et al., 2018).

The VIGS methods are continuously being refined to attain higher efficiency. A root wounding immersion method was applied in *N. benthamiana* and tomato, resulting in a silencing rate of 95%–100% (Li et al., 2024). The extracts from agroinfiltrated *N. benthamiana* leaves were developed for soybean VIGS infiltration; the apical puncture inoculation method achieved 60% efficiency compared with 12.5% by rub-inoculation method in the Jack

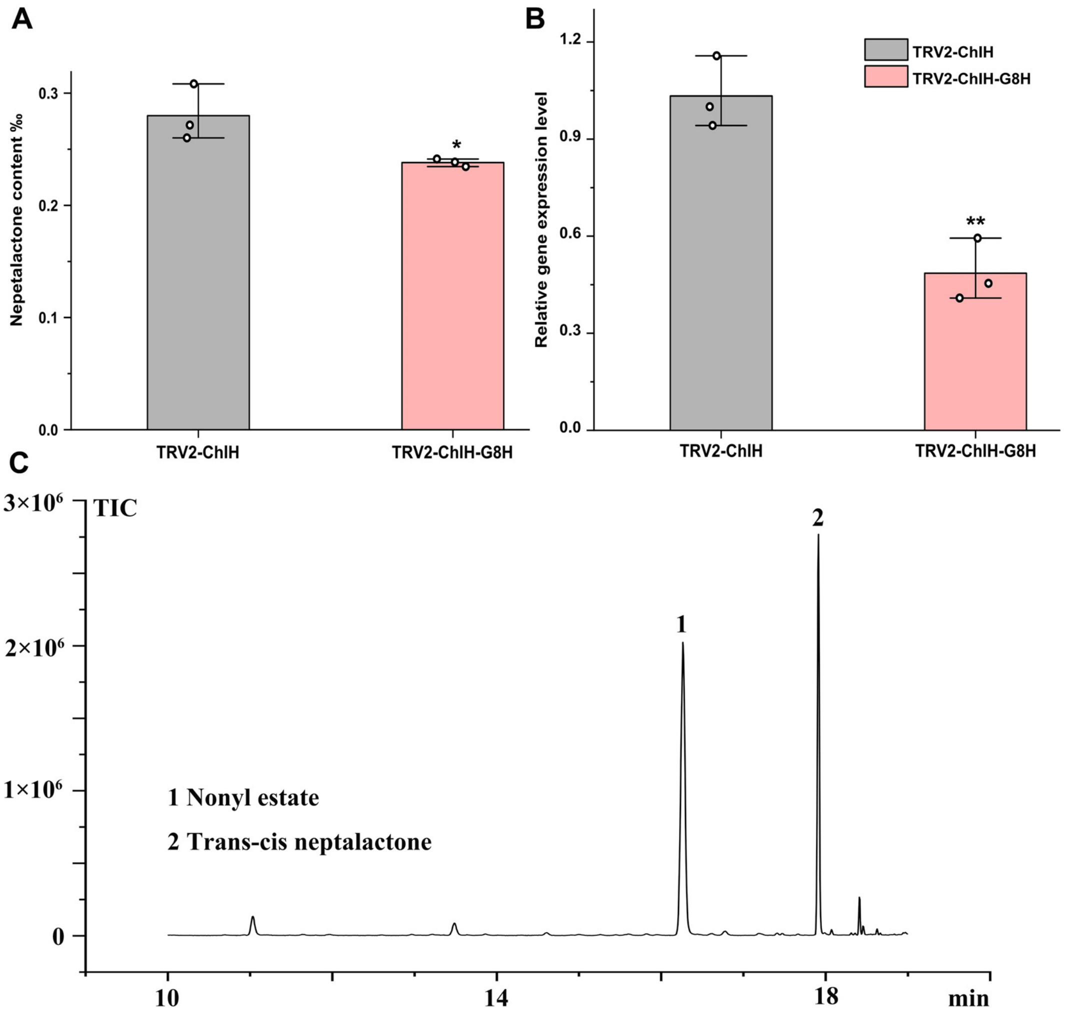


FIGURE 6
Characterization of the function of *G8H in vivo* by virus-induced gene silencing (VIGS). The content of *trans–cis* nepetalactone **(A)** and *G8H* expression level **(B)** were measured in the cotyledon-based visual VIGS experiment. **(C)** Representative gas chromatography–mass spectrometry (GC–MS) diagram from the control TRV2-ChlH. Nonyl acetate was used as the internal standard. Error bars were the standard error (n = 3). * and **above the bracket indicate *p* value < 0.05 and 0.01 respectively in *t* test.

genotype soybean (Gedling et al., 2018). In *Nepeta*, the wounded node method was established, but the efficiency remained relatively low (between 25% and 32%), and the chlorosis phenotype was often confined to the vicinity of the infection site (Palmer et al., 2022). In our case, the cotyledon-based VIGS achieved an average efficiency of 84.4%, and the phenotype reliably spread to the first two pairs of true leaves. Overall, we provided an easy yet effective cotyledon-based VIGS and verified the role of *G8H* in nepetalactone synthesis in *N. cataria*, which will significantly facilitate the gene functional analysis in *Nepeta* and holds the potential to be applied for other purposes in many plants.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

ZY: Conceptualization, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft. RL: Data curation, Methodology, Validation, Writing – review & editing. BH: Validation, Visualization, Writing – original draft. LY: Data curation, Funding acquisition, Methodology, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. ZY acknowledges the National Natural Science Foundation of China (No. 32360111), Program of China Scholarship Council (No. 202008210144), Science Fund for Distinguished Young Scholars of Jiangxi Province (No. 20224ACB215004), Double Thousand Plan of Jiangxi Province (No. JXSQ2023101106), Key Research and Development Program of Jiangxi Province (No. 20243BBI91010), and Open Project Program of Shanghai Key Laboratory of Plant Functional Genomics and Resources (No. PFGR202303). LY acknowledges Yunnan Revitalization Talent Support Program "Top Team" Project (Grant No. 202305AT350001), Science and Technology Commission of Shanghai Municipality (YDZX20223100001003 and YDZX20243100004002), and the Special Fund for Shanghai Landscaping Administration Bureau Program (Grant No. G232401).

Acknowledgments

The authors are thankful for the great help and useful discussion with Benjamin R. Lichman, Samuel J. Smit, Caragh Whitehead, Lira Palmer, Tomasz Czechowski, Kaouthar Eljounaidi, Sandesh H. Swamidatta, and Daniel C Jeffares. The authors acknowledge the Centre for Novel Agricultural Products, Department of Biology, University of York, for the use of resources and materials.

Conflict of interest

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

Generative AI statement

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1514614/full#supplementary-material

SUPPLEMENTARY FIGURE 1

Tissue expression pattern of *CHIH* reporter genes in *Nepeta*. The samples from 'mature leaf, closed flower buds, immature leaf, open flowers, petiole, root, stem' were shown as 'ML, CFB, IML, OF, P, R, S' for short.

SUPPLEMENTARY DATA SHEET 1

The cDNA sequences of Nepeta genes.

SUPPLEMENTARY DATA SHEET 2

Primers for vector construction and qRT-PCR analysis.

SUPPLEMENTARY DATA SHEET 3

The protein sequences of ChlHs used for phylogenetic analysis.

References

Boachon, B., Buell, C. R., Crisovan, E., Dudareva, N., Garcia, N., Godden, G., et al. (2018). Phylogenomic mining of the mints reveals multiple mechanisms contributing to the evolution of chemical diversity in Lamiaceae. *Mol. Plant* 11, 1084–1096. doi: 10.1016/j.molp.2018.06.002

Carbonell, A., and Carrington, J. C. (2015). Antiviral roles of plant ARGONAUTES. *Curr. Opin. Plant Biol.* 27, 111–117. doi: 10.1016/j.pbi.2015.06.013

Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., et al. (2023). TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol. Plant* 16, 1733–1742. doi: 10.1016/j.molp.2023.09.010

Cheuk, A., and Houde, M. (2019). A new barley stripe mosaic virus allows large protein overexpression for rapid function analysis. *Plant Physiol.* 180, 2306–2311. doi: 10.1104/pp.19.00759

Czechowski, T., Forestier, E., Swamidatta, S. H., Gilday, A. D., Cording, A., Larson, T. R., et al. (2022). Gene discovery and virus-induced gene silencing reveal branched pathways to major classes of bioactive diterpenoids in Euphorbia peplus. *Proc. Natl. Acad. Sci. United States America* 119, e2203890119. doi: 10.1073/pnas.2203890119

Gedling, C. R., Ali, E. M., Gunadi, A., Finer, J. J., Xie, K., Liu, Y., et al. (2018). Improved apple latent spherical virus-induced gene silencing in multiple soybean genotypes through direct inoculation of agro-infiltrated Nicotiana benthamiana extract. *Plant Methods* 14, 19–29. doi: 10.1186/s13007-018-0286-7

Jiang, Z. Q., Zhao, Q. Q., Bai, R. Y., Yu, R. N., Diao, P. F., Yan, T., et al. (2021). Host sunflower-induced silencing of parasitism-related genes confers resistance to invading Orobanche cumana. *Plant Physiol.* 185, 424–440. doi: 10.1093/plphys/kiaa018

Lei, J. F., Dai, P. H., Li, Y., Zhang, W. Q., Zhou, G. T., Liu, C., et al. (2021). Heritable gene editing using FT mobile guide RNAs and DNA viruses. *Plant Methods* 17, 20. doi: 10.1186/s13007-021-00719-4

Li, T. D., Yang, X. P., Yu, Y., Si, X. M., Zhai, X. W., Zhang, H. W., et al. (2018). Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* 36, 1160–1163. doi: 10.1038/nbt.4273

Li, G. D., Li, Y., Yao, X. Z., and Lu, L. T. (2023). Establishment of a virus-induced gene-silencing (VIGS) system in tea plant and its use in the functional analysis of CsTCS1. *Int. J. Mol. Sci.* 24, 392–406. doi: 10.3390/ijms24010392

Li, X. Y., Tao, N., Xu, B., Xu, J. Q., Yang, Z. A., Jiang, C. Q., et al. (2024). Establishment and application of a root wounding-immersion method for efficient virus-induced gene silencing in plants. *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1336726

Lichman, B. R., Kamileen, M. O., Titchiner, G. R., Saalbach, G., Stevenson, C. E. M., Lawson, D. M., et al. (2019). Uncoupled activation and cyclization in catmint reductive terpenoid biosynthesis. *Nat. Chem. Biol.* 15, 71–79. doi: 10.1038/s41589-018-0185-2

Lichman, B. R., Godden, G. T., Hamilton, J. P., Palmer, L., Kamileen, M. O., Zhao, D. Y., et al. (2020). The evolutionary origins of the cat attractant nepetalactone in catnip. *Sci. Adv.* 6, eaba0721. doi: 10.1126/sciadv.aba0721

Liu, N., Xie, K., Jia, Q., Zhao, J. P., Chen, T. Y., Li, H. G., et al. (2016). Foxtail mosaic virus-induced gene silencing in monocot plants. *Plant Physiol.* 171, 1801–1807. doi: 10.1104/pp.16.00010

Liu, Y. L., Schiff, M., and Dinesh-Kumar, S. P. (2002). Virus-induced gene silencing in tomato. *Plant J.* 31, 777–786. doi: 10.1046/j.1365-313X.2002.01394.x

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15, 394–408. doi: 10.1038/nrg3683

Misra, R. C., Sharma, S., Sandeep Garg, A., Chanotiya, C. S., and Ghosh, S. (2017). Two CYP716A subfamily cytochrome P450 monooxygenases of sweet basil play similar but nonredundant roles in ursane- and oleanane-type pentacyclic triterpene biosynthesis. *New Phytol.* 214, 706–720. doi: 10.1111/nph.2017.214.issue-2

Palmer, L., Chuang, L., Siegmund, M., Kunert, M., Yamamoto, K., Sonawane, P., et al. (2022). *In vivo* characterization of key iridoid biosynthesis pathway genes in catnip. *Planta* 256. doi: 10.1007/s00425-022-04012-z

Ratcliff, F., Martin-Hernandez, A. M., and Baulcombe, D. C. (2001). Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant J.* 25, 237–245. doi: 10.1046/j.0960-7412.2000.00942.x

Rössner, C., Lotz, D., and Becker, A. (2022). VIGS goes viral: how VIGS transforms our understanding of plant science. *Annu. Rev. Plant Biol.* 73, 703–728. doi: 10.1146/annurev-arplant-102820-020542

Shi, G. Y., Hao, M. Y., Tian, B. M., Cao, G. Q., Wei, F., and Xie, Z. Q. (2021). A methodological advance of tobacco rattle virus-induced gene silencing for functional genomics in plants. *Front. Plant Sci.* 12, 671091–671106. doi: 10.3389/fpls.2021.671091

Si, Z. F., Wu, H. T., Tian, Y., Zhang, Z. Y., Zhang, T. Z., and Hu, Y. (2022). Visible gland constantly traces virus-induced gene silencing in cotton. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1020841

Smit, S. J., and Lichman, B. R. (2022). Plant biosynthetic gene clusters in the context of metabolic evolution. *Natural Product Rep.* 39, 1465–1482. doi: 10.1039/D2NP00005A

Taochy, C., Gursanscky, N. R., Cao, J. L., Fletcher, S. J., Dressel, U., Mitter, N., et al. (2017). A genetic screen for impaired systemic RNAi highlights the crucial role of DICER-LIKE 2. *Plant Physiol.* 175, 1424–1437. doi: 10.1104/pp.17.01181

Uenoyama, R., Miyazaki, T., Hurst, J. L., Beynon, R. J., Adachi, M., Murooka, T., et al. (2021). The characteristic response of domestic cats to plant iridoids allows them to gain chemical defense against mosquitoes. *Sci. Adv.* 7, eabd9135. doi: 10.1126/sciadv.abd9135

Wu, H. H., Li, B. S., Iwakawa, H. O., Pan, Y. J., Tang, X. L., Ling-hu, Q. Y., et al. (2020). Plant 22-nt siRNAs mediate translational repression and stress adaptation. *Nature* 581, 89–93. doi: 10.1038/s41586-020-2231-y

Wu, C. J., Jia, L. L., and Goggin, F. (2011). The reliability of virus-induced gene silencing experiments using tobacco rattle virus in tomato is influenced by the size of the vector control. *Mol. Plant Pathol.* 12, 299–305. doi: 10.1111/j.1364-3703.2010.00669.x

Yao, M. Q., Chen, W. W., Kong, J. H., Zhang, X. L., Shi, N. N., Zhong, S. L., et al. (2020). METHYLTRANSFERASE1 and ripening modulate vivipary during tomato fruit development. *Plant Physiol.* 183, 1883–1897. doi: 10.1104/pp.20.00499

Zhang, J., Yu, D. S., Zhang, Y., Liu, K., Xu, K. D., Zhang, F. L., et al. (2017). Vacuum and co-cultivation agroinfiltration of (germinated) seeds results in tobacco rattle virus (TRV) mediated whole-plant virus-induced gene silencing (VIGS) in wheat and maize. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00393

Frontiers in Plant Science

TYPE Original Research
PUBLISHED 06 February 2025
DOI 10.3389/fpls.2025.1452533

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Ardha Apriyanto,
University of Potsdam, Germany
Yun Tian,
Hunan Agricultural University, China

*CORRESPONDENCE
Qiang Liu
liu.qiangcs@163.com
Yan Yang
yangyan@hnlky.cn

†These authors have contributed equally to this work

RECEIVED 21 June 2024
ACCEPTED 22 January 2025
PUBLISHED 06 February 2025

CITATION
Zeng W, Xie B, Chen Y, Chen J, Li P, Jiang L, Li C, Liu Q and Yang Y (2025) Regulatory mechanism of carbohydrate metabolism pathways on oil biosynthesis of oil plant *Symplocos paniculata*.
Front. Plant Sci. 16:1452533.
doi: 10.3389/fpls.2025.1452533

Regulatory mechanism of carbohydrate metabolism pathways on oil biosynthesis of oil plant *Symplocos paniculata*

Wenbin Zeng[1†], Beilei Xie[2†], Yunzhu Chen[2], Jingzhen Chen[2], Peiwang Li[2], Lijuan Jiang[1], Changzhu Li[2], Qiang Liu[1*] and Yan Yang[2*]

[1]College of Life and Environmental Sciences, Central South University of Forestry and Technology, Changsha, China, [2]State Key Laboratory of Utilization of Woody Oil Resource, Hunan Academy of Forestry, Changsha, China

The mechanism underlying oil synthesis in oil plant fruits remains elusive, as sugar metabolism provides the essential carbon skeleton without a clear understanding of its intricate workings. The transcriptome and oil and sugar metabolites' content of *Symplocos paniculate*, an extraordinary oil plant with immense ecological significance, were subjected to a comparative analysis throughout fruit development. The findings unveiled that the impact of sugar metabolism on oil synthesis varied throughout distinct stages of fruit development. Remarkably, during the initial phase of fruit development from 10 to 90 days after flowering (DAF), pivotal genes involved in starch biosynthesis, such as ADP-glucose pyrophosphorylase (*AGP*), starch synthase (*SS*), and starch branching enzyme (*SBE*), facilitated an earlier accumulation of starch within the fruit. Whereas, during the fruit maturation stage (from 90 DAF to 170 DAF), the expression of phosphofructokinase 1 (*PFK-1*), pyruvate kinase *(PK)* and pyruvate dehydrogenase (*PDH*) enzyme genes involved in the glycolysis pathway was significantly upregulated, thereby facilitating a rapid and substantial accumulation of oil. The sugar metabolism activity of *S. paniculata* fruit exerts a crucial influence on the process of oil synthesis, which is highly dependent on the specific developmental stage. These significant discoveries provide potential candidate genes for advanced genetic improvement using molecular biotechnology, thus enhancing both fruit oil production and modifying the composition of fatty acids.

KEYWORDS

***Symplocos paniculata*, carbohydrate metabolism, oil biosynthesis, transcriptome, regulatory mechanism**

1 Introduction

Symplocos paniculata (Thunb.) Miq., a member of the *Symplocaceae* family, is a woody oil plant indigenous to China and holds significant ecological and economic importance. The aforementioned statement has been recognized as one of the most auspicious contenders for the production of bioenergy (Liu et al., 2016). A fully matured tree possesses the capacity to yield a plentiful harvest of up to 20 kg of fruit. The entire fruit showcases a 36.6% oil content with a multitude of oil cells predominantly located in the pericarp, thereby determining the optimal utilization of the entire fruit during oil extraction (Liu et al., 2015). The abundant yield and affluent oil content of *S. paniculata* showcase its immense potential for application in biodiesel production, as well as other industrial sectors such as surfactants, soap, and lubricants (Li et al., 2023). Furthermore, *S. paniculata* demonstrates remarkable adaptability across varying temperature zones and soil conditions. It flourishes in desolate, saline, and severely drought-stricken soil (Liu et al., 2015). Moreover, its naturally stunted growth requires minimal management for commercial-scale cultivation. Nevertheless, there have been scant reports on the exploitation and utilization of this species.

The economic value of this woody plant is often determined by the fruit oil content and quality, as it serves as the primary metabolite of *S. paniculata*. Therefore, comprehending the biosynthesis pathway and regulatory mechanisms of oil becomes crucial in order to enhance seed oil content and composition in plants. Many studies have demonstrated that the synthesis of oil is a complex process involving the distribution of carbon sources, which encompasses numerous enzyme genes and their regulated expression in various cellular compartments such as plastids, endoplasmic reticulum, and cytoplasm (Xu and Shanklin, 2016). This intricate process includes pathways for fatty acid biosynthesis, triacylglycerol (TAG) assembly, as well as glycometabolism (starch biosynthesis and metabolism, glycolysis, gluconeogenesis, and citrate cycle) (Lin et al., 2017).

Carbohydrate metabolism plays a pivotal role in seed development, as exemplified by the carbohydrate polymers cellulose and hemicellulose that constitute the major components of cell walls, while starch serves as the primary reserve for seed development (Weichert et al., 2010). Not only do carbohydrates provide fundamental nutrients and energy for seed growth and development, but they also furnish precursors (glycolate pyruvate) for oil biosynthesis derived from carbohydrate decomposition and metabolism (Hill, 2003). The tricarboxylic acid cycle furnishes the ATP required for fat biosynthesis, while Glucose-6-phosphate in Glycolysis pathway serves as the primary carbon source for fatty acid biosynthesis (Liu et al., 2013). The study of developmental gene expression profile of Jatropha (*Jatropha curcas* L.) seeds has underscored the crucial regulatory effects of glucose metabolism, starch metabolism and other carbohydrate metabolic pathways on oil biosynthesis (Jiang et al., 2012). Bourgis et al. have reported that the remarkable increase of oil in oil palm is not only correlated with elevated transcription levels of fatty acid synthetases, but also associated with key enzymes involved in specific plastid carbohydrate metabolism (such as phosphofructose kinase, pyruvate kinase and pyruvate dehydrogenase, etc.) (Bourgis et al., 2011). Furthermore, numerous studies have demonstrated that the manipulation of genes associated with carbohydrate metabolism can significantly enhance oil accumulation in plants' seeds (Allen et al., 2011).

Unfortunately, the close relationship between carbohydrate metabolism and oil biosynthesis has received little attention (Dunahay et al., 1995). Furthermore, there continues to be a dearth of comprehensive investigations into this intricate correlation. Zhao et al. concluded that in Yellowhorn (*Xanthoceras sorbifolia* Bunge), oil is primarily derived from sugar based on the negative correlation observed between changes in soluble sugar quantity dynamics and oil content during seed development (Zhao et al., 2015). Whereas, previous studies have shown a positive correlation between the dynamic changes in oil and carbohydrate quantities (Kennedy et al., 2011). Furthermore, Borek et al. suggest that the inconsistent correlation observed between carbohydrate metabolism and oil biosynthesis in different plant species' seeds may be attributed to their genetic backgrounds (Borek et al., 2009). Generally speaking, soluble carbohydrates (primarily sucrose) are synthesized by plant leaves through photosynthesis and subsequently transported to the seeds. The relationship between seeds and leaves is one of sink and source in terms of carbohydrate supply and demand. Sucrose serves as the principal form for translocating carbohydrates from source to sink regions, acting as a physiological precursor for fatty acid biosynthesis in oilseeds (Zhao et al., 2015). Hence, the endeavor to scrutinize the impact of carbohydrate metabolism on oil biosynthesis solely based on their quantity dynamic change correlation is insufficient. The formation of triacylglycerols in developing oilseeds is one of the primary biosynthetic events in the cytoplasm of storage cells but the flow of carbon through different carbohydrate metabolism pathways affecting oil biosynthesis is not fully understood. and whether there is a competition of intermediate precursor utilization among them is still unknown.

In this study, fresh fruits of *S. paniculata* at different developmental stages from the same tree were selected as experimental materials. and aim to (1) Examine the temporal fluctuations in the quantitative relationship between oil and sugars; (2) Investigate functional unigenes encoding the key enzymes associated with carbohydrate metabolism and lipid biosynthesis. (3) Identify functional unigenes through complete transcriptome annotation; (4) Reconstruct carbohydrate metabolism and lipid biosynthesis pathways using identified enzymes; and (5) Elucidate the molecular regulatory mechanisms of carbohydrate metabolism on lipid biosynthesis in *S. paniculata*.

2 Materials and methods

2.1 Plant materials

The fresh fruits (approximately 100 g) of *S. paniculata* were harvested from accession C3 at the experimental station (28° 07'10.38" N, 113°02'53.16"E, and 94.5 m) of Hunan Academy of Forestry in Changsha, Hunan, China every 10 days after flowering (DAF) throughout 2022. The sampling was terminated at 170 DAF due to fruit shedding. The fruit samples were split into two groups: The fruits in one group were subjected to drying at a temperature of 70°C for a duration of three days, followed by oil extraction and subsequent fatty acid analysis, while another group of samples was immediately removed from the mother tree and frozen in liquid nitrogen before being stored at -80°C for sugars determination and transcriptome analysis.

2.2 Oils component determination

The oil content of the fruit was determined through meticulous soxhlet extraction. The dried fruits were finely ground into a delicate powder and meticulously weighed (M_1, g), followed by meticulous extraction using petroleum ether (99.7%) at 60°C for a duration of 8 hours. The residual sample underwent vacuum drying at a temperature of 105°C for a duration of two hours, followed by weighing (M_2, g), Oil content (%) $=(M_1-M_2)/M_1 \times 100\%$. Oil extraction was repeated three times for each fruit development stage.

The saponification of 30 mg of fruit oil was carried out in a 2 mL blending solution (0.5 mol/L), which was prepared by adding 28.1 g of potassium hydroxide (KOH) to one liter of methyl alcohol (99.9%). The resulting saponified fruit oil was separated into layers by the addition of 10 mL each of petroleum ether (99.9%) and deionized water, followed by centrifugation at 3000 r/min for 5 minutes using a refrigerated centrifuge model 3H16RI from Hunan Hexi Instruments Co., Ltd., Chang Sha, China. The fruit oil's fatty acid components were analyzed using a Nexis GC-2030 gas spectrometer (Shimadzu, Japan), and the resulting supernatant was injected into an Agilent HP-88 column with dimensions of 0.25mm×100m for free fatty acid polyester separation. The Supelco 37 Component FAME Mix was utilized for the identification of the fatty acid components.

2.3 Sugars content determination

The sugars, including starch, sucrose, fructose, and glucose, were extracted using the modified methods described in a previous study (Fu et al., 2023). In brief, the fresh fruit tissues were finely powdered, weighed accurately, and incubated with 80% ethanol at a temperature of 80°C for a duration of 2 hours. The resulting supernatant was collected for the quantification of sucrose, fructose, and glucose levels. The pellet underwent digestion using a diluted hydrochloric acid solution (3 mol/L) and subsequently neutralization with sodium hydroxide to facilitate starch determination. The determination of starch, sucrose, fructose, and glucose was executed utilizing corresponding detection reagent kits (Solarbio Science & Technology Co., Ltd, Beijing, China) in strict accordance with the provided protocols. Each measurement was conducted with six replicates to ensure accuracy and reliability.

2.4 Microscopic observation

The collected fruits were fixed with FAA fixative (a mixture of formaldehyde, acetic acid, and ethyl alcohol, volume ratio 5:5:90), washed, immersed in a glycerol solution, and subjected to vacuum infiltration to ensure complete penetration of glycerol into the tissue voids of the fruit samples. The frozen microtome was set at a temperature of -30°C for sectioning. Fruit samples were rapidly frozen for 30 seconds, and sections with a thickness of 25 μm were obtained as temporary slides (Wang and Guo, 2021). These slides were then stained with Sudan III and K-I staining solution for 20 minutes and subsequently examined using the Motic digital microscopy system (Er et al., 2000).

2.5 Transcriptome sequencing analysis

The total RNA of the fruits from different representative developmental stages was extracted in accordance with the manufacturer's protocol, utilizing the Spin Column Plant Total RNA Purification kit (Sangon Biotech, China). The cDNA library construction was performed following the operational instructions provided by the mRNA-Seq Sample Preparation Kit (Illumina Biotech, USA). The Illumina HiseqTM 2000 sequencing platform was employed for the purpose of sequencing, while Trinity was utilized to *de novo* assemble the transcriptome in order to acquire the non-redundant Unigene (Grabherr et al., 2011). The high-quality reads were deposited in the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) database, and their corresponding accession numbers SRA357712 were assigned.

The well-assembled unigenes have been aligned to the following esteemed public protein databases: the Non-redundant (NR) protein database, the Gene Ontology (GO) protein database (Conesa et al., 2005), the Clusters of Orthologous Groups (COGs) protein database (Tatusov et al., 2003), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) protein database (Kanehisa and Goto, 2000). And the expression levels of the unigenes were statistically calculated using fragments per kilobase transcriptome per million mapped reads (FPKM) in accordance with Mortazavi et al.'s methodology from 2008. Subsequently, a hierarchical clustering analysis was conducted based on the expression levels of various unigenes.

2.6 qRT-PCR verification

The extraction of fruit RNA and the construction of cDNA were carried out as previously described in section 2.4. Twelve key genes associated with glucose metabolism and lipid anabolism (*DGAT1, LPAAT, PADT, SAD, FATA, ACC, PDH, PFK-1, SS, HK MDH* and *GPI*) were chosen as the target genes. The primers used in the assay were designed by Primer Premier 5.0 which are provided in Supplementary Table S1. Each reaction being performed in triplicate.

The qRT-PCR reaction program was conducted as follows: an initial denaturation step at 95°C for 2 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds, annealing and extension at 60°C for 15 seconds, and a final extension step at 95°C for 15 seconds. The relative gene expression levels were calculated with the comparative cycle threshold method (ΔΔCt), where ΔΔCt = (Ct value of target genes in the test group - Ct value of housekeeping genes in the test group) - (Ct value of target genes in the control group - Ct value of housekeeping genes in the control group). The ΔΔCt value and FPKM value of the target unigene at 10 DAF were normalized to 1, and the data at 90, 130, and 170 DAF in other time points were calculated as multiples of the gene expression levels at 10 DAF (Schmittgen and Livak, 2008).

3 Results

3.1 The dynamic changes of oil content and components

The temporal dynamics of morphological features, oil and fatty acid in the *S. paniculata* fruit were studied throughout its developmental stages. The elongated oval green fruit begins to take shape from 10 days after flowering (DAF), followed by a rapid expansion in volume into yellow-green ovoid fruits up to 90 DAF. Subsequently, there is a slight alteration in the fruit's volume, accompanied by a transformation of the peel color from yellow-green to shades of yellow-purple and purple, ultimately leading to the fruit's maturation (Figure 1A). While only a small amount of oil (0.23~2.56%) was accumulated in the fruit from 10 to 70 DAF, a remarkable shift occurred during the period between 90 to 140 DAF, with an average daily increase of 0.5% in oil content. Thereafter, the oil content continued to gradually increase until 170 DAF, at which point it reached its maximum fruit oil content of 37% (Figure 1B).

The composition of *S. paniculata* fruit oil was predominantly comprised of palmitic acid (C16:0), stearic acid (C18:0), oleic acid (C18:1), linoleic acid (C18:2), and linolenic acid (C18:3) in abundance (Figure 1C). The most significant changes entailed a remarkable surge in the proportion of C18:1, skyrocketing from 12% to an impressive 50%, while concurrently witnessing a substantial decline in C16:0 content, plummeting from 35% to a mere 18%. Although alterations in C18:2 content were relatively scarce, they occurred at a quantitatively moderate yet noticeably accelerated pace (27%), surpassing the rate observed during the initial stages of development (22%). The proportions of C18:3 and C18:0 exhibited a continuous decline throughout its development, albeit remaining very low but not negligible in the mature stage (Figure 1D).

Collectively, the morphological changes and oil accumulation in *S. paniculata* fruits development can be categorized into three distinct stages.: the Fruit Rapid Expansion (FES) stage (10-90 DAF), the Oil

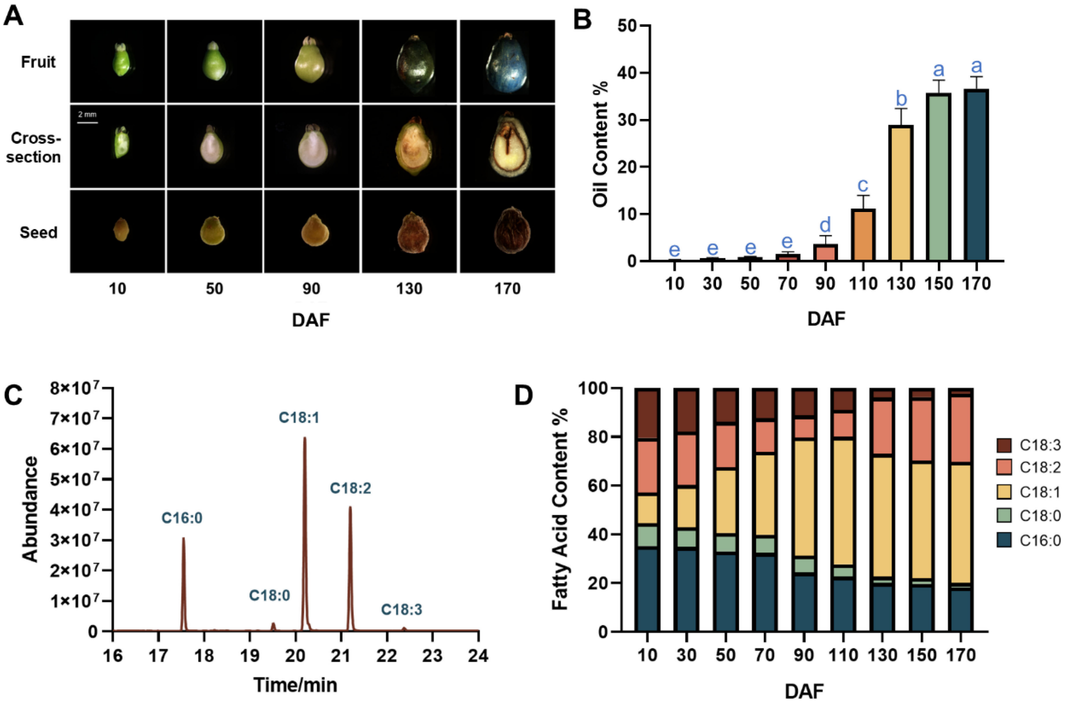


FIGURE 1
The dynamic changes of oil content and components during *S. paniculata* fruit development. **(A)** the fruit morphological changes, from top to the bottom of figure row illustrates the fruit external morphology, fruit cross-section and seed; **(B)** Dynamic change of fruit oil content, letters indicate differences between developmental stages by Tukey's honestly significant difference multiple comparison at P < 0.05; **(C)** The gas chromatograph of fruit oil; **(D)** Dynamic change of main fatty acid in fruit oil.

Rapid Accumulation (ORA) stage (90-130 DAF), and the Fruit Dis-Coloration (FDC) stage (130-170 DAF). Fresh fruits at four distinct developmental stages, namely 10, 90, 130 and 170 DAF, were meticulously handpicked for further transcriptomic analysis.

3.2 Temporal pattern of sugar metabolites

The accumulation of oil in fruits is accompanied by the metabolism of carbohydrate substances. The levels of starch, sucrose, glucose, and fructose were determined during the development of *S. paniculata* fruit. The starch content increased progressively during fruit development, reaching a peak of 0.56 mg/g at 50 DAF. Subsequently, it remained relatively stable throughout the ORA stage (90-130 DAF), followed by a rapid decline during the FDC stage. At maturity, the final fruit exhibited a starch content of 0.31 mg/g (Figure 2A). The sucrose content exhibited a significant decrease, reaching a minimum of 0.32 mg/g during the FES stage (10-90 DAF), followed by a rapid increase in the fast stage, peaking at 3.346 mg/g at 130 DAF. Throughout the FDC stage (130-170 DAF), there was minimal variation in sucrose content, which remained consistently stable (Figure 2B). The glucose and fructose contents exhibited similar patterns, characterized by a significant decrease in the 10-50 DAF content, minimal changes during rapid oil accumulation, and a continuous decline during fruit ripening (Figures 2C, D). These findings suggest that the

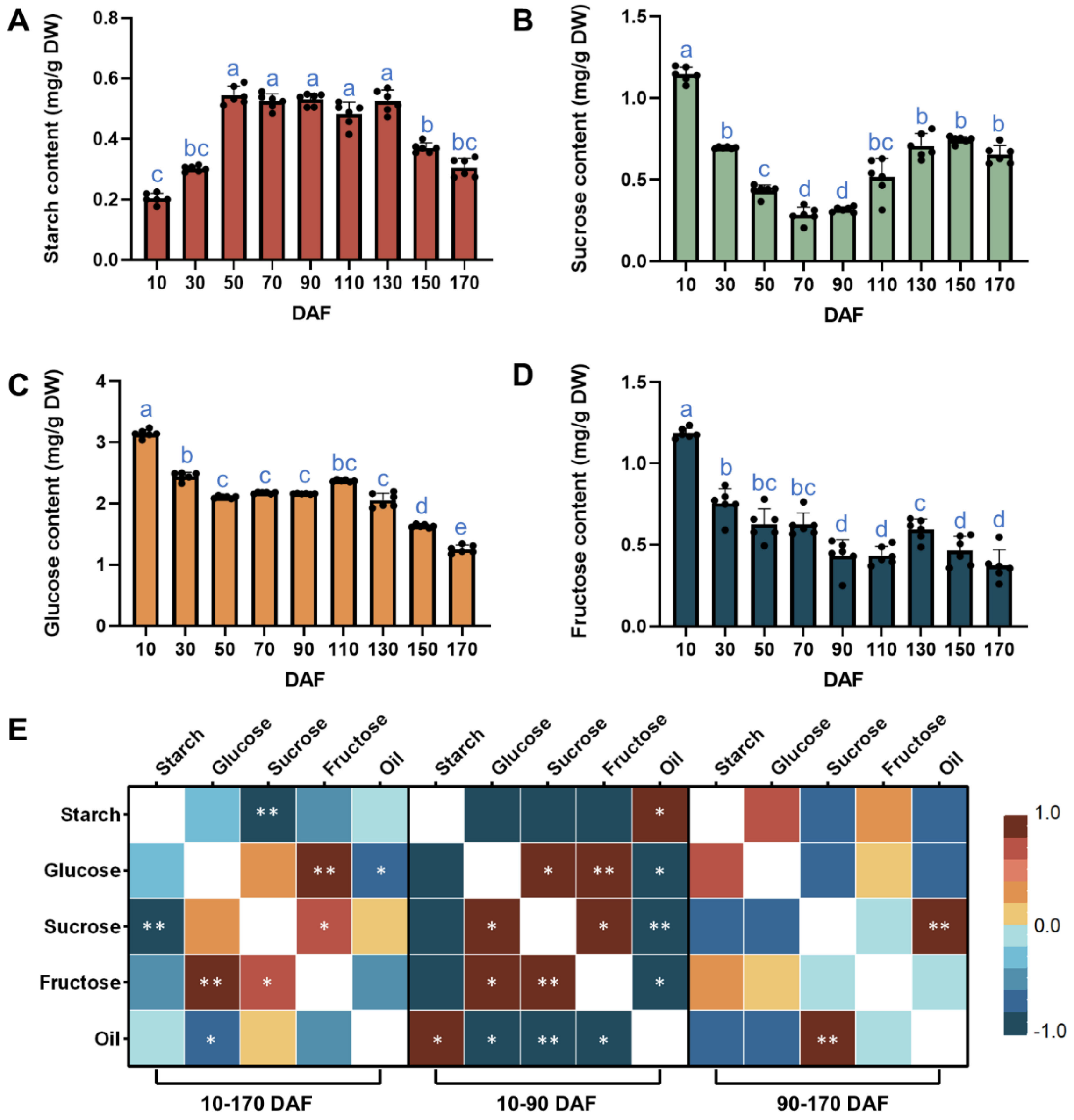


FIGURE 2
Temporal pattern of sugar metabolites in *S. paniculata* fruit **(A)** Dynamic change of starch content; **(B)** Dynamic change of sucrose content; **(C)** Dynamic change of glucose and content; **(D)** Dynamic change of fructose content; **(E)** Correlation analysis of dynamic changes of main sugar and oil metabolites. Letters indicate differences between developmental stages by Tukey's honestly significant difference multiple comparison at $P < 0.05$; ** and * indicates significant correlation at $p < 0.01$ and $p < 0.05$ by two-sided test.

alterations in various sugar contents were distinct and did not align with variations in oil contents.

The correlation between sugar and oil content was further examined. Throughout the entire fruit development stage (10-170 DAF), no significant correlation was observed between oil content and sugar content, except for a significant negative correlation between oil content and glucose. Fructose and glucose exhibited a highly significant positive correlation, while starch and sucrose displayed a highly significant negative correlation. During the 10-90 DAF period of fruit development, there was a positive correlation between oil content and starch content, while a negative correlation was observed with sucrose, glucose, and fructose levels. However, during the critical phase of oil synthesis from 90 to 170 DAF, a highly significant positive correlation between oil and sucrose was observed (Figure 2E). No significant correlation was observed with the other three sugars, suggesting that there exists a differential association between sugars and oils at different stages of fruit development.

3.3 Temporal pattern of sugar and oil microstructure in fruit cell

Microstructural analysis at different stages of fruit development provides a comprehensive understanding of the oil synthesis process. Microscopic examination revealed that *S. paniculata* fruits (mainly mesocarp and seeds) contained oil, which was stored as oil cells (OC) and cellular lipid droplets (LD). Starch granules (SG) served as the storage structure for starch, while soluble sugars such as sucrose, fructose, and glucose did not exhibit any discernible microstructure (Figure 3).

During the initial stage of fruit development (10 DAF), OC were observed in the mesocarp of fruit, scattered throughout the tissue and exhibiting a pale-yellow coloration. The flesh parenchyma cells contained SG, while seed cells lacked both SG and LD. Considering Figure 1B, it can be inferred that oil synthesis primarily originated from the mesocarp OC during this early phase of fruit development. Upon rapid fruit oil accumulation (90 DAF), the color of oil cells in the mesocarp transitioned to an orange hue, accompanied by a substantial presence of mature SG in parenchyma cells, which exhibited an oblong or kidney-shaped morphology. Additionally, small SG also emerged within seed cells. This observation was further supported by metabolic content analysis depicted in Figure 2A, confirming that fruit starch content reached its peak level.

At the stage of discoloration (130 DAF), a substantial accumulation of orange LD became evident in the parenchyma cells of the mesocarp, accompanied by a gradual reduction in SG. Additionally, an abundance of LD was observed within the seed cells. At 170 days after fruit maturation (DAF), the OC exhibited an orange-red coloration, while the fusion of small and medium LD resulted in the formation of large LD within parenchyma cells.

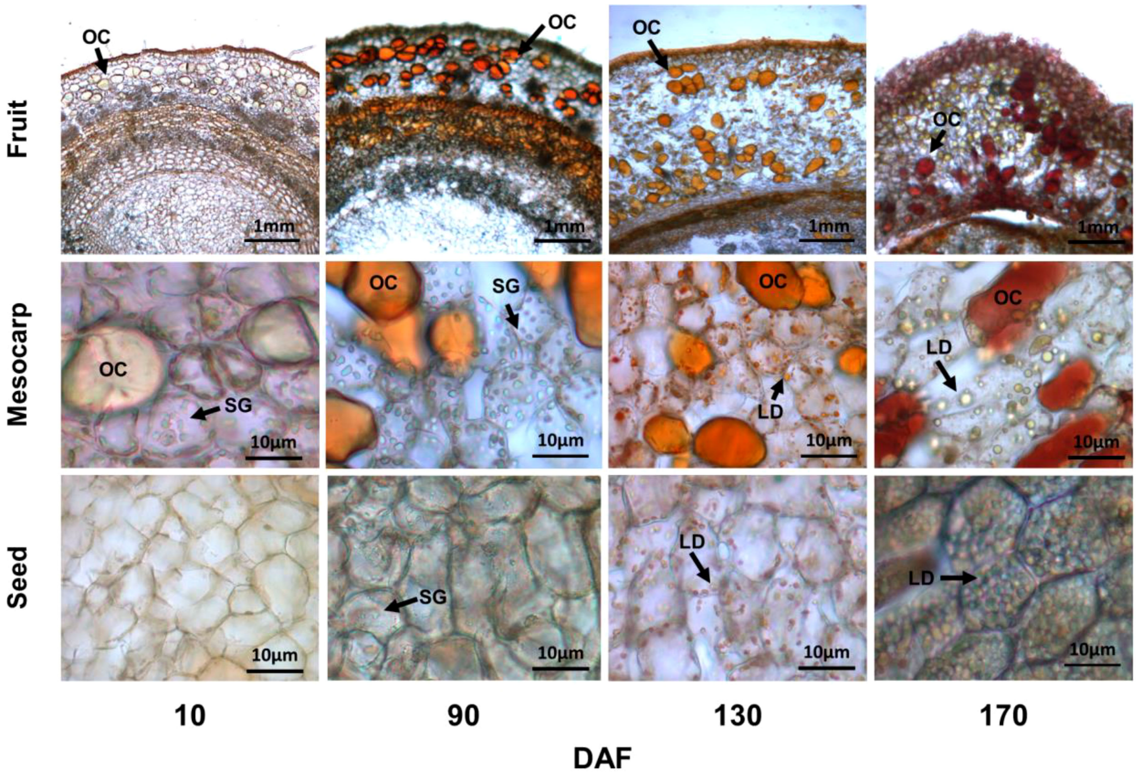


FIGURE 3
The various developmental stages of fruit cell microstructure. All Sample sections were obtained by frozen sections at 25um, Sudan III (0.5% Dissolved in ethanol solution) staining of oil cells (OC), lipid droplets (LD) and starch granules (SG) by iodine staining (0.5g/L Potassium iodide solution).

Furthermore, the seed cells contained spherical-shaped LD that completely occupied the cell volume. These findings suggest that oil synthesis in *S. paniculata* primarily occurs in both mesocarp and seed tissues, and the oil synthesis in the mesocarp precedes that in the seed; The synthesis of a substantial quantity of SG precedes the synthesis of oil, whether in the mesocarp or seeds. Additionally, there is an additive effect of OC on fruit's oil synthesis.

3.4 Functional annotation of non-redundant unigenes

In this study, a total of approximately 27,827,593 high-quality clean reads (representing 93.86% of the raw reads) were successfully obtained. Following filtration and culling processes, an impressive set of 182,904 non-redundant unigenes were generated. Notably, as the length of unigene sequences increased gradually without any discernible discontinuity (as depicted in Figure 4). The aforementioned findings imply that the RNA sequencing was executed with impeccable continuity and exceptional quality, thereby facilitating subsequent analysis.

Blast was performed for the unigenes' functional annotation. The functional annotations of unigenes were obtained by comparing public databases (NR, GO, COG and KEGG). The hit unigenes were meticulously annotated across a total of 32 KEGG pathways. Among these, an impressive count of 1,776 unigenes were precisely mapped to the intricate realm of lipid metabolic pathways, while an astounding number of 3,493 unigenes triumphantly traversed the vast expanse of carbohydrate metabolic pathways (Figure 5).

Unigene annotation findings of metabolic pathways related to carbohydrate and lipid metabolism were further analyzed. The pathway of "fatty acid metabolism" boasted the most abundant repertoire of unigenes, totaling 249 in number, this is followed by a substantial number of unigenes in the areas of "glycerolipid metabolism" (212 unigenes), "fatty acid biosynthesis" (148 unigenes), and "glycerophospholipid metabolism" (142 unigenes) (Figure 6A). Whereas, among the 15 carbohydrate metabolism pathways, the most frequently annotated unigenes were associated with "Starch and sucrose metabolism (731 unigenes)", "Glycolysis and Gluconeogenesis (718 unigenes)", "Pyruvate metabolism (579 unigenes)", and "Amino sugar and nucleotide glucose metabolism (579 unigenes)". there are additional pathways related to fatty acid synthesis and metabolism: 'Pentose phosphate pathway' has 307 unigenes and 'Fructose and mannose metabolism' has 378 unigenes (Figure 6B).

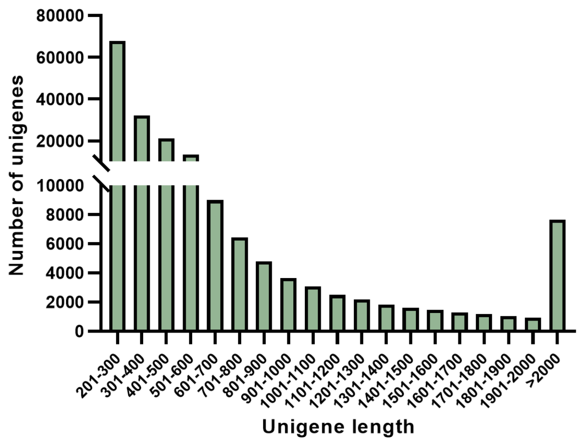


FIGURE 4
Frequency of *S. paniculata* unigenes.

The RPKM values of annotated unigenes associated with carbohydrate and lipid metabolism were utilized to conduct a hierarchical cluster analysis, resulting in the elegant classification of these unigenes into four distinct groups. Unigenes within each cluster exhibited identical or similar expression patterns during fruit oil accumulation stages (Figure 7A). The expression pattern of Cluster I and Cluster II exhibited an up-regulated trend, with the FPKM value of the unigenes continuously increasing throughout fruit development. In contrast, Cluster III showed a stable trend during fruit development. Additionally, Cluster IV displayed a down-regulated pattern (Figure 7B). The clustering results indicate a relatively distinct expressional profile during the *S. paniculata* fruit development.

3.5 The gene expression profiles of sugar and oil metabolism

The KEGG pathway assignment and functional annotation of the unigenes have revealed the presence of pivotal enzymes involved in carbohydrate and lipid metabolism pathways, which are comprehensively presented in Supplementary Table S2. Based on these identified enzymes, an elaborate depiction of the intricate processes governing carbohydrate and lipid metabolism in *S. paniculata* fruit has been meticulously illustrated (Figures 8, 9). In order to gain a comprehensive understanding of the distinct expressional patterns exhibited by specific genes associated with fruit development and oil accumulation, we conducted a comparative analysis of the RPKM values of unigenes across different phases of *S. paniculata*'s fruit oil accumulation.

At the FES stage (10-90 DAF), the oil accumulation exhibited a gradual pace (Figure 1), while the pivotal genes involved in glycolysis and tricarboxylic acid cycle (TCA) associated with glucose metabolism displayed low or even down-regulated expression levels. While the enzyme genes responsible for starch biosynthesis, such as ADP-glucose pyrophosphorylase (*AGP*), starch synthase (*SS*), and starch branching enzyme (*SBE*) have shown an up-regulated pattern. Notably, at 90 DAF, their expression levels were observed to be 1.5-fold, 2.4-fold, and 2.1-fold higher than those at 10 DAF.

In conjunction with Figure 2, it was observed that the starch content exhibited a rapid surge from 10 to 90 DAF, while the levels of sucrose, fructose, and glucose all experienced a decline. Furthermore, microanalysis in Figure 3 revealed substantial accumulation of starch granules (SG) within the mesocarp cells at 90 DAF. These findings suggest that during this stage, the primary metabolic activity of the fruit involves the conversion of soluble

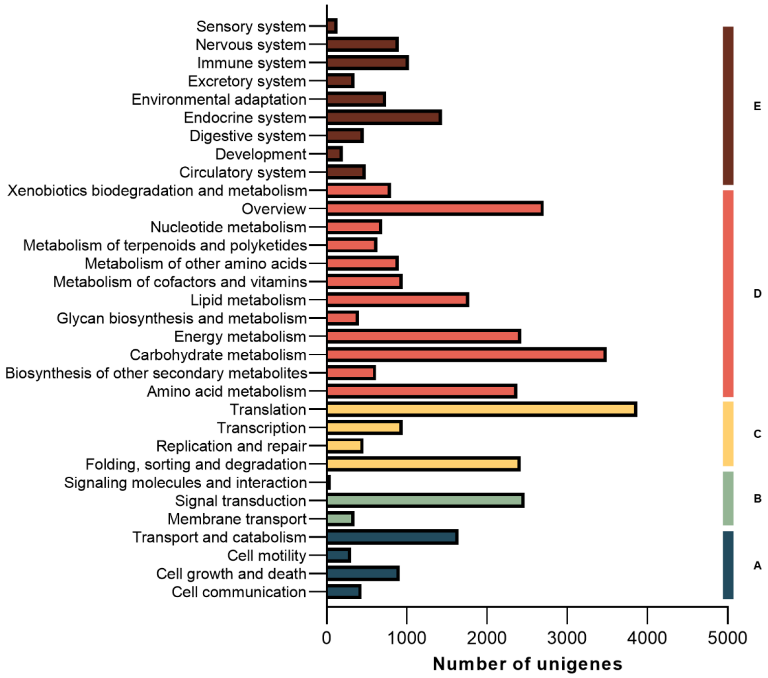


FIGURE 5
Unigene classifid by KEGG pathways. **(A)** Cellular Processes; **(B)** Environmental Information Processing; **(C)** Genetic Information Processing; **(D)** Metabolism; **(E)** Organismal Systems.

sugars into starch for storage. At the same time, it was discovered that the genes associated TAG assembly and metabolism were not actively engaged with the exception of acetyl-CoA carboxylase carboxyl transferase (*ACC*) and stearoyl-ACP desaturase (*SAD*). The pivotal gene for fatty acid desaturation, *SAD*, exhibited high expression at 10 DAF to ensure timely C18:0 into C18:1, potentially leading to substantial production of C18:1 during fruit development. However, limited accumulation of C18:0 is the underlying cause. These findings suggest a relative inactivity in genes related to fatty acids or TAG metabolisms. With fruit development, the phosphatidylcholine (PC) pathway revealed a significant up-regulation of lysophosphatidylcholine acyltransferase (*LPCAT*) and phospholipid: diacylglycerol acyltransferase 1 (*PDAT1*) between 10 and 90 DAF, indicating

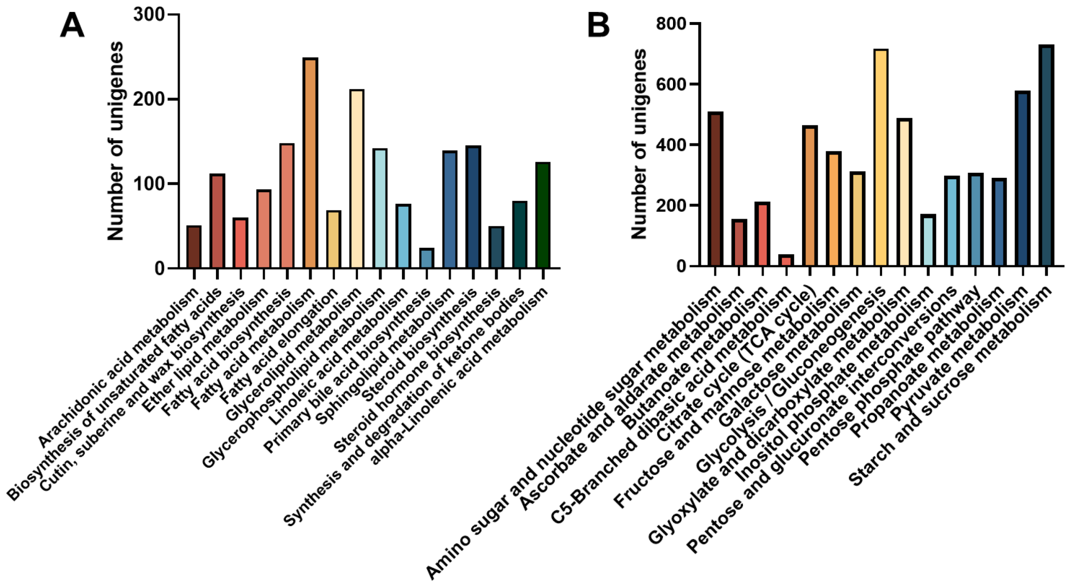


FIGURE 6
Number of unigene classified to carbohydrate and lipid metabolism and pathway. **(A)** Lipid metabolism pathways; **(B)** Carbohydrate metabolism.

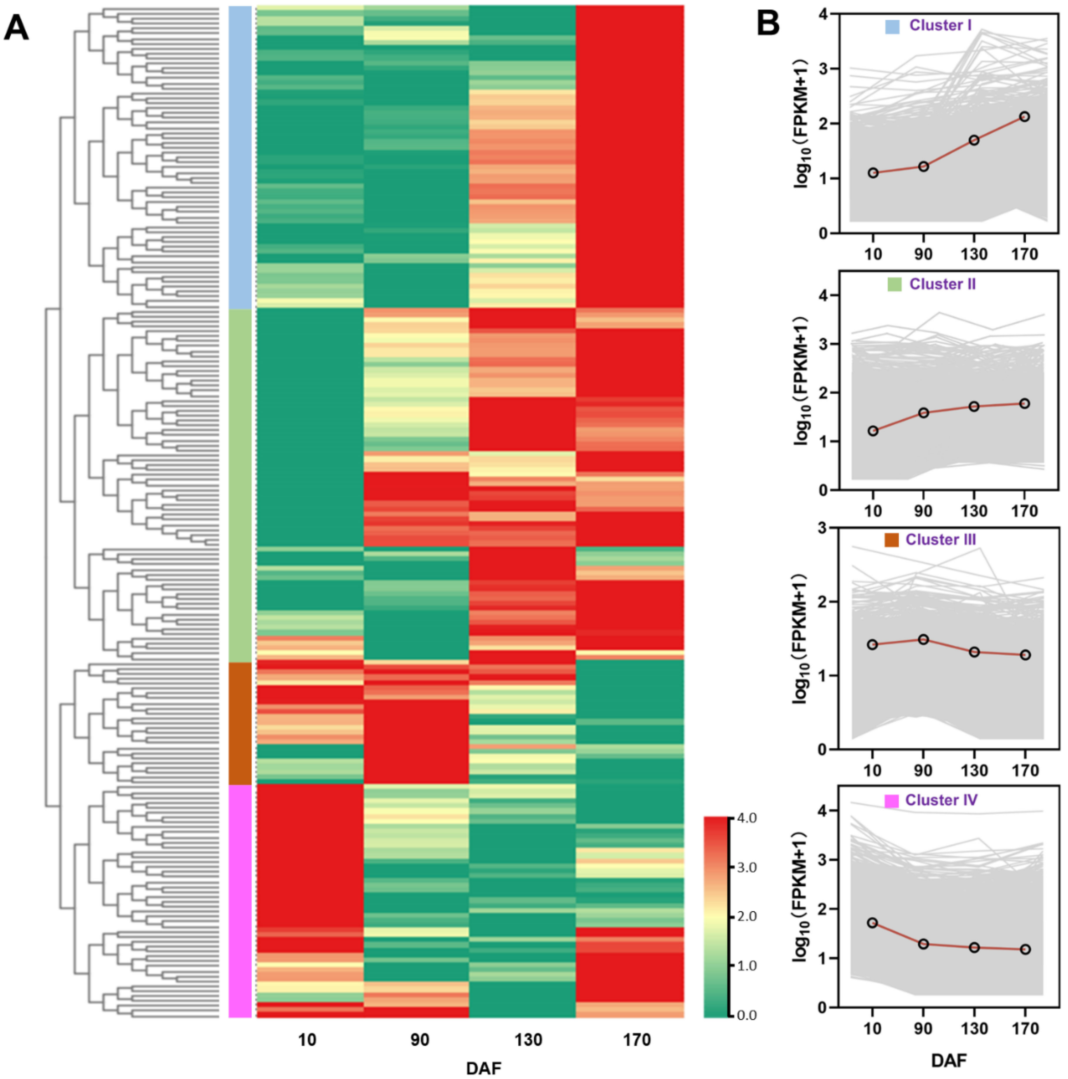


FIGURE 7
The classification of unigenes associated with carbohydrate and lipid metabolism pathways in *S. paniculata*. **(A)** Clustering tree the unigenes; **(B)** The four different expression trends.

that fruit development may derive a small quantity of oil synthesis from the PC pathway. Additionally, Figure 3 Microstructure unveiled that oil in fruits at this stage originates from mesocarp oil emphasizing the crucial role of the developmental PC pathway in facilitating mesocarp oil accumulation in Avocado (*Persea americana*) [33]. In conclusion, during FES stage (10-90 DAF), the up-regulation of major starch synthesis genes stimulates starch accumulation while key genes (*LPCAT* and *PDAT1*) in the PC pathway promote modest amounts of oil accumulation within the mesocarp.

Subsequently, during the ORA stage (90-130 DAF), characterized by rapid oil accumulation (Figure 1), the genes associated with glycolysis and TCA metabolism, such as phosphofructokinase 1 (*PFK-1*), pyruvate kinase (*PK*), and pyruvate dehydrogenase (*PDH*), exhibited up-regulation. *PFK-1* and *PK* served as pivotal enzymes in the two most crucial irreversible reactions of glycolytic metabolism, facilitating increased conversion of glucose into pyruvate. The up-regulation of *PDH* facilitates the swift conversion of pyruvate into acetyl-coA, as well as malonyl-CoA-ACP transacylase (*MAT*), 3-ketoacyl ACP synthase I/II/III (*KAS I/II/III)*, enoyl-ACP reductase I (*EAR*), 3R-hydroxymyristoyl ACP dehydrase (*HAD*), fatty acyl-ACP thioesterase A (*FATA*) and fatty acid desaturase 2 (*FAD2*) enzymes associated with fatty acid and TAG synthesis. Remarkably enhanced expression levels of glycerol-3-phosphate acyltransferase (*GPAT*), phosphatidate phosphatase (*PP*) and diacylglycerol O-acyltransferase 1 (*DGAT1*) genes indicate that the abundant acetyl-coA during this stage primarily serves oil biosynthesis. Enhanced with Figure 3, it was also discovered that an abundance of lipid droplets materialized within the fruit cells during this stage. Furthermore, the gluconeogenesis-associated genes phosphoenolpyruvate carboxylase (*PEPCK*) and pyruvate carboxylase (*PYC*) exhibited down-regulation, while no significant disparity was observed between fructose-1, 6-bisphosphatase (*FBP*) and glucose-6-phosphatase (*G6PC*). The enzyme genes involved in starch synthesis pathway such as SS, SBE, starch phosphorylase

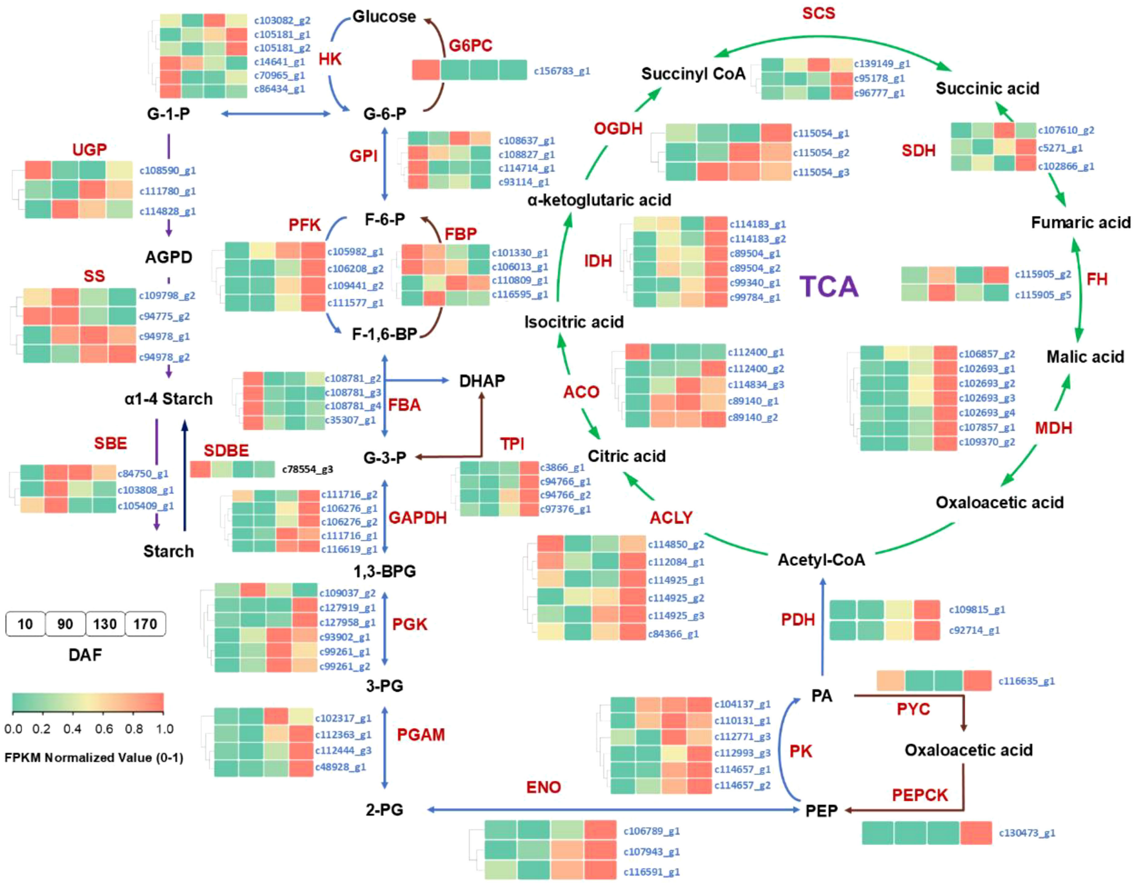


FIGURE 8
The temporal pattern of key enzyme genes' expression in sugars metabolism pathways. The blocks in the figure represent the FPKM normalized value at different stages. From left to right: 10, 90, 130 and 170 DAF; FPKM value of all unigenes were normalized to 0 to 1 for comparison. The identified key enzymes involved in carbohydrate metabolism include Hexokinase (HK); hexose phosphate isomerase (GPI); phosphofructokinase 1 (PFK-1); fructose diphosphate aldulase (FBA); triose phosphate isomerase (TPI); glyceraldehyde-3-phosphate dehydrogenase (GAPDH); phosphoglycerate kinase (PGK); phosphoglycerate mutase (PGAM); enolase (ENO); pyruvate kinase (PK); pyruvate carboxylase (PC); phosphoenolpyruvate carboxylase (PEPCK); fructose-1, 6-bisphosphatase (FBP); glucose-6-phosphatase (G6PC); phosphate glucose mutase (PGM); ADP-glucose pyrophosphorylase (AGP); starch synthase (SS); starch branching enzyme (SBE); starch phosphorylase (PYG); starch debranching enzyme (SDBE); pyruvate dehydrogenase (PDH); citrate synthase (ACLY); aconitate hydratase (ACO); isocitrate dehydrogenase (IDH); ketoglutarate dehydrogenase (OGDH); succinyl-coA synthetase (SCS); succinate dehydrogenase (SDH); fumarate esterase (FH); malate dehydrogenase (MDH); sugar substrates are abbreviated: Glucose-6-phosphate (G-6-P); fructose-6-phosphate (F-6-P); fructose-1,6-diphosphate(F-1,6-BP); glyceraldehyde-3-phosphate (G3P); dihydroxyacetone phosphate (DHAP); 1,3-diphosphoglycerate (1,3-BPG); 3-phosphoglycerate (3-PG); 2-phosphoglycerate (2-PG); enolpyruvate (PEP).

(*PYG*), and phosphate glucose mutase (*PGM*) were found to be down-regulated. This observation signifies that gluconeogenesis and starch synthesis remain inactive at this particular stage, thereby ensuring the seamless progression of glycolysis and facilitating substantial oil synthesis.

During the FDC stage (130-170 DAF), oil accumulation continued, and glycolysis-related genes *PFK-1*, glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), enolase (*ENO*), phosphoglycerate mutase (*PGAM*), *PK* and *PDH* remained highly expressed. Additionally, TCA metabolism-related genes PDH, aconitate hydratase (*ACO*), ketoglutarate dehydrogenase (*OGDH*), citrate synthase (*ACLY*), succinyl-coA synthetase (*SCS*), fumarate esterase (*FH*) and malate dehydrogenase 2 (*MDH2*) exhibited high expression levels. Considering the changes metabolic substance content (Figure 2), that active glycolysis and TCA metabolism are likely the primary factors contributing to the simultaneous rapid decline in soluble sugars and starch content during this FDC stage. The genes associated with starch synthesis, namely *SBE* and *SS* down-regulated while the genes related to starch metabolism, *AGP* and *PGM* enzymes, experienced an up-regulation. of starch degradation, thereby signifying that acetyl-coA generated through glycolysis predominantly enters the tricarboxylic acid cycle. Simultaneously, there was a continuous up-regulation in the expression of *KASIII, KAR, EAR, FATA, FATB, FAD2, DGAT1* and other genes involved in fatty acid carbon chain elongation and desaturation pathway. Consequently, leading to a consistent increase in oil content. However, the tricarboxylic acid cycle activity in this stage may also impede oil growth.

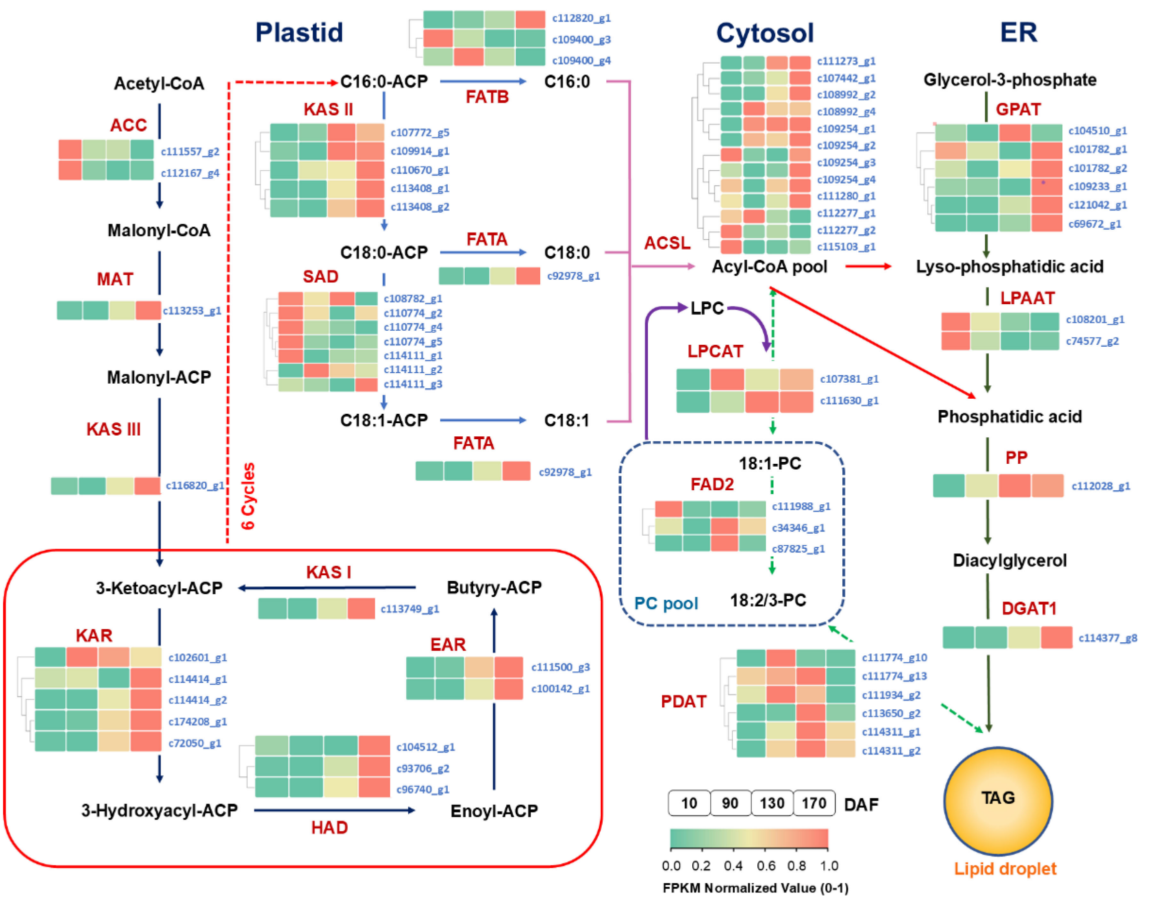


FIGURE 9
The temporal pattern of key enzyme genes' expression in oil biosynthesis pathways. The icons close to each enzyme show the results of FPKM normalized value, from left to right were 10, 90, 130 and 170 DAF. FPKM value of all unigenes were normalized to 0 to 1 for comparison. The identified key enzymes involved in lipid metabolism include acetyl-CoA carboxylase carboxyl transferase (ACC); acyl carrier protein (ACP); Malonyl-CoA-ACP transacylase (MAT); 3-Ketoacyl ACP synthase I,II,III (KASI,II,II); 3-Ketoacyl ACP reductase (KAR); 3R-hydroxymyristoyl ACP dehydrase (HAD); enoyl-ACPreductase I (EAR); fatty acyl-ACP thioesterase A (FATA); fatty acyl-ACP thioesterase B (FATB); Stearoyl-ACP desaturase (SAD); long-chain acyl-CoA synthetase (ACSL); fatty acid desaturase (FAD2); glycerol kinase (GK); glycerol-3-phosphate acyltransferase (GPAT); lysophosphatidyl acyltransferase (LPAAT); phosphatidate phosphatase (PP); diacylglycerol O-acyltransferase (DGAT); phospholipid: diacylglycerol acyltransferase (PDAT); lysophosphatidylcholine acyltransferase (LPCAT). Lipid substrates are abbreviated: palmitic acid (16:0); stearic acid (18:0); oleic acid(18:1); linoleic acid(18:2) and linolenic acid (18:3).

3.6 qRT-PCR validation for sequencing results

To validate the transcriptome sequencing accuracy of *S. paniculata* fruit, this study employed fluorescent quantitative PCR technology to investigate the dynamic changes in relative expression patterns of 12 genes associated with sugar metabolism and fat synthesis (*DGAT1, LPAAT, PADT, SAD, FATA, ACC, PDH, PFK-1, SS, HK, MDH* and *GPI*) (Figure 10). The results demonstrated a consistent dynamic change trend in the relative expression of these 12 genes, as determined by qRT-PCR (ΔΔCt), with that observed through transcriptome sequencing and RNA-Seq sequencing (FPKM). These findings further validate the accuracy and reliability of the transcriptome sequencing expression profile, thereby indicating the feasibility of utilizing RNA-Seq technology for differential gene expression pattern analysis in *S. paniculata*.

4 Discussion

4.1 The influence of sugar metabolism on oil synthesis exhibited diverse patterns across discrete phases of fruit maturation

The sugars present in plant fruits serve as the fundamental basis for oil accumulation, and there exists a close association between carbohydrate metabolism and oil biosynthesis (John and Elspeth, 2003). The investigation on sagebrush (*Salvia splendens* Ker Gawl.) seed development revealed a negative correlation between soluble sugar content and oil content, suggesting the conversion of soluble sugars into oil during seed maturation. Similar findings were observed in olive (*Olea europaea* L.), where an increase in oil content was accompanied by a decrease in sugar levels, indicating the transformation of sugars into oil. The key principle underlying the transformation of oils by sugars is that glycolysis can supply

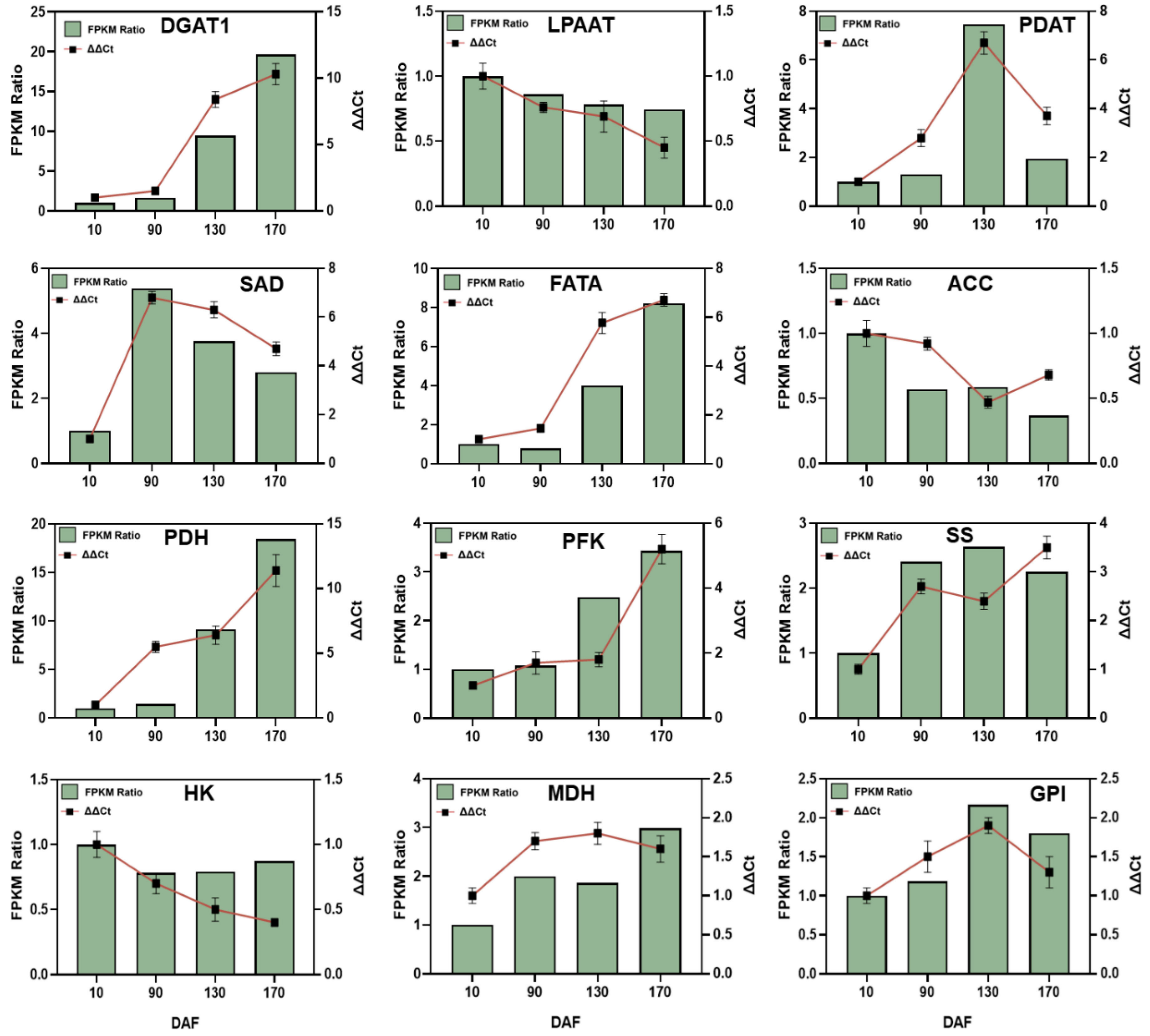


FIGURE 10
The key genes' qRT-PCR validation involved in sugar metabolism and oil biosynthesis of *Symplocos paniculate*. The key genes include diacylglycerol O-acyltransferase (DGAT); lysophosphatidyl acyltransferase (LPAAT); phospholipid: diacylglycerol acyltransferase (PDAT); stearoyl-ACP desaturase (SAD); fatty acyl-ACP thioesterase A (FATA); acetyl-CoA carboxylase carboxyl transferase (ACC); pyruvate dehydrogenase (PDH); phosphofructokinase (PFK); starch synthase (SS); Hexokinase (HK); malate dehydrogenase (MDH); and hexose phosphate isomerase (GPI).

acetyl-CoA, dihydroacetone phosphate (the synthetic precursor of glycerol 3-phosphate), ATP, and NADH for lipid synthesis. Additionally, the TCA primarily generates ATP while the pentose phosphate pathway mainly provides the necessary NADH for fatty acid synthesis. However, this study revealed stage-specific variations and correlations in the sugars and oils' contents of *S. paniculata*, indicating the absence of an absolute mutual transformation relationship between these two metabolites. Similar stepwise relationships have been observed during the developmental stages of other oil seeds (Borisjuk et al., 2004). The variations in physiological functions throughout fruit development stages may account for the observed phenomenon. Sugars are not only closely associated with the accumulation of storage substances such as oil and starch (Zhang et al., 2017), but also play a crucial role in providing nutrients and energy for fruit growth and development (Gambetta et al., 2010).

During the FRE stage (10-90 DAF) of *S. paniculata* fruits, starch accumulation exhibited rapid kinetics, while oil accumulation remained minimal. Notably, no significant correlation was observed between oil content and the levels of starch, sucrose, fructose, or glucose. The strong negative correlation observed between sucrose and starch suggests that sucrose is predominantly converted into starch during this stage, a finding further supported by the microscopic examination of mesocarp cell structures (Figure 3). Additionally, it should be noted that starch serves as a primary substrate for glycolysis in both plastids and avocado (*Persea americana* Mill.) mesocarp (Kilaru et al., 2015). During the ORA stage (90-130 DAF), oil accumulated rapidly, while starch content remained elevated and sucrose and oil content exhibited significant increases. Furthermore, research has demonstrated that during the intermediate phase of fruit development, a substantial amount of starch accumulates to establish a stable carbon reservoir for oil biosynthesis, thereby facilitating oil synthesis (Vigeolas et al., 2004). The absence of this carbon pool function leads to an increase in soluble sugar content, which in turn limits the rate of sugar transfer in leaves due to high sugar levels. Consequently, during the rapid synthesis of oil, the sustainability of glycolysis process is compromised and there is insufficient carbon source available to maintain optimal activity of related enzymes, resulting in a significant reduction in oil content oilseed rape (*Brassica napus* L.) (Silva et al., 1997). During the FDC period (130-170 DAF), there is a rapid decrease in starch, sucrose, fructose, and glucose levels, which promotes stable oil synthesis through sugar decomposition (Norton and Harris, 1975). Additionally, a study on *X. sorbifolium* fruit revealed a negative correlation between sucrose changes and oil content during the later stages of fruit development, indicating that sucrose decomposition facilitates oil synthesis (Zhang and Liu, 2016).

4.2 The pivotal genes within the sugar metabolism pathway of *S. paniculata* fruit exert a profound influence on the synthesis of oil

The synergistic effects between other pathways and lipid anabolic pathways are often overlooked in the study of oil plant biosynthetic pathways. Further investigation is warranted to explore the regulation of enzyme activities related to lipid synthesis, while also considering the contribution of reaction substrates produced by the glucose metabolism pathway. Additionally, prior research has demonstrated a close correlation between glucose metabolism and oil synthesis, with the manipulation of genes associated with glucose metabolism resulting in significant enhancements to seed oil accumulation (Allen et al., 2011). In the investigation of *Arabidopsis thaliana* (L.) Heynh., a model plant, over 800 enzyme genes associated with lipid anabolism were identified, including a substantial number of key genes involved in glycolysis and the TCA pathway (Beisson et al., 2003). The regulatory role of carbon circulation, starch metabolism, and other pathways in oil synthesis was highlighted in the gene expression profile analysis of *Jatropha curcas* seed development (Jiang et al., 2012). The enhancement of oil palm (*Elaeis guineensis* Jacq.) oil is not only associated with the upregulation of fatty acid synthase transcription, but also linked to specific plastid transport factors and key enzymes involved in plastid glucose metabolism, such as phosphofructokinase, pyruvate kinase, and pyruvate dehydrogenase (Bourgis et al., 2011).

In this study, we observed an upregulation in the expression of starch biosynthesis-related enzyme genes during FRE stage (10-90 DAF), including *AGP* (1.5-fold), *SS* (2.4-fold), and *SBE* (2.1-fold), among the key enzymes involved in carbohydrate metabolism pathway. Starch also serves as a primary substrate for glycolysis in plastids, and within the avocado mesocarp, transcripts of genes related to starch synthesis and degradation are abundantly present throughout mesocarp development (Kilaru et al., 2015). Theoretically, starch synthesis reduces the content of monosaccharides which required for glycolysis and would construct an inescapable competition with oil biosynthesis (Bettey and Smith, 1990), but it is found that the synthesis of starch does not restrict but promotes the biosynthesis of oil (Norton and Harris, 1975). Other studies have also found that restricted starch synthesis slowed down or even reduced the accumulation of seed oil. For example, the *PGM* is a key enzyme in starch anabolism, which catalyzed the reversible transformation between glucose-1-phosphate (G-1-P) and glucose-6-phosphate (G-6-P), a noticeable 40% reduction of oil content in Arabidopsis mutant plant with *PGM* gene knocked compare with control (wild type) (Periappuram et al., 2000). Vigeolas et al. reported that the embryo-specific reduction of *AGP* which catalyzed first step of starch synthesis leads to an inhibition of starch synthesis and a delay in oil accumulation in developing seeds of oilseed rape (*Brassica napus*) (Vigeolas et al., 2004). Therefore, these findings suggest that a substantial accumulation of starch occurs during the development of *S. paniculata* fruit, serving as a carbon reservoir for oil synthesis and facilitating enhanced sucrose transport from leaves to fruits, thereby promoting subsequent fruit oil biosynthesis.

A substantial flux through the glycolytic pathway is anticipated to supply the requisite quantities of pyruvate essential for efficient oil synthesis in oil plants. In this investigation, we observed a consistent upregulation of *PFK-1*, *PK*, and *PDH* enzyme genes involved in the glycolysis pathway during the fruit's ORA stage (90-170 DAF). *PFK-1* and *PK* serve as pivotal enzymes in the two most crucial irreversible reactions within the glycolytic metabolic pathway. The upregulation of *PFK-1* and *PK* facilitates the progression of glycolysis, leading to enhanced conversion of glucose into pyruvate (Chapman and Ohlrogge, 2012). Meanwhile, upregulating *PDH* can expedite the swift transformation of pyruvate into acetyl-CoA, a common precursor for both fatty acid synthesis and TCA cycle, thereby promoting fatty acid biosynthesis (Norton and Harris, 1975). The high oil content in the oil palm fruit is not solely attributed to increased gene expression for enzymes involved in TAG synthesis, but rather exhibits a stronger correlation with plastid glycolytic enzymes (*PFK-1, PK* and *PDH*). The cytoplasmic *PFK-1*, an ATP-dependent enzyme, exhibited a 3.4-fold higher abundance in oil palm mesocarp compared to date palm (a variety primarily accumulating sugar), and demonstrated increased levels during the ripening process of oil palm mesocarp (Bourgis et al., 2011). Therefore, our findings suggest that an upregulation in glycolytic expression plays a pivotal role in facilitating pyruvate provision for the efficient synthesis of fatty acids at high rates.

The TCA serves as a critical link between sugar metabolism and oil biosynthesis. Pyruvate, generated through glycolysis, functions as a shared intermediate in both the oil biosynthesis and TCA cycle pathways. Concurrently, during lipid breakdown, fatty acids undergo β-oxidation while glycerol undergoes the Embden-Meyerhof-Parnas (EMP) pathway to generate acetyl CoA, which subsequently enters the tricarboxylic acid cycle for complete oxidation. In this study, the TCA-related enzyme genes, including *PDH, ACO, OGDH, ACLY, IDH, SCS, FH* and *MDH2*, exhibited significantly elevated expression levels during the FDC stage (150-170 DAF). The possible explanation lies in the sharp increase of ripening hormones, such as ethylene, upon fruit maturation, which subsequently impacts the activities of respiratory enzymes and triggers the respiratory climacteric phenomenon (Chapman and Ohlrogge, 2012). The hydrolysis of a substantial quantity of starch and sucrose into glucose for entry into glycolysis results in a rapid reduction in the levels of starch and sucrose. Simultaneously, the upregulation of TCA-related enzyme genes results in enhanced acetyl-CoA production through glycolysis, predominantly fueling the tricarboxylic acid cycle. However, this metabolic shift may impede oil content growth despite its increase during this stage. Previous studies have demonstrated that the addition of hydrogen peroxide (H_2O_2) and malonic acid ($C_3H_4O_4$) during *Chlorella vulgaris* cultivation effectively inhibits intracellular TCA cycle activity, resulting in a significant enhancement of oil content by 34.1% and 28.3%, respectively (Chu et al., 2019). Therefore, modulation of TCA metabolic-related enzyme expression in fruits or reduction of endogenous ethylene levels to attenuate the respiratory climacteric during fruit ripening could potentially serve as a novel approach for augmenting oil synthesis in fruits.

4.3 Analysis of the expression patterns of key genes in the lipid synthesis pathway

In the oil synthesis pathway, acetyl-CoA participates in glyceride biosynthesis via soluble proteins or membrane-bound transfer proteins (Yin et al., 2007). The down-regulation of *ACC* was observed in this study, with its expression peaking during the early stage of fruit development. Notably, similar patterns of *ACC* enzyme gene expression were reported in studies on *Prunus sibirica* L (Niu et al., 2015). and *Vernicia fordii* (Hemsl.) Airy Shaw (Chen et al., 2013), corroborating our findings. Relevant studies have demonstrated that the concentration of oleic acid inhibits ACC enzyme activity. Andre et al. (2012). introduced oleic acid into a suspension of European rape liquid culture cells, resulting in the inhibition of ACC enzyme activity in plastid and subsequent reduction in fatty acid synthesis. However, oleic acid constitutes the predominant fatty acid in *S. paniculata* fruit oil. It is worth investigating whether this phenomenon observed during the synthesis of *S. paniculata* fruit oil is associated with the inhibitory effect of oleic acid on ACC enzyme gene expression. Additionally, the transfer of the *ACC* gene into algae for enhancing oil production did not yield any discernible effect (Dunahay et al., 1995). Furthermore, it has been established that the activity of the ACC enzyme can be influenced by the availability of reaction substrates generated through glucose metabolism.

FATA/B primarily functions in the hydrolysis of fatty acyl ACP, facilitating the liberation of free fatty acids. Specifically, *FATA* predominantly catalyzes the release of stearoyl ACP and oleoyl ACP (C18:0/1-ACP) to yield stearic acid and oleic acid (C18:0/1), respectively, while FATB mainly mediates the cleavage of palmitoyl ACP (C16:0-ACP) into palmitic acid (C16:0). In this study, we observed a consistent expression of *FATA* from 10 to 90 DAF stage, followed by a subsequent increase from 90 to 170 DAF stage, accompanied by a significant elevation in oleic acid content. While the up-regulation of *FATB* during fruit development stage resulted in a higher palmitic acid content compared to stearic and oleic acid contents at FRE (10-90 DAF) stage. The over-expression of *FATB* in seeds has been shown to result in a four-fold increase in palmitic acid fatty acids in *Arabidopsis* seeds (Dörmann et al., 2000), while disrupting the expression of *FATB* leads to a significant reduction in palmitic acid content, consequently reducing the overall saturated fatty acid content (Bonaventure et al., 2003). Therefore, *FATA/B* would serve as a potential candidate gene for the regulation of fatty acid composition, thereby offering the possibility to genetically modify the relative content of C16 and C18 fatty acids in *S. paniculata* fruits.

The desaturation of fatty acids in the fruit of *S. paniculata* involves two key enzymes. One enzyme, *SAD*, facilitates the conversion of C18:0-ACP to C18:1-ACP within plastids. The other enzyme is present in the cell membrane, endoplasmic reticulum, and chloroplast, and catalyzes the introduction of a second double bond at specific positions along the carbon chain of mono-unsaturated fatty acids. For instance, *FAD2* mediates the desaturation process from C18:1-ACP to generate C18:2-ACP (Tasaka et al., 1996). The expression of the *SAD* gene was observed to be up-regulated by 5-fold at the 10 DAF stage, followed by a sustained and high expression from the 90 to 170 DAF stages. This ensured the timely conversion of stearic acid into oleic acid and facilitated a rapid increase in oleic acid content. The expression of *FAD2* was initially down-regulated and subsequently up-regulated, which played a pivotal role in the biosynthesis of linoleic acid during the late FDC phase (130-170 DAF) of oil accumulation. Modulating the expression of crucial target genes involved in fatty acid desaturation through RNA interference technology is a widely employed approach in genetic engineering for regulating the level of unsaturation in oil plant seeds (Knutzon et al., 1992). In studies on *B. napus* (Knutzon et al., 1992) and *Gossypium hirsutum* L (Liu et al., 2002), silencing the *SAD* gene expression resulted in a significant 38% increase in stearic acid content in seeds of both plants. Similarly, genetic interference targeting the *FAD_2* gene led to substantial increases in oleic acid content in seeds of *Brassica napus* L. (26%), *Brassica juncea* (L.) Czern. (28%) (Stoutjesdjik et al., 2000), soybean (*Glycine max* (L.) Merr.) (76%) (Wang et al., 2012), and *G. hirsutum* (62%) (Liu et al., 2002). In the context of industrial production and application, biodiesel and edible oil production raw materials have distinct fatty acid composition requirements. An optimal biodiesel should exhibit a high abundance of monounsaturated fatty acids (Wang et al., 2012). Therefore, the incorporation of *SAD* and *FAD2* genes presents a promising avenue for plant genetic modification, enabling precise modulation of *S. paniculata*'s unsaturation degree to cater to diverse production requirements in edible oil and biodiesel sectors.

GPAT is the key rate-limiting enzyme that catalyzes the initial step of TAG anabolic pathway, thereby playing a pivotal role in regulating oil synthesis in plants (Shockey et al., 2016). The observed upregulation of the *GPAT* gene expression between 130 and 170 DAF aligns with the trend for rapid increase in oil content during this developmental stage. To enhance oil content in plant varieties, the expression of *GPAT* gene was modulated through genetic engineering techniques (Liu et al., 2012). The safflower plastid *GPAT* gene and *GPAT* gene of *Escherichia coli* were introduced into Arabidopsis plants for targeted expression, resulting in a significant increase of 22% and 15%, respectively, in seed oil content (Jain et al., 2000). Thus, reinforcing the expression

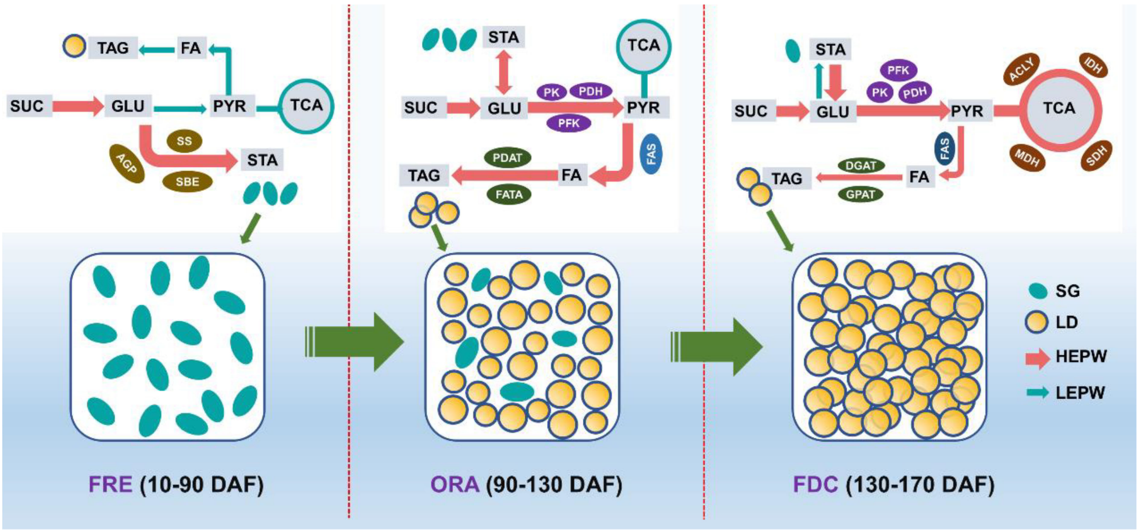


FIGURE 11
The schematic diagram of the key genes' temporal expressional patterns involved in the sugar and oil pathways. The width of each arrow is strictly proportional to the level of gene expression in related pathways. HEPW: the high expressional pathway; LEPW: the low expressional pathway; LD: lipid droplets; SG: Starch granules; FES: Fruit rapid expansion stage; ORA: Oil rapid accumulation stage; FDC: Fruit dis-coloration stage; SUC: sucrose; GLU: glucose; PYR: pyruvic acid; FA: fatty acid; TCA: tricarboxylic acid cycle; TAG: triacylglycerol; STA: starch; AGP: ADP-glucose pyrophosphorylase; SS: starch synthase; SBE: starch branching enzyme; PDH: pyruvate dehydrogenase; PK: pyruvate kinase; PFK: phosphofructokinase; FATA: fatty acyl-ACP thioesterase A; DGAT:diacylglycerol O-acyltransferase; PDAT: phospholipid: diacylglycerol acyltransferase; GPAT: glycerol-3-phosphate acyltransferase; MDH: malate dehydrogenase; ACLY: citrate synthase; IDH: isocitrate dehydrogenase; SDH: succinate dehydrogenase.

of *GPAT* enzyme gene can effectively enhance the oil content of *S. paniculata*.

The *PDAT1* and *DGAT1* genes play pivotal roles in the biosynthesis of TAGs. These two genes exhibit overlapping functions and mutually complement each other [45]. The expression of *PDAT1* was observed to be upregulated at FRE stage (10-90 DAF), followed by downregulation at FDC stage (130-170 DAF). These findings suggest that the PC pathway may contribute to a minor fraction of oil synthesis during early fruit development. The microstructure depicted in Figure 3 also reveals that the presence of oil in fruits at this stage originates from mesocarp oil cells, highlighting the crucial role of the developmental PC pathway in facilitating mesocarp oil accumulation in Avocado. The *DGAT* gene is more likely to enhance oil content in oil seeds (Kilaru et al., 2015). In this study, we observed a consistent up-regulation of the expression of the *DGAT1* enzyme gene from 90 DAF to 170 DAF. Depletion of the *DGAT1* gene resulted in a significant reduction in oil content in Arabidopsis (*Arabidopsis thaliana*) (Jako et al., 2001), maize *(Zea mays* L.*)* (Zheng et al., 2008), and soybean (*Glycine max*) seeds (Lardizabal et al., 2008).

5 Conclusion

In this study, we conducted a comprehensive quantitative analysis and employed advanced cell microscopic observation techniques to investigate the dynamic changes in sugar and oil content during different developmental stages of *S. paniculata* fruit. Additionally, we performed a transcriptome analysis to unravel the intricate gene regulatory network underlying these metabolic processes (Figure 11). During the FRE stage (10-90 DAF) of *S. paniculata* fruit development, genes *AGP*, *SS*, and *SBE* were upregulated to enhance starch accumulation, a process that involves the conversion of soluble sugars (sucrose, fructose, and glucose) into starch. Concurrently, we observed limited oil accumulation in the mesocarp oil cells, which we attribute in part to the upregulation of the *LPCAT* gene within the PC pathway. Transitioning into the ORA stage (90-130 DAF), we noted high glucose levels serving as a "carbon pool" for oil synthesis. During this period, the upregulation of glycolytic genes PK, PFK-1, and PDH facilitated rapid oil accumulation. Furthermore, genes associated with oil synthesis pathways (Fatty acid biosynthesis and TAG assembly) were generally upregulated, further promoting oil synthesis. In this stage, we observed substantial accumulation of oil bodies in both mesocarp and seed cells. By the FDC stage (130-170 DAF), oil accumulation gradually increased to 36% under the regulation of fatty acid synthesis (*KASIII, KAR, EAR, FATA, FATB, FAD2*) and TAG assembly (*DGAT1*) pathwasys. It is noteworthy that the upregulation of TCA cycle genes during this stage could potentially decelerate the rate of oil accumulation. The study has successfully achieved a comprehensive understanding of lipid synthesis and examined key regulatory genes, thereby laying a solid theoretical foundation for the rational and efficient development of *S. paniculata* resources.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, SRA357712.

Author contributions

WZ: Formal analysis, Investigation, Writing – original draft, Writing – review & editing. BX: Investigation, Visualization, Writing – original draft. YC: Software, Visualization, Writing – review & editing. JC: Formal analysis, Methodology, Project administration, Writing – review & editing. PL: Project administration, Resources, Software, Writing – review & editing. LJ: Resources, Supervision, Writing – review & editing. CL: Resources, Supervision, Writing – review & editing. QL: Conceptualization, Supervision, Writing – review & editing. YY: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was financially supported by the Key Scientific and Technological Innovation Platform of Hunan Province (2024PT1001), the Forestry Science and Technology Research and Innovation Project (XLKY202310), the Innovation Demonstration Project of ChenZhou City (2022sfq53), the natural Science Foundation of Changsha City (kq2208098), the national Natural Science Foundation of China (32401626); the natural Science Foundation of Hunan Province (2024JJ7283), the State Key Laboratory of Utilization of Woody Oil Resource (GZKF202202).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1452533/full#supplementary-material

References

Allen, S. M., Damude, H. G., Everard, J. D., Meyer, K., and Yoo, B. C. 2011*Sucrose transporter genes for increasing plant seed lipids*. Available online at: https://pubchem.ncbi.nlm.nih.gov/patent/US-2018245092-A1 (Accessed 19 January 2024).

Andre, C., Haslam, R. P., and Shanklin, J. (2012). Feedback regulation of plastidic acetyl-CoA carboxylase by 18,1-acyl carrier protein in *Brassica napus*. Proc. *Natl. Acad. Sci. U.S.A.* 109, 10107–10112. doi: 10.1073/pnas.1204604109

Beisson, F., Koo, A. J. K., Ruuska, S., Schwender, J., Pollard, M., Thelen, J. J., et al. (2003). Arabidopsis genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. *Plant Physiol.* 132, 681–697. doi: 10.1104/pp.103.022988

Bettey, M., and Smith, A. M. (1990). Nature of the effect of the r locus on the lipid content of embryos of peas (*Pisum sativum* L.). *Planta.* 180, 420–428. doi: 10.1007/BF00198795

Bonaventure, G., Salas, J. J., Pollard, M. R., and Ohlrogge, J. B. (2003). Disruption of the FATB gene in Arabidopsis demonstrates an essential role of saturated fatty acids in plant growth. *Plant Cell.* 15, 1020–1033. doi: 10.1105/tpc.008946

Borek, S., Pukacka, S., and Michalski, K. (2009). Lipid and protein accumulation in developing seeds of three lupine species, *Lupinus luteus* L, *Lupius albus* L, and *Lupinus mutabilis* sweet. *J. Exp. Bot.* 60, 3353–3366. doi: 10.1093/jxb/erp186

Borisjuk, L., Rolletschek, H., Radchuk, R., Weschke, W., Wobus, U., and Weber, H. (2004). Seed development and differentiation, a role for metabolic regulation. *Plant Biol.* 6, 375–386. doi: 10.1055/s-2004-817908

Bourgis, F., Kilaru, A., Cao, X., Ngando-Ebongue, G. F., Drira, N., Ohlrogge, J. B., et al. (2011). Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12527–12532. doi: 10.1073/pnas.1106502108

Chapman, K. D., and Ohlrogge, J. B. (2012). Compartmentation of triacylglycerol accumulation in plants. *J. Biol. Chem.* 287, 2288–2294. doi: 10.1074/jbc.R111.290072

Chen, H., Jiang, G. X., Long, H. X., and Tan, X. F. (2013). Analysis of oil synthesis metabolism pathways based on transcriptome changes in tung oil tree's seeds during three different development stages. *Hereditas.* 35, 1403–1414. doi: 10.3724/SP.J.1005.2013.01403

Chu, F. F., Cheng, J., Zhang, X. D., Ye, Q., Chen, S. T., Zhou, J. H., et al. (2019). Transcriptome and key gene expression related to carbon metabolism and fatty acid synthesis of *Chlorella vulgaris* under a nitrogen starvation and phosphorus repletion regime. *J. Appl. Phycol.* 31, 2881–2893. doi: 10.1007/s10811-019-01811-y

Conesa, A., Gotz, S., Garcia-Gornez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO, a universal tool for annotation, visualization and analysis of functional genomics research. *Bioinformatics.* 219, 3674–3676. doi: 10.1093/bioinformatics/bti610

Dörmann, P., Voelker, T. A., and Ohlrogge, J. B. (2000). Accumulation of palmitate in Arabidopsis mediated by the acyl-acyl carrier protein thioesterase FATB1. *Plant Physiol.* 123, 637–644. doi: 10.1104/pp.123.2.637

Dunahay, T. G., Jarvis, E. E., and Roessler, P. G. (1995). Genetic transformation of the diatoms *Cyclotella cryptica* and *Navicula saprophila*. J. *Phycol.* 31, 1004–1012. doi: 10.1111/j.0022-3646.1995.01004.x

Er, Q., Da, X. P., and Liu, G. Q. (2000). Sudan III/IV staining for simple frozen section. *Chin. J. Forensic Med.* 15, 229. doi: 10.13618/j.issn.1001-5728.2000.04.020

Fu, X., Jin, Y., Paul, M. J., Yuan, M., Liang, X., Cui, R., et al. (2023). Inhibition of rice germination by ustiloxin A involves alteration in carbon metabolism and amino acid utilization. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1168985

Gambetta, G. A., Matthews, M. A., Shaghasi, T. H., McElrone, A. J., and Castellarin, S. D. (2010). Sugar and abscisic acid signaling orthologs are activated at the onset of ripening in grape. *Planta.* 232, 219–234. doi: 10.1007/s00425-010-1165-2

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Hill, L. M. (2003). Metabolism of sugars in the endosperm of developing seeds of oilseed rape. *Plant Physiol.* 131, 228–236. doi: 10.1104/pp.010868

Jain, R., Coffey, M., Lai, K., Kumar, A., and MacKenzie, S. (2000). Enhancement of seed oil content by expression of glycerol-3-phosphate acyltransferase genes. *Biochem. Soc Trans.* 28, 958–961. doi: 10.1042/BST0280958

Jako, C., Kumar, A., Wei, Y., Zou, J., Barton, D. L., Giblin, E. M., et al. (2001). Seed-specific overexpression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol.* 126, 861–874. doi: 10.1104/pp.126.2.861

Jiang, H., Wu, P., and Zhang, S. (2012). Global analysis of gene expression profiles in developing physic nut (*Jatropha curcas* L.) seeds. *PloS One* 7, e36522. doi: 10.1371/journal.pone.0036522

John, E. L., and Elspeth, M. R. (2003). New complexities in the synthesis of sucrose. *Curr. Opin. Plant Biol.* 6, 208–214. doi: 10.1016/s1369-5266(03)00033-5

Kanehisa, M., and Goto, S. (2000). KEGG, kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Kennedy, Y., Yokoi, S., Daimon, H., Daimon, H., Nishida, I., and Takahata, Y. (2011). Genetic variation of storage compounds and seed weight in rapeseed (*Brassica napus* L.) germplasms. *Breed. Sci.* 61, 311–315. doi: 10.1270/jsbbs.61.311

Kilaru, A., Cao, X., Dabbs, P. B., Sung, H. J., Rahman, M. M., Thrower, N., et al. (2015). Oil biosynthesis in a basal angiosperm, transcriptome analysis of *Persea americana* mesocarp. *BMC Plant Biol.* 15, 203. doi: 10.1186/s12870-015-0586-2

Knutzon, D. S., Thompson, G. A., Radke, S. E., Johnson, W. B., Knauf, V. C., and Kridl, J. C. (1992). Modification of Brassica seed oil by antisense expression of a stearoylacyl carrier protein desaturase gene. *Proc. Natl. Acad. Sci. U.S.A.* 89, 2624–2628. doi: 10.1073/pnas.89.7.2624

Lardizabal, K., Effertz, R., Levering, C., Mai, J., Pedroso, M. C., Jury, T., et al. (2008). Expression of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean. *Plant Physiol.* 148, 89–96. doi: 10.1104/pp.108.123042

Li, W. J., Jiang, L. J., Chen, Y. Z., Li, C. Z., Li, P. W., Yang, Y., et al. (2023). Transcriptome analysis unveiled the intricate interplay between sugar metabolism and lipid biosynthesis in Symplocos paniculate fruit. *Plants (Basel).* 12, 2703. doi: 10.3390/plants12142703

Lin, Z., An, J., Wang, J., Niu, J., Ma, C., Wang, L. B., et al. (2017). Integrated analysis of 454 and Illumina transcriptomic sequencing characterizes carbon flux and energy source for fatty acid synthesis in developing *Lindera glauca*, fruits for woody biodiesel. *Biotechnol. Biofuels.* 10, 134. doi: 10.1186/s13068-017-0820-2

Liu, Q., Li, C. Z., Jiang, L. J., Li, H., Chen, J. Z., and Yi, X. Y. (2015). The oil accumulation of oil plant *Symplocos paniculata. J. Biobased Materials Bioenergy* 9, 1–5. doi: 10.1166/JBMB.2015.1490

Liu, Q., Siloto, R. M. P., Lehner, R., Stone, S. J., and Weselake, R. J. (2012). Acyl-CoA, diacylglycerol acyltransferase, molecular biology, biochemistry and biotechnology. *Prog. Lipid Res.* 51, 350–377. doi: 10.1016/j.plipres.2012.06.001

Liu, Q., Singh, S. P., and Green, A. G. (2002). High-stearic and high-oleic cotton seed oils produced by hairpin RNA-mediated post-transcriptional gene silencing. *Plant Physiol.* 129, 1732–1743. doi: 10.1104/pp.001933

Liu, Q., Sun, Y. P., Chen, J. C., Li, P. W., Li, C. Z., Niu, G. H., et al. (2016). Transcriptome analysis revealed the dynamic oil Accumulation in *Symplocos paniculata* fruit. *BMC Genomics* 17, 929. doi: 10.1186/s12864-016-3275-0

Liu, H., Wang, C., and Komatsu, S. (2013). Proteomic analysis of the seed development in *Jatropha curcas*, from carbon flux to the lipid accumulation. *J. Proteomics.* 91, 23–40. doi: 10.1016/j.jprot.2013.06.030

Niu, J., An, J., Wang, L., Fang, C., Ha, D., Fu, C., et al. (2015). Transcriptomic analysis revealed the mechanism of oil dynamic accumulation during developing Siberian apricot (*Prunus sibirica* L.) seed kernels for the development of woody biodiesel. *Biotechnol. Biofuels.* 8, 29. doi: 10.1186/s13068-015-0213-3

Norton, G., and Harris, J. F. (1975). Compositional changes in developing rape seed (*Brassica napus* L.). *Planta.* 123, 163–174. doi: 10.1007/BF00383865

Periappuram, C., Steinhauer, L., Barton, D. L., Taylor, D. C., Chatson, B., and Zou, J. (2000). The plastidic phosphogluco-mutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control. *Plant Physiol.* 122, 1193–1199. doi: 10.1104/pp.122.4.1193

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Shockey, J., Regmi, A., Cotton, K., Adhikari, N., Browse, J., and Bates, P. D. (2016). Identification of Arabidopsis GPAT9 (At5g60620) as an essential gene involved in triacylglycerol biosynthesis. *Plant Physiol.* 170, 163–179. doi: 10.1104/pp.15.01563

Silva, P.M.F.R.D., Eastmond, P. J., Hill, L. M., Smith, A. M., and Rawsthorne, S. (1997). Starch metabolism in developing embryos of oilseed rape. *Planta.* 203, 480–487. doi: 10.1007/s004250050217

Stoutjesdijk, P. A., Hurlstone, C., Singh, S. P., and Green, A. G. (2000). High oleic Australian Brassica napus and B. juncea varieties produced by co-suppression of endogenous 12-desaturases. Biochem. Soc. *Transl.* 28, 938–940. doi: 10.1042/BST0280938

Tasaka, Y., Gombos, Z., Nishiyama, Y., Mohanty, P., Ohba, T., Ohki, K., et al. (1996). Targeted mutagenesis of acyl-lipid desaturases in Synechocystis, evidence for the important roles of poly-unsaturated membrane lipids in growth, respiration and photosynthesis. *EMBO J.* 15, 6416–6425. doi: 10.1002/j.1460-2075

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., et al. (2003). The COG database, an updated version includes eukaryotes. *BMC Bioinf.* 4, 41. doi: 10.1186/1471-2105-4-41

Vigeolas, H., Möhlmann, T., Martini, N., Neuhaus, H. E., and Geigenberger, P. (2004). Embryo-specific reduction of ADP-Glc pyrophosphorylase leads to an inhibition of starch synthesis and a delay in oil accumulation in developing seeds of oilseed rape. *Plant Physiol.* 136, 2676–2686. doi: 10.1104/pp.104.046854

Wang, X. K., and Guo, Q. M. (2021). Applications of cryo-sectioning in higher plants. *Plant Physiol. J.* 57, 1047–1054. doi: 10.13592/j.cnki.ppj.2020.0622

Wang, L. D., Yu, H. Y., He, X. H., and Liu, R. Y. (2012). Influence of fatty acid composition of woody biodiesel plants on the fuel properties. *J. Fuel Chem. Technology.* 40, 397–404. doi: 10.1016/S1872-5813(12)60018-8

Weichert, N., Saalbach, I., Weichert, H., Kohl, S., Erban, A., Kopka, J., et al. (2010). Increasing sucrose uptake capacity of wheat grains stimulates storage protein synthesis. *Plant Physiol.* 152, 698–710. doi: 10.1104/pp.109.150854

Xu, C. C., and Shanklin, J. (2016). Triacylglycerol metabolism, function, and accumulation in plant vegetative tissues. *Annu. Rev. Plant Biol.* 67, 179–206. doi: 10.1146/annurev-arplant-043015-111641

Yin, D., Deng, S., Zhan, K., and Cui, D. (2007). High-oleic peanut oils produced by HpRNA mediated gene silencing of oleate desaturase. *Plant Mol. Biol. Rep.* 25, 154–163. doi: 10.1007/s11105-007-0017-0

Zhang, Y., and Liu, A. Z. (2016). The correlation between soluble carbohydrate metabolism and lipid accumulation in castor seeds. *Biotechnol. Bull.* 32, 120–129. doi: 10.13560/j.cnki.biotech.bull.1985.2016.06.017

Zhang, K., Wu, Z. D., Tang, D. B., Luo, K., Lu, H. X., Liu, Y. Y., et al. (2017). Comparative transcriptome analysis reveals critical function of sucrose metabolism related-enzymes in starch accumulation in the storage root of sweet potato. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00914

Zhao, N., Zhang, Y., Wang, J., Liu, X., Zhao, C. G., and Guo, H. H. (2015). Seed development, lipid accumulation and its relationship with carbohydrates and protein in *Xanthoceras sorbifolia* bunge. *Bull. Botanical Res.* 35, 133–140. doi: 10.7525/j.issn.1673-5102.2015.01.020

Zheng, P., Allen, W. B., Roesler, K., Williams, M. E., Zhang, S., Li, J., et al. (2008). A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat. Genet.* 40, 367–372. doi: 10.1038/ng.85

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 04 March 2025
DOI 10.3389/fpls.2025.1525226

EDITED BY
Xin Fang,
Chinese Academy of Sciences (CAS), China

REVIEWED BY
Qian Song,
Guangxi Academy of Agricultural Science, China
Caowen Sun,
Nanjing Forestry University, China
Fangyan Chen,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Duo Chen
chenduo@fjnu.edu.cn
Ting Xue
xueting@fjnu.edu.cn

RECEIVED 09 November 2024
ACCEPTED 07 February 2025
PUBLISHED 04 March 2025

CITATION
Chen D, Xiao Y, Zheng X, Sun H, Zhang C, Zhu J and Xue T (2025) Seasonal dynamics and molecular regulation of flavonoid biosynthesis in *Cyclocarya paliurus* (Batal.) Iljinsk.
Front. Plant Sci. 16:1525226.
doi: 10.3389/fpls.2025.1525226

Seasonal dynamics and molecular regulation of flavonoid biosynthesis in *Cyclocarya paliurus* (Batal.) Iljinsk

Duo Chen*, Yixin Xiao, Xuehai Zheng, Huamiao Sun, Cifeng Zhang, Jinmao Zhu and Ting Xue*

The Public Service Platform for Industrialization Development Technology of Marine Biological Medicine and Products of the Department of Natural Resources, Fujian Key Laboratory of Special Marine Bioresource Sustainable Utilization, Southern Institute of Oceanography, College of Life Sciences, Fujian Normal University, Fuzhou, China

Introduction: *Cyclocarya paliurus*, an economically important species known for its high flavonoid content, has potential for industrial applications. Understanding the seasonal dynamics and molecular regulation of flavonoid biosynthesis in this species is crucial for optimizing its production.

Methods: We conducted an integrated analysis of transcriptomic and metabolomic data to identify key genes involved in flavonoid biosynthesis and regulation. Seasonal variation in flavonoid content and gene expression was examined, with a focus on the genes involved in the flavonoid synthesis pathway and their correlation with flavonoid levels.

Results: Flavonoid content peaked in August and declined towards November, with quercetin and kaempferol glycosides being the most abundant compounds. Pearson correlation analysis revealed significant relationships between the functional genes of the flavonoid synthesis pathway and flavonoid content. Seasonal variations in the expression of key biosynthetic genes (*CHS, CHI, F3H, DFR, FLS*) and regulatory transcription factors (*MYB11, MYB12, MYB111, MYB75, MYB90, bHLH, WD40*) were strongly correlated with flavonoid levels, particularly under environmental stress.

Discussion: These findings provide insights into the genetic regulation of flavonoid biosynthesis in *C. paliurus* and highlight the importance of seasonal and environmental factors. This knowledge has practical implications for industrial breeding and biotechnological applications, particularly in enhancing the functional properties of *C. paliurus* for industrial use. Our study establishes a foundation for future research aimed at optimizing flavonoid production in this species and exploring its potential for bioactive compound production.

KEYWORDS

Cyclocarya paliurus, flavonoid biosynthesis, seasonal variation, transcriptomics, metabolomics

1 Introduction

Cyclocarya paliurus (Batal.) Iljinsk., also known as the wheel wingnut or golden leaf tree, is a deciduous species in the Juglandaceae family, native to central and southern China. This tree can grow up to 30 meters tall and features a broad canopy with pinnate leaves that turn golden-yellow in autumn, enhancing its ornamental value (Shang et al., 2015; Zou et al., 2018; Zhao, 2020; Zheng et al., 2020). *C. paliurus* produces small monoecious flowers and winged nuts for wind dispersal, thriving in warm, humid climates and mixed forests at elevations of 500-2000 meters (Sun et al., 2021). Its ecological roles include providing wildlife habitat and aiding in soil stabilization.

The leaves of *C. paliurus* are rich in flavonoids, bioactive compounds known for their antioxidant, anti-inflammatory, and hypoglycemic properties. These flavonoids contribute to the plant ecological adaptability and have significant medicinal potential (Xie et al., 2015). In traditional Chinese medicine, the leaves are used for their health benefits, including anti-inflammatory and hypoglycemic effects (Zhou et al., 2019). Despite not being endangered, *C. paliurus* faces threats from habitat loss and overharvesting, making conservation efforts necessary (Chen et al., 2022).

C. paliurus shows seasonal variations in leaf production, which is most abundant during summer and autumn, particularly from August to November. This period is critical for studying the leaf biochemical and physiological properties, as production decreases in spring and winter. Our research focuses on differences in the metabolome and transcriptome of leaves collected from August to November, providing insights into the biosynthetic pathways regulating flavonoid production during peak growth (Fu et al., 2015).

Flavonoids, present in many fruits and vegetables, are essential for plant physiology, aiding in UV filtration, nitrogen fixation, and floral pigmentation to attract pollinators (Kumar and Pandey, 2013). They also protect plants from oxidative stress and act as defense agents against herbivores and pathogens (Liu et al., 2021). Additionally, flavonoids are involved in signal transduction, regulating various developmental pathways and stress responses (Falcone Ferreyra et al., 2012). In humans, flavonoids help neutralize free radicals, reducing the risk of chronic diseases like cancer and cardiovascular conditions, while supporting metabolic health (Liu et al., 2018; Hussain et al., 2020).

While previous studies, such as Sheng et al., provided insights into the flavonoid biosynthesis in *C. paliurus*, they did not examine how this process varies across different seasons (Sheng et al., 2021). Our study aims to fill this gap by investigating the temporal regulation of flavonoid biosynthesis from August to November, offering new insights into seasonal variations and regulatory mechanisms. By integrating transcriptomic, metabolomic, and seasonal data, we seek to optimize the harvesting periods for flavonoids, a novel perspective not explored in earlier studies.

Additionally, we previously reported the genome assembly of *C. paliurus* (~634.90 Mb), providing a foundation for further research on its nutrient metabolites (Zheng et al., 2021). This study aims to investigate the seasonal variation in flavonoid content from August to November, compare the transcriptomic and metabolomic profiles involved in flavonoid synthesis, and identify key changes in flavonoid compounds and the genes responsible for their synthesis. These insights are crucial for agriculture and medicine by enhancing our understanding of flavonoid biosynthesis regulation and contributing to the sustainable use of *C. paliurus*.

2 Materials and methods

2.1 Plant material collection

The study site for collecting *C. paliurus* leaves is located in Taozhou Township, Anxi County, Fujian Province, China (E117° 45′42′, N25°22′48′). Samples were collected from August to November 2021, with 4 to 6 plants of similar height and morphology selected from the site. Midsections of leaves (5-10 cm in length) from 10 to 30 leaves per plant were harvested and pooled to form experimental samples. After collection, the samples were immediately placed in zip-lock bags, frozen in liquid nitrogen, and stored at -80°C (Zheng et al., 2021). The leaf samples were divided into three biological replicates for the determination of total flavonoid content. Leaf samples from August, September, October, and November were divided into three biological replicates (8M-ML1/2/3, 9M-ML1/2/3, 10M-ML1/2/3, 11M-ML1/2/3) for metabolomics and transcriptomics analyses. Prior to detection, the samples were freeze-dried.

2.2 Flavonoid content and metabolite analysis for *C. paliurus*

2.2.1 Sample preparation and extraction

Biological samples were freeze-dried using a vacuum freeze-dryer (Scientz-100F). The freeze-dried samples were then crushed using a mixer mill (MM 400, Retsch) with a zirconia bead for 1.5 minutes at 30 Hz. A 100 mg portion of the lyophilized powder was dissolved in 1.2 ml of 70% methanol solution. The mixture was vortexed for 30 seconds every 30 minutes, repeated six times, and stored at 4°C overnight. After centrifugation at 12,000 rpm for 10 minutes, the extracts were filtered through a 0.22 μm pore size filter (SCAA-104; ANPEL, Shanghai, China) before undergoing UPLC-MS/MS analysis (Luo et al., 2013).

2.2.2 Determination of flavonoid content

For the rutin standard solution, 0.05 g of rutin was dissolved in anhydrous ethanol in a 50 mL volumetric flask. Aluminum nitrate solution was prepared by dissolving 17.60 g of aluminum nitrate nonahydrate in water, and potassium acetate solution was prepared by dissolving 9.814 g of potassium acetate in water, with both solutions adjusted to 100 mL. Calibration was performed by placing 1.0 to 5.0 mL of the rutin solution in separate 50 mL volumetric flasks and filling them to 15 mL with ethanol. Subsequently, 1.00 mL each of the aluminum nitrate and potassium acetate solutions was added, and the final volume was adjusted with water. After standing for 1 hour, absorbance was measured at 420 nm. *C. paliurus* leaves were freeze-dried for 24 hours, and then 1.50 g of the dried sample was ground. The sample was extracted with 30 mL of ethanol at 40°C using ultrasound for 1 hour, with shaking every 20 minutes. The solution was then filtered and brought to a final volume of 50 mL with ethanol. For analysis, 1.0 mL of this extract was diluted to 15.0 mL with ethanol, and the flavonoid content was measured using the standard curve (Liu et al., 2016).

2.2.3 Metabolite analysis for *C. paliurus*

The sample extracts were analyzed using a UPLC-ESI-MS/MS system (UPLC, SHIMADZU Nexera X2; MS, Applied Biosystems 4500 Q TRAP). The UPLC utilized an Agilent SB-C18 column (1.8 µm, 2.1 mm x 100 mm). The mobile phase consisted of solvent A (pure water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid). The gradient program started with 95% A and 5% B, transitioning linearly to 5% A and 95% B within 9 minutes, maintaining this ratio for 1 minute. The program then reverted to 95% A and 5% B over 1.1 minutes, held for 2.9 minutes. The flow rate was 0.35 ml/min, the column oven was set to 40°C, and the injection volume was 4 µl. The effluent was connected to an ESI-triple quadrupole-linear ion trap (QTRAP)-MS. LIT and triple quadrupole (QQQ) scans were acquired using an AB4500 Q TRAP UPLC/MS/MS System with an ESI Turbo Ion-Spray interface, operating in both positive and negative ion modes, controlled by Analyst 1.6.3 software (AB Sciex). The ESI source parameters were: source temperature 550°C, ion spray voltage 5500 V (positive)/-4500 V (negative), ion source gas I and II at 50 and 60 psi, respectively, curtain gas at 25 psi, and high collision-activated dissociation (CAD). Instrument tuning and mass calibration were performed with 10 and 100 µmol/L polypropylene glycol solutions in QQQ and LIT modes. QQQ scans were acquired as MRM experiments with nitrogen collision gas set to medium. Specific MRM transitions were monitored for each period based on the eluted metabolites (Zheng et al., 2022).

2.3 Transcriptomics Analysis *C. paliurus*

Total RNA was extracted using the RNeasy Plant Mini Kit (Tiangen Bio, Beijing, China). For library construction, both rRNA-depleted stranded RNA-seq and small RNA sequencing were performed using the TruSeq Stranded Total RNA Prep Kit (Illumina, San Diego, USA). Sequence alignment to the reference genome was conducted using HISAT2, and FPKM values for mRNAs were calculated with StringTie (version 1.3.1). Differential expression analysis between the two conditions/groups was carried out using the DESeq R package (version 1.10.1). DESeq provides statistical routines for determining differential expression in digital gene expression data, utilizing a model based on the negative binomial distribution. The resulting p-values were adjusted for multiple testing using the Benjamini and Hochberg approach to control the false discovery rate. Genes with an adjusted p-value of <0.01 and an absolute log2(fold change) >1, as identified by DESeq, were designated as differentially expressed. Gene function annotation was performed using several databases, including Nr (NCBI non-redundant protein sequences), Pfam (Protein family), KOG/COG (Clusters of Orthologous Groups of proteins), Swiss-Prot (a manually annotated and reviewed protein sequence database), and KEGG (Kyoto Encyclopedia of Genes and Genomes). Gene Ontology (GO) enrichment analysis of the differentially expressed genes (*DEGs*) was conducted using the topGO R package. Additionally, statistical enrichment of differentially expressed genes in KEGG pathways was tested using KOBAS software (Mao et al., 2005; Trapnell et al., 2009).

2.4 Pearson correlation analysis of gene-flavonoid relationships

We selected 39 flavonoid substances with high content from 144 flavonoid substances, which accounted for more than 80% of the total flavonoid substances. Pearson correlation analysis was carried out between the seasonal variation of the contents of 39 flavonoids and the seasonal variation of the expression dose of 66 functional genes related with the flavonoid synthesis pathway and 11 transcription factors. Pearson correlation coefficient and pvalue matrix are generated. When the Pearson correlation coefficient of a pair of flavonoid content and gene expression dose data was greater than 0.8, and pvalue was less than 0.05, a significant correlation was determined.

3 Results

3.1 Determination of total flavonoid content in *C. paliurus* during different months

Based on Figure 1A, the total flavonoid content in *C. paliurus* leaves shows significant seasonal variation from August to November. In August, the flavonoid content is at its peak, reaching approximately 10 mg/g. This high level decreases sharply in September, dropping to around half of the August value. However, in October, the flavonoid content increases again to a level comparable to that in August, indicating a partial recovery. By November, the flavonoid level declines slightly from October, reaching about 80% of the October content. Statistical analysis reveals significant differences between these months, as indicated by the asterisks (**$p < 0.01$, *$p < 0.05$), suggesting that environmental factors and seasonal cues may influence flavonoid biosynthesis in *C. paliurus* leaves across these months. This pattern highlights a dynamic adjustment in flavonoid production, possibly linked to the plant's adaptive response to seasonal changes.

3.2 Flavonoid metabolic differences of *C. paliurus* leaves during different months

The analysis revealed the presence of 751 secondary metabolites in *C. paliurus*, including amino acids and their derivatives, phenolic acids, nucleotides and their derivatives, flavonoids, lignans,

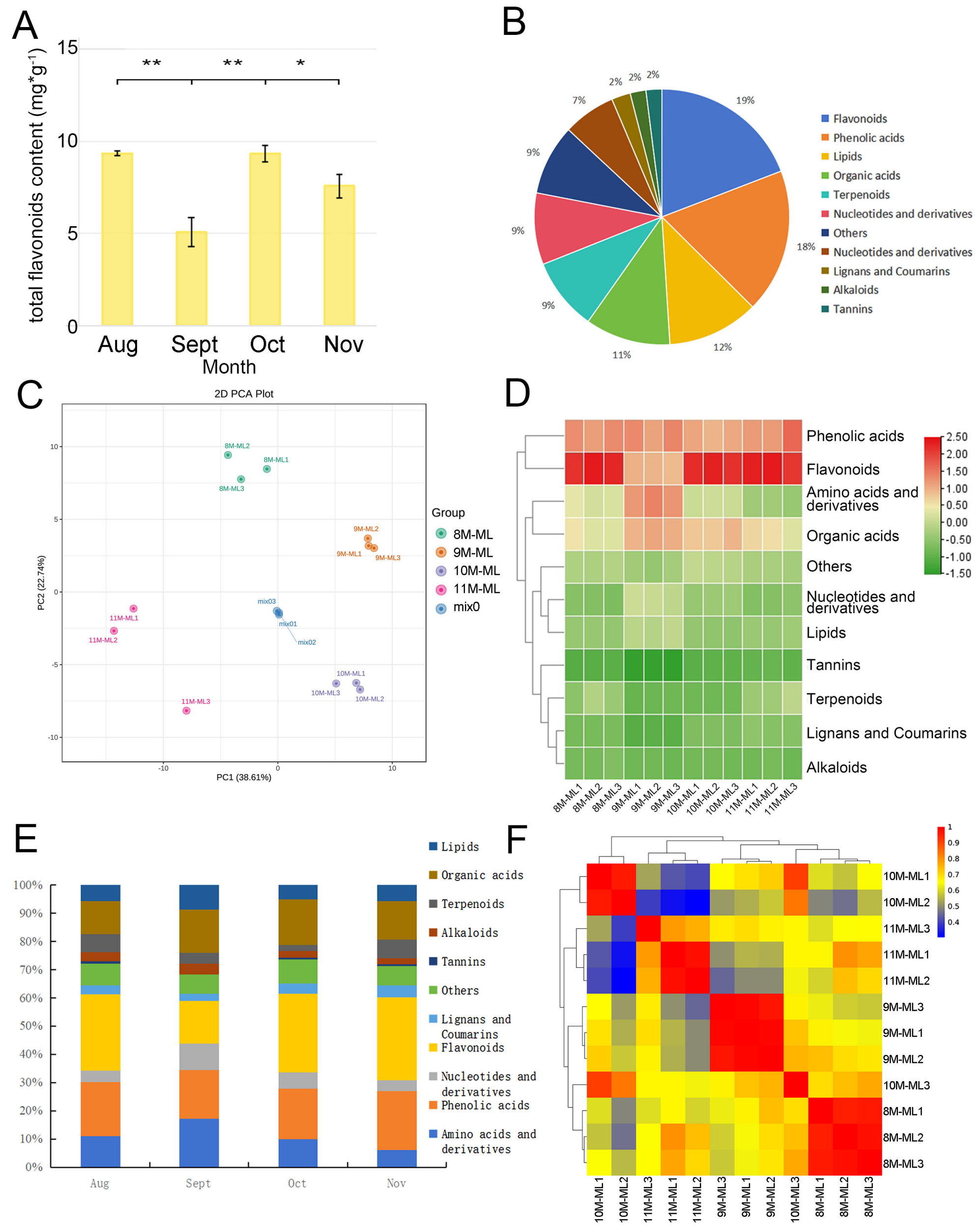


FIGURE 1

Characteristics of flavonoids in *C. paliurus*. **(A)** Total flavonoid content in *C. paliurus* leaves across different months. **(B)** Distribution of various metabolites in *C. paliurus* leaves. **(C)** PCA score plot showing the clustering of different leaf samples. **(D)** Heatmap clustering of the identified compounds. **(E)** Trends in the relative abundance of different compound classes from August to November. **(F)** Correlation matrix among leaf samples based on their metabolite profiles.

coumarins, tannins, alkaloids, terpenes, organic acids, lipids, and other metabolites. Figure 1B illustrates the distribution of various metabolite categories in *C. paliurus* leaves, with flavonoids and nucleotides dominating the distribution, totaling 144 and 137 metabolites, respectively, accounting for more than 19.17% of the total metabolites. Other significant categories include terpenoids (87), amino acids (81), and lipids (69).

The PCA score plots for each sampling time (Figure 1C) show clear separation between samples collected during different months, indicating significant variation in substance accumulation at the metabolomic level. The data of flavonoid metabolites in *C. paliurus* leaves from different harvesting periods were analyzed by PCA, and the results are presented in Figure 1C, where each point represents one sample. The results indicate that the samples from the four harvesting periods were distinctly separated, highlighting the differences in flavonoid components across these periods. The groups from August (8M-ML) and September (9M-ML) were closely positioned on the score map, suggesting that the flavonoid components in the leaves from these two months were similar. Although PCA and the Pearson correlation coefficient algorithm are different, both methods reveal significant differences in terpenoid composition from August to November.

In our study, the data were normalized, and the accumulation patterns of metabolites in different samples were analyzed through

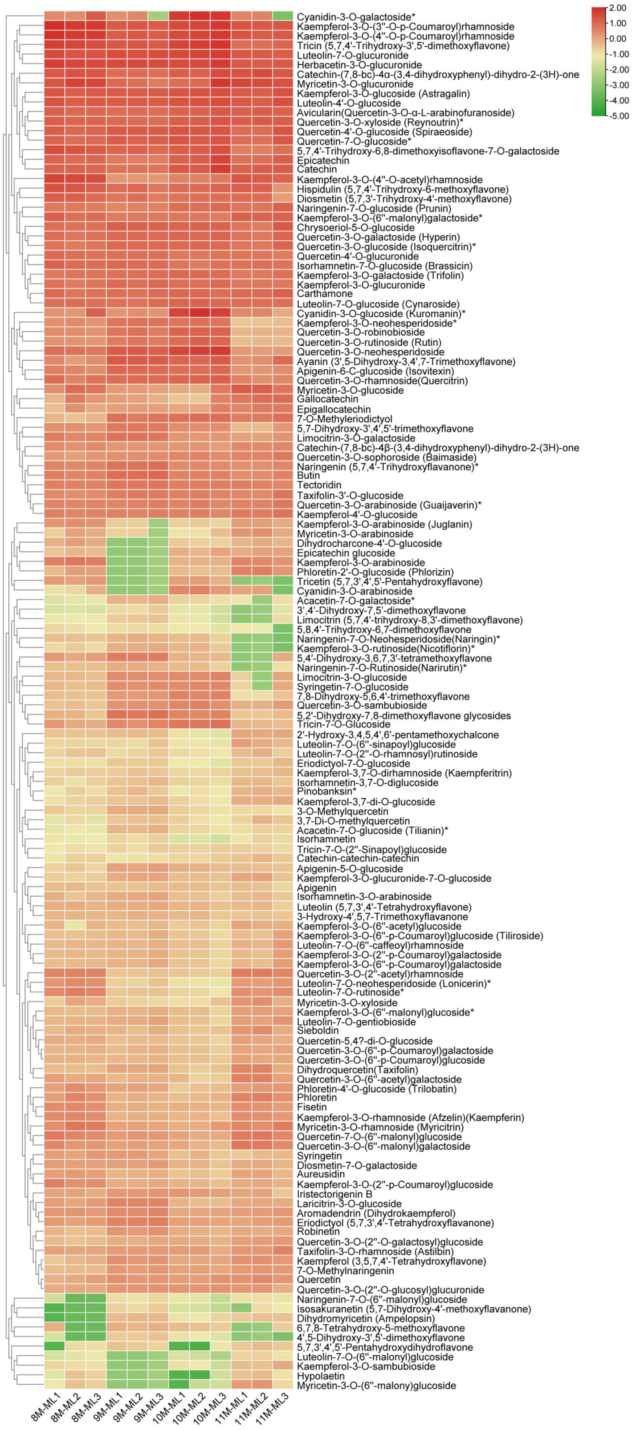


FIGURE 2
Clustering of 144 flavonoid compounds identified in *C. paliurus* leaves.

clustering. Figure 1D presents the clustering results, which show good intra-group repeatability for the leaf samples from the four harvesting periods. A significant difference in the distribution of leaf components in September was observed compared to the other months, while the distribution in the other three months was relatively similar. The metabolites during the four harvesting stages were primarily concentrated in phenolic acids, flavonoids, amino acids, and organic acids.

All flavonoids in the metabolome of *C. paliurus* from August to November were selected for further analysis. The bar accumulation diagram in Figure 1E illustrates the seasonal variation in the proportion of flavonoids in *C. paliurus* leaves across the four harvesting periods: August, September, October, and November. In August, flavonoids represented approximately 26%~28% of the total metabolites, reflecting peak biosynthetic activity during the summer, when environmental conditions were optimal for flavonoid production. As the season transitioned into September, the proportion of flavonoids slightly decreased to around 15%, likely due to cooler weather and shorter daylight hours, which may have shifted the plant's metabolic focus. In October, the flavonoid content increased again, reaching 27%~29%, and in November, the flavonoid content reached the highest level, at 27%~31% (Supplementary Table S1). This trend highlights *C. paliurus*' adaptive strategy to modulate flavonoid biosynthesis in response to seasonal changes, optimizing its metabolic processes to ensure survival and sustainability under varying environmental conditions.

By calculating the Pearson correlation coefficient, as shown in Figure 1F, the heatmap visually represents the correlation between leaf samples of *C. paliurus* collected from August to November, with the color gradient indicating the strength of these correlations (red for high, blue for low). Samples collected within the same month, such as those from August, September, October, and November, exhibited strong correlations, indicated by red areas, suggesting consistency in their metabolic profiles during each respective period. In contrast, samples from different months, particularly between August and November, showed lower correlations, depicted by blue and green areas, indicating significant seasonal variations in metabolite composition. Interestingly, the samples from September and October displayed moderate correlations (yellow and orange areas), suggesting some similarity in their metabolic profiles during these transitional months. Overall, the heatmap effectively captures both the internal consistency of each sampling period and the dynamic shifts in metabolite composition as the seasons progress.

Figure 2 shows the content distribution of 144 flavonoid metabolites in each sample across different groups. The clustering results indicate that the proportion of metabolites increased with the progression of months, with the highest accumulation of flavonoids observed in November. The accumulation of kaempferol and its glycosides was concentrated across the four different harvesting periods. For example, the relative contents of kaempferol-3-O-(3''-O-p-coumaryl) rhamnoside, kaempferol-3-O-(4''-O-p-coumaryl) rhamnoside, and kaempferol-3-O-glucoside (Astragalin) were high across all four harvesting periods (Figure 2, Supplementary Table S1).

The analysis of flavonoid metabolism in *C. paliurus* leaves revealed significant seasonal variations in the concentrations of flavonoid metabolites from August to November. In August, the concentrations of several flavonoid metabolites peaked. Notably, quercetin-3-O-glucoside and quercetin-3-O-rutinoside (rutin) exhibited high levels, driven by optimal conditions for biosynthesis. These compounds are crucial for antioxidant defense, aiding the plant in coping with oxidative stress. Similarly, kaempferol-3-O-glucoside and kaempferol-3-O-rutinoside were abundant, reflecting their roles in UV protection and antioxidative activities (Lin et al., 2020). As the season transitioned into September, a noticeable decline in some metabolites was observed, for example, kaempferol-3-O-arabinoside, phloretin-2'-O-glucoside and tricetin. The reduced levels of these metabolites suggest a shift in metabolic activity, possibly due to lower environmental stress and the onset of dormancy preparation (Zhang et al., 2020). quercetin-3-O-glucoside (Isoquercitrin) and kaempferol-3-O-glucoside (astragalin) always maintain a high level. By October, the concentrations of these flavonoids continued to increase, reaching a peak. This trend was consistent for quercetin-3-O-neohesperidoside, tricin, as well as tricin-7-O-glucoside. and luteolin-7-O-glucuronide, known for their anti-inflammatory properties, also showed significant reductions from their peak levels in August. The content level of catechin and epicatechin, important for antioxidative protection, remained consistently high, especially in October, reaching a peak. By November, flavonoid concentrations were at their low level, reflecting the plant's transition to dormancy and significantly slowed metabolic activity. Flavonoids such as quercetin-3-O-rutinoside (rutin), cyanidin-3-O-galactoside and cyanidin-3-O-glucoside showed minimal levels in November, consistent with the overall downregulation of flavonoid biosynthesis. (Figure 2, Supplementary Tables S2-S4). As the season progresses, *C. paliurus* reduces flavonoid synthesis, conserving resources and preparing for dormancy. This adaptive strategy underscores the dynamic regulation of flavonoid metabolism in response to environmental changes, highlighting the ecological and physiological importance of these compounds (Mohammadi et al., 2014; Sun et al., 2024).

3.3 Transcriptomics analysis for *C. paliurus* leaves in different months

3.3.1 Screening and analysis of differentially expressed genes

In the transcriptome data collected across four different harvesting periods from August to November, we focused on the expression analysis of differentially expressed genes (*DEGs*) in three specific comparisons: 8MML_vs_9MML, 9MML_vs_10MML, and 10MML_vs_11MML. DEGs were identified by comparing expression levels between adjacent months using DEGseq software, with the screening criteria set to a p-value of less than 0.01 and a log2 fold change greater than 1.5. The screening results are presented in the form of volcano plots, as shown in Figures 3A, C, E. The log2 fold change values on the horizontal axis indicate upregulated genes (greater than 0) and downregulated genes (less than 0), while the vertical axis represents the significance of differential gene expression between the two sample groups. Specifically, in the 8MML_vs_9MML group, we identified 1,226 upregulated and 1,184 downregulated genes (Figure 3A). In the 9MML_vs_10MML group, 5 genes were upregulated and 38 genes were downregulated (Figure 3C). In the 10MML_vs_11MML comparison, 413 genes were upregulated and 59 genes were downregulated (Figure 3E).

Further analysis was performed using Gene Ontology (GO) enrichment for the *DEGs* identified in the 8MML_vs_9MML, 9MML_vs_10MML, and 10MML_vs_11MML comparisons, with

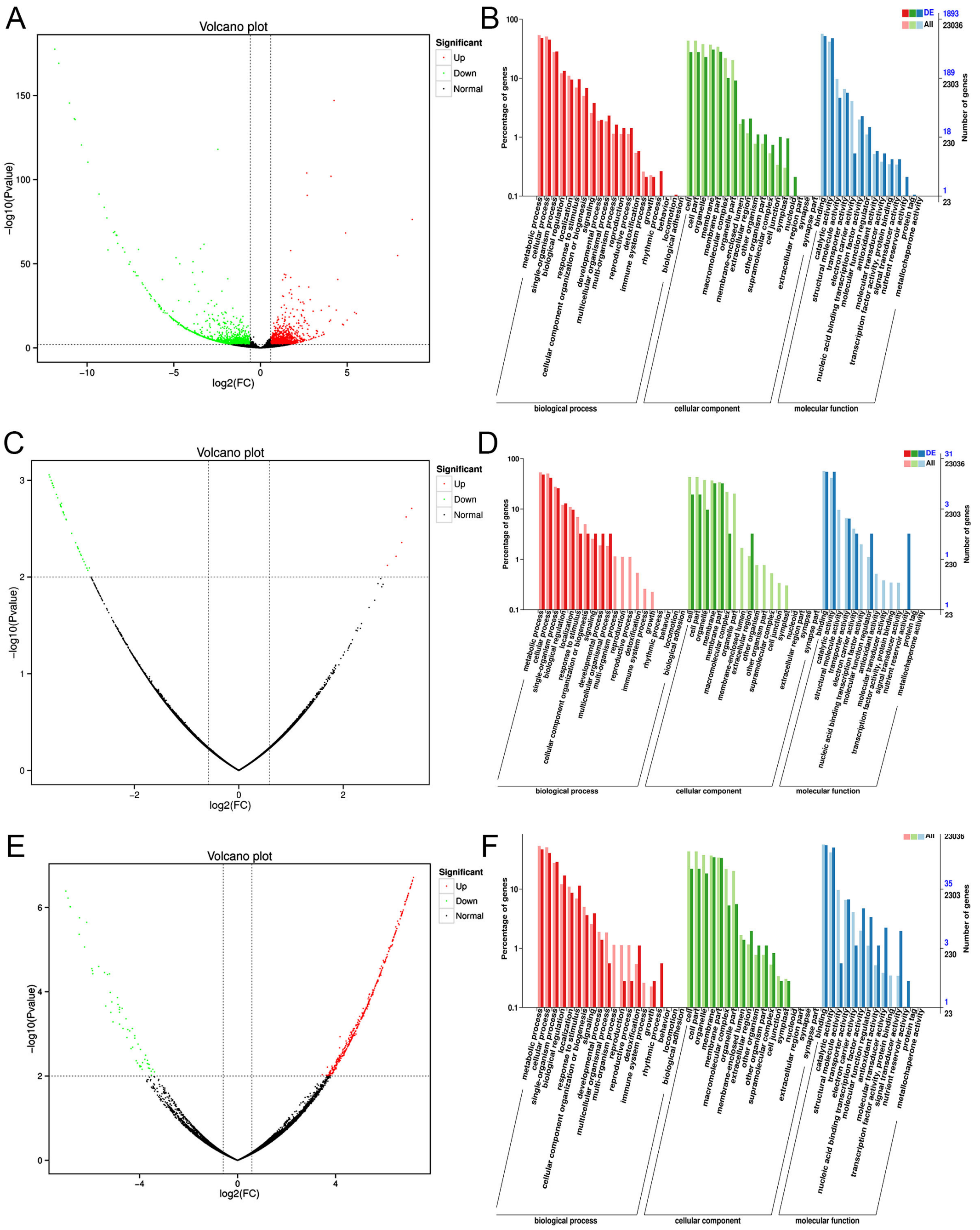


FIGURE 3

Volcano plots and GO enrichment analysis for differentially expressed genes (*DEGs*) in *C. paliurus*. **(A)** Volcano plot of *DEGs* between 8M-ML and 9M-ML samples. **(B)** GO enrichment analysis of *DEGs* between 8M-ML and 9M-ML samples. **(C)** Volcano plot of *DEGs* between 9M-ML and 10M-ML samples. **(D)** GO enrichment analysis of *DEGs* between 9M-ML and 10M-ML samples. **(E)** Volcano plot of *DEGs* between 10M-ML and 11M-ML samples. **(F)** GO enrichment analysis of *DEGs* between 10M-ML and 11M-ML samples.

the results illustrated in Figures 3B, D, F. The GO enrichment analysis highlighted significant biological processes, with most differentially expressed genes (*DEGs*) annotated to processes primarily involved in metabolic and cellular functions. This underscores the essential role of these processes in regulating flavonoid biosynthesis and related metabolic pathways across the different seasonal stages. In the biological process category, *DEGs* were predominantly enriched in metabolic and cellular processes, suggesting that these genes are crucial in regulating various metabolic pathways during the seasonal fluctuations in flavonoid production. The cellular component category showed a high concentration of *DEGs* associated with the cell membrane, particularly in processes related to the formation of cell membrane components. This finding indicates the importance of cellular structure and membrane dynamics in maintaining metabolic functions during the seasonal transitions. Moreover, in the molecular function category, *DEGs* were most enriched in binding and catalytic activities, further emphasizing the role of these genes in driving key biochemical reactions involved in flavonoid biosynthesis. These results not only highlight the key molecular functions associated with flavonoid metabolism but also provide a deeper understanding of the regulatory mechanisms underlying the seasonal changes observed in *C. paliurus*. The GO enrichment analysis thus reveals important insights into the molecular framework governing the plant's adaptive responses, reinforcing the seasonal modulation of flavonoid biosynthesis.

Subsequently, KEGG pathway enrichment analysis was conducted for the *DEGs* identified in the 8MML_vs_9MML, 9MML_vs_10MML, and 10MML_vs_11MML comparisons, with the results depicted in Supplementary Figures S1A–C. In the 8MML_vs_9MML group, a total of 569 *DEGs* were enriched, with a significant concentration in metabolism-related pathways, including plant-pathogen interaction, starch and sucrose metabolism, and notably, 9 *DEGs* related to the flavonoid biosynthesis pathway (Supplementary Tables S6), including *CYP90A1*, cinnamoyl-CoA reductase, *VSR6*, Flavonoid 3'-monooxygenase, *CYP714C2, CYP71AN24*, Chalcone and stilbene synthases, *CYP87A3* and *CYP86A22*. In contrast, in the 9MML_vs_10MML group, 38 *DEGs* were enriched, with the majority being related to plant-pathogen interaction, and 2 *DEGs* were annotated to the flavonoid biosynthesis pathway (Supplementary Tables S7), including *CYP71D9* and shikimate O-hydroxycinnamoyltransferase. Similarly, in the 10MML_vs_11MML group, 420 *DEGs* were enriched, with a strong emphasis on metabolic and organismal systems pathways, including 4 *DEGs* related to the flavonoid biosynthesis pathway (Supplementary Tables S8), including stemmadenine O-acetyltransferase, flavonoid 3'-monooxygenase, Isoflavone reductase and trans-cinnamate 4-monooxygenase.

The comprehensive analysis of *DEGs* in *C. paliurus* leaves across the critical periods from August to November identified numerous genes implicated in both the upstream biosynthesis and downstream metabolism of flavonoids. Key enzymes in the flavonoid biosynthesis pathway, such as Chalcone Synthase (*CHS*), Chalcone Isomerase (*CHI*), and Flavanone 3-Hydroxylase (*F3H*), were prominently upregulated during the transition from August to September, suggesting an early surge in flavonoid production. Additionally, Dihydroflavonol 4-Reductase (*DFR*) and Anthocyanidin Synthase (*ANS*) were significantly upregulated from September to October, correlating with increased production of anthocyanins and other flavonoid derivatives. Notably, Flavonol Synthase (*FLS*), which facilitates the conversion of dihydroflavonols to flavonols, was also upregulated in this period, indicating a diversification of flavonoid types. In the later months, from October to November, genes such as Leucoanthocyanidin Reductase (*LAR*) and Flavonoid 3'-Hydroxylase (*F3'H*) were downregulated, indicating a decrease in the synthesis of proanthocyanidins and other flavonoid metabolites as the leaves transitioned into winter. Additionally, genes involved in the regulation and modification of flavonoids, including UDP-glucose-3-O-glucosyltransferase (*UFGT*) and Cytochrome P450 enzymes like *CYP75B1*, which are critical for the hydroxylation and glycosylation of flavonoids, exhibited varying expression patterns, underscoring the complex regulation of flavonoid metabolism. This dynamic expression of flavonoid-related genes highlights the intricate balance between biosynthesis and metabolism that is modulated in response to seasonal changes, offering valuable insights into the optimal harvesting period and potential genetic targets for enhancing flavonoid content in *C. paliurus* leaves (Supplementary Tables S6-S8).

3.3.2 Integrated weighted gene co-expression network analysis of transcriptomic and metabolomic data

To analyze gene-metabolite correlations related to flavonoid biosynthesis, we conducted Weighted Gene Co-expression Network Analysis (WGCNA) using flavonoid-related metabolites and transcriptomic data from *C. paliurus* leaves collected between August and November. We selected an optimal soft threshold of β=19, ensuring a scale-free network with a fitting index above 0.8 (Langfelder and Horvath, 2008). This threshold allowed us to construct a robust co-expression network (Figures 4A, B).

We identified 20 distinct gene modules, each visualized by a unique color in Figure 4C. Highly correlated genes were grouped within the same module, while the gray module contained genes without significant correlation patterns. The lightsteelblue1 module had the most genes (2,479), while the thistle1 module had the fewest (41). Figure 4C illustrates the correlations between modules, with blue shades indicating weaker correlations and red indicating stronger ones. As shown in Figures 4D, E, we examined correlations between gene modules and specific flavonoids detected in the metabolomic analysis, including compounds like naringin, geranolin, kaempferol-3-O-arabinoside, kaempferol-3-O-rhamnoside, and naringin-7-O-glucoside. The tan module demonstrated a high correlation (correlation coefficient > 0.6) with key flavonoids, such as kaempferol-3-O-glucoside, kaempferol-4'-O-glucoside, and kaempferol-3,7-di-O-glucoside. Other modules, including thistle1 and plum2, were associated with naringenin and specific kaempferol derivatives, such as kaempferol-3-O-(2"-p-coumaryl) galactoside and kaempferol-3-O-(6"-p-coumaryl) galactoside (Xie et al., 2015).

To further investigate regulatory genes, we constructed an interaction network of genes within the tan module using Cytoscape, focusing on the top 200 genes with the highest co-expression weights. As visualized in Figure 4F, we identified ten key genes, *CpBBX* (newGene_23931), *CpIP5P3* (newGene_15401), *CpLHY* (newGene_27003), *CpSGPP* (newGene_29931), *CpCIPK3* (newGene_19518), *CpAPF1* (newGene_24020), *CpHSFB3* (newGene_10195), *CpLOC5396* (newGene_2447), *CpMANES* (newGene_24731), and *CpLOC7593* (newGene_28372), as central regulators within this network (Shannon et al., 2003). Their high connectivity suggests pivotal roles in regulating flavonoid biosynthesis. For example, Cyclocarya_paliurus_newGene_27003 and newGene_29931 were notably active in October, with corresponding upregulation of metabolites like quercetin-3-O-(2"-acetyl) rhamnoside and luteinin-7-O-rutinoside in November. This pattern suggests that these genes may contribute to the synthesis of flavonoids as seasonal conditions change.

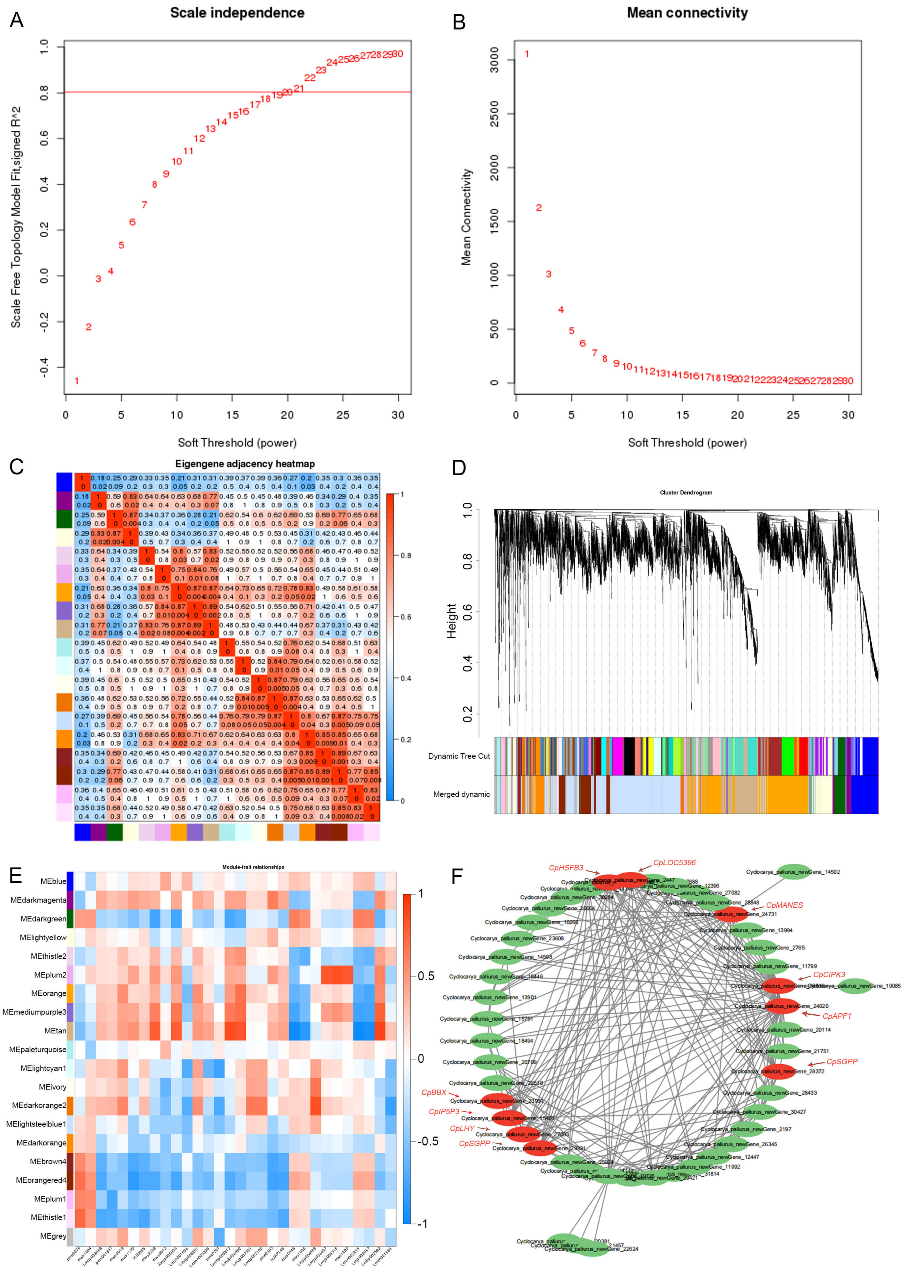


FIGURE 4
Gene co-expression network analysis. **(A)** Scale independence as a function of soft-thresholding power. **(B)** Mean connectivity across soft-thresholding powers. **(C)** Heatmap showing correlations between gene expression modules. **(D)** Clustering dendrogram of expressed genes, with modules indicated by different colors. **(E)** Heatmap illustrating correlations between gene modules and various traits. **(F)** Co-expression network of genes in a selected module, highlighting hub genes.

3.3.3 Identification of key genes in the flavonoid biosynthesis pathway

Comprehensive differential expression analysis revealed several key genes involved in the flavonoid biosynthesis pathway in *C. paliurus*, with notable seasonal variations across the studied months. Chalcone Synthase (*CHS*), a critical enzyme responsible for initiating flavonoid biosynthesis, exhibited peak expression in August, aligning with the highest flavonoid content observed during this period. This suggests that *CHS* serves as a major regulatory enzyme driving flavonoid production in response to favorable environmental conditions. Similarly, Chalcone Isomerase (*CHI*), which converts naringenin chalcone to naringenin, displayed elevated expression in August and September, ensuring an adequate supply of naringenin as a precursor for downstream flavonoid biosynthesis (Wan et al., 2024). Flavanone 3-Hydroxylase (*F3'H*), critical for diversifying flavonoid structures, peaked in September, while Dihydroflavonol 4-Reductase (*DFR*) showed heightened expression in both August and September, supporting anthocyanin biosynthesis during this period. Flavonol Synthase (*FLS*), which catalyzes the conversion of dihydroflavonols into flavonols such as quercetin and kaempferol, also exhibited peak expression in August, coinciding with the accumulation of these antioxidant compounds (Figures 5A, B) (Jaakola, 2013).

Figure 5A illustrates the flavonoid biosynthetic pathway in *C. paliurus*, mapping the transformation of precursors like cinnamoyl-CoA and p-coumaroyl-CoA into diverse flavonoids through the activity of key enzymes, including *CHS, CHI, F3'H, FLS*, Anthocyanidin Synthase (*ANS*), and Anthocyanidin Reductase (*ANR*) (Winkel-Shirley, 2001; Hichri et al., 2010). The heat maps integrated into this pathway visually represent the differential expression of these genes, with color gradients highlighting the upregulation and downregulation patterns in response to developmental and environmental stimuli (Falcone Ferreyra et al., 2012). These expression patterns underline the dynamic regulation of flavonoid biosynthesis, driven by the plant's adaptive responses to oxidative stress, UV radiation, and other environmental challenges, while simultaneously contributing to human health benefits through antioxidant and anti-inflammatory properties (Wang et al., 2024). Genes associated with the biosynthesis of procyanidins, such as Leucoanthocyanidin Reductase (*LAR*), *DFR, ANS*, and *ANR*, were also dynamically expressed. These enzymes play essential roles in synthesizing catechins and epicatechins, as shown in Figure 5B, which emphasizes the significance of flavonol glycosides and procyanidins as key flavonoid classes with diverse biological functions (Xie et al., 2015; Xu et al., 2024). Notably, the heat maps reveal seasonal changes in gene expression, with green and red denoting down- and upregulation, respectively, reflecting the plant capacity to modulate flavonoid synthesis in response to environmental or developmental cues (Hichri et al., 2010).

From August to November, dynamic shifts in the expression of flavonoid-related genes were observed. Genes such as CHS, CHI, and F3'H exhibited prominent upregulation during the transition from August to September, corresponding to an early surge in flavonoid production. In September and October, elevated expression levels of *DFR* and *ANS* coincided with increased

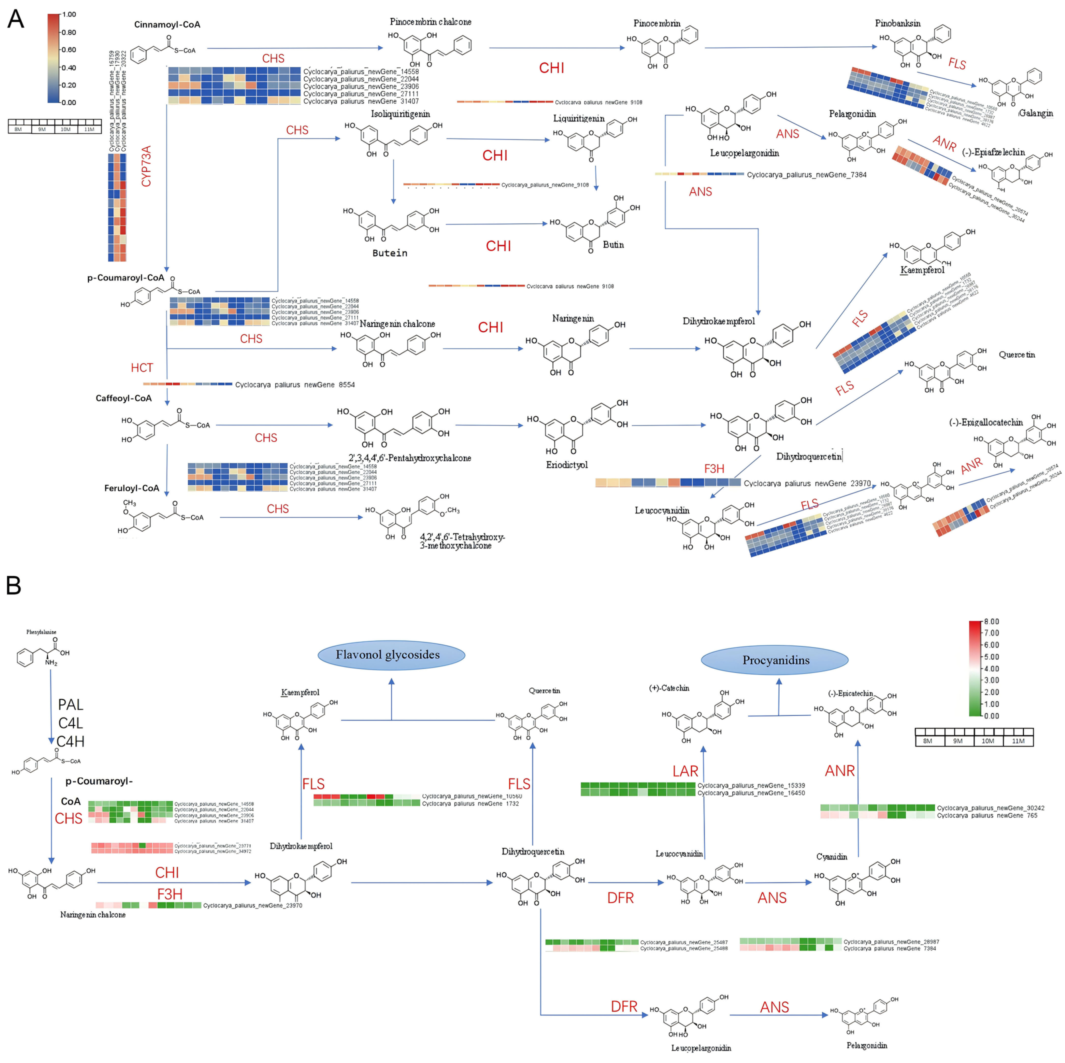


FIGURE 5

Transcriptional expression profiles of key functional genes involved in flavonoid biosynthesis. **(A)** Expression profiles of genes involved in the transformation from precursor molecules to flavonoids. **(B)** Expression profiles of genes responsible for flavonoid biosynthesis.

production of anthocyanins, while FLS activity diversified flavonoid profiles by facilitating the synthesis of flavonols. By late autumn (October to November), genes such as LAR and F3'H were downregulated, marking a decline in proanthocyanidin synthesis as the plant transitioned into dormancy. Additionally, genes involved in flavonoid regulation and modification, including UDP-glucose-3-O-glucosyltransferase (*UFGT*) and Cytochrome P450 enzymes like *CYP75B1*, exhibited fluctuating expression patterns, underscoring the complexity of flavonoid metabolism.

The interaction of 57 genes linked to 16 enzymes, as noted in the KEGG pathway (ko00941), highlights the extensive regulatory control underlying flavonoid biosynthesis. These regulatory networks enable *C. paliurus* to adapt to seasonal changes, bolstering its resilience to environmental stressors while optimizing flavonoid production for both plant defense and human health benefits (Falcone Ferreyra et al., 2012). The findings presented in Figures 5A, B; Supplementary Tables S6-S8 not only emphasize the intricate balance between flavonoid biosynthesis and metabolism but also provide valuable insights for optimizing the harvesting period and identifying potential genetic targets for enhancing flavonoid content in *C. paliurus*.

3.3.4 Analysis of gene-flavonoid relationships and their seasonal variations

The total number of flavonoid substances in the leaves of *C. paliurus* was 144, among which 39 major flavonoid substances with high content accounted for more than 80% of the total flavonoid content from August to November (Table 1). Therefore, we selected these 39 flavonoid substances for the analysis of gene-flavonoid relationships and seasonal variations, Pearson correlation studies were carried out between them and the function genes of the flavonoid synthesis pathway (Supplementary Table S9) and the significantly different transcription factors in *DEGs* (Supplementary Table S10). The regulation of flavonoid biosynthesis in *C. paliurus* was thoroughly examined through a comprehensive Pearson correlation analysis, which linked the expression of functional biosynthetic genes and regulatory transcription factors to seasonal variations in flavonoid content. The data from Figures 6, 7 provide critical insights into how these genes exhibit coordinated fluctuations across different months, significantly influencing the production of key flavonoids such as quercetin and kaempferol glycosides. This seasonal regulation sheds light on the plants intricate molecular responses to environmental stressors and climatic changes (Supplementary Tables S11-S12).

Beginning with the analysis of functional genes as illustrated in Figure 6, it is evident that several core biosynthetic genes play pivotal roles in driving the flavonoid biosynthesis pathway. Among these, chalcone synthase (*CHS*) stands out as the key enzyme catalyzing the initial step in flavonoid biosynthesis, converting malonyl-CoA and p-coumaroyl-CoA into chalcones. The strong positive correlation between *CHS* expression and flavonoid accumulation, particularly in late summer, suggests that this gene is crucial for the overall regulation of flavonoid biosynthesis in *C. paliurus*. This observation is consistent with previous studies conducted in white apruce, where it was revealed that accumulation of *CHS* mRNA in needle tissue following mechanical wounding, or application of signal molecules, such as jasmonic acid or methyl jasmonate (Richard et al., 2000).

Similarly, the role of chalcone isomerase (*CHI*), which converts chalcones into flavanones, is highlighted by its high expression during the peak flavonoid accumulation period. This gene plays a vital role in facilitating downstream biosynthesis, particularly in the synthesis of flavonols. The expression patterns observed in *C. paliurus* align with findings in *Glycine max*, where *CHI* activity increased under environmental stress, enabling the plant to boost flavonoid production as a means of stress adaptation (Seehaus and Tenhaken, 1998). Furthermore, flavanone 3'-hydroxylase (*F3'H*), which catalyzes the conversion of flavanones into dihydroflavonols, demonstrated a strong correlation with both quercetin and kaempferol derivatives. This gene is integral to the overall balance of flavonol types produced, and its seasonal expression pattern mirrors those observed in other species such as *Malus domestica*, where *F3'H* plays a key role in regulating flavonoid biosynthesis during fruit development (Torres et al., 2020).

Moving further down the biosynthetic pathway, dihydroflavonol 4-reductase (*DFR*) exhibited significant correlations with anthocyanin content, particularly with cyanidin-3-O-glucoside. This suggests that *DFR* plays a vital role in anthocyanin biosynthesis during the late summer months, likely as part of the plant's adaptive response to high light conditions. Such findings are corroborated by studies in *Vitis vinifera* (grape), where ubiquitin ligase *VvWRKY24* could directly interact with the promoters of dihydroflavonol-4-reductase (*DFR*) to inhibit proanthocyanidins biosynthesis (Zhao et al., 2024). In parallel, flavonol synthase (*FLS*), which catalyzes the final conversion of dihydroflavonols into flavonols such as quercetin and kaempferol, also demonstrated strong seasonal expression patterns, correlating with the high levels of flavonols observed in late summer. This observation is consistent with research on *Arabidopsis*, where *FLS* was shown to be upregulated in response to environmental stimuli, enhancing the synthesis of flavonols for drought inducing (Ma et al., 2024).

Turning to the analysis of transcription factors in Figure 7, it is apparent that the regulatory machinery governing flavonoid biosynthesis is equally complex and dynamic. Members of the *MYB* family, including *MYB11, MYB12*, and *MYB111*, exhibited strong positive correlations with quercetin and kaempferol glycosides, indicating their central roles in regulating flavonol synthesis. These transcription factors have been widely documented for their involvement in controlling flavonoid biosynthesis, particularly in response to environmental stresses such as UV radiation. In *Arabidopsis thaliana*, for instance, similar *MYB* transcription factors have been shown to directly regulate the expression of flavonol biosynthetic genes, enhancing flavonol accumulation under high-light conditions (Stracke et al., 2007).

MYB75 (*PAP1*) and *MYB90* (*PAP2*) were closely associated with anthocyanin production, particularly cyanidin-3-O-glucoside. These transcription factors are well-known for their roles in regulating anthocyanin biosynthesis in a wide range of plant species. For example, in *Vitis vinifera*, in 'Malbec' grape skins, *VviMyb4a* and *VviMyb4*b likely regulate the early genes in the anthocyanin biosynthesis pathway, while *VviMybC2-L2* and *VviMyb4-like* are

TABLE 1 The content of the main flavonoids in *C. paliuru* leaves.

Compounds	8M		Ratio	9M		Ratio	10M		Ratio	11M		Ratio
	Average	STDEV	(%)	Average	STDEV	(%)	Average	STDEV	(%)	Average	STDEV	(%)
Ayanin (3',5-Dihydroxy-3,4',7-Trimethoxyflavone)	7.56E+05	3.27E+05	0.22	7.87E+06	3.61E+05	3.95	5.63E+06	5.71E+05	1.33	2.41E+06	2.77E+06	0.59
Quercetin-4'-O-glucoside (Spiraeoside)	5.43E+06	5.85E+05	1.55	7.58E+06	5.28E+05	3.81	9.98E+06	1.68E+06	2.36	9.62E+06	5.92E+06	2.35
Quercetin-3-O-neohesperidoside	2.77E+06	2.82E+05	0.79	7.03E+06	1.37E+06	3.53	2.30E+07	3.10E+06	5.44	4.28E+05	4.21E+04	0.1
Quercetin-7-O-glucoside*	4.40E+06	4.58E+05	1.26	6.10E+06	4.26E+05	3.06	9.16E+06	1.97E+06	2.17	8.36E+06	5.63E+06	2.04
Quercetin-3-O-galactoside (Hyperin)	2.79E+06	3.35E+05	0.8	3.38E+06	2.96E+05	1.7	4.92E+06	5.83E+05	1.16	4.50E+06	2.06E+06	1.1
Quercetin-3-O-glucoside (Isoquercitrin)*	2.37E+06	3.96E+05	0.68	3.35E+06	1.93E+05	1.68	4.44E+06	7.11E+05	1.05	4.56E+06	2.49E+06	1.12
Quercetin-3-O-xyloside (Reynoutrin)*	5.68E+06	2.02E+05	1.62	3.31E+06	4.17E+05	1.66	7.63E+06	4.39E+05	1.81	1.37E+07	4.64E+06	3.34
Isorhamnetin-7-O-glucoside (Brassicin)	4.24E+06	6.92E+05	1.21	3.14E+06	2.12E+05	1.58	2.63E+06	3.68E+05	0.62	4.47E+06	1.18E+06	1.09
Quercetin-3-O-rhamnoside(Quercitrin)	2.46E+06	9.32E+05	0.7	3.01E+06	2.96E+05	1.51	5.55E+06	4.61E+05	1.31	1.23E+06	4.01E+05	0.3
Avicularin(Quercetin-3-O-α-L-arabinofuranoside)	4.88E+06	2.32E+05	1.4	2.88E+06	2.70E+05	1.45	6.58E+06	3.12E+05	1.56	1.18E+07	3.67E+06	2.89
Quercetin-4′-O-glucuronide	3.30E+06	1.42E+05	0.94	2.42E+06	1.28E+05	1.22	2.51E+06	1.19E+05	0.59	3.04E+06	8.90E+05	0.74
Quercetin-3-O-robinobioside	7.59E+05	1.08E+05	0.22	1.81E+06	7.09E+05	0.91	5.72E+06	8.31E+05	1.35	1.07E+05	1.45E+04	0.03
Quercetin-3-O-rutinoside (Rutin)	5.99E+05	6.71E+04	0.17	1.38E+06	3.94E+05	0.69	5.11E+06	1.25E+06	1.21	7.71E+04	1.23E+04	0.02
8-Hydroxykaempferol(Herbacetin)-3-O-glucuronide	2.31E+07	1.68E+06	6.59	1.84E+07	1.13E+06	9.23	1.80E+07	7.00E+05	4.27	1.76E+07	4.14E+06	4.29
Kaempferol-3-O-(3''-O-p-Coumaroyl)rhamnoside	4.27E+07	2.90E+06	12.2	8.38E+06	1.72E+06	4.21	2.01E+07	8.26E+05	4.76	1.96E+07	5.19E+06	4.79
Kaempferol-3-O-(4''-O-p-Coumaroyl)rhamnoside	3.93E+07	7.97E+05	11.21	8.25E+06	1.77E+06	4.14	2.00E+07	1.03E+06	4.73	1.82E+07	5.01E+06	4.46
Kaempferol-3-O-glucoside (Astragalin)	8.53E+06	9.96E+05	2.44	7.17E+06	2.63E+05	3.6	1.13E+07	1.95E+06	2.67	1.76E+07	2.27E+06	4.31
Kaempferol-3-O-glucuronide	2.44E+06	3.76E+05	0.7	2.89E+06	4.39E+05	1.45	3.60E+06	5.53E+04	0.85	2.78E+06	2.99E+05	0.68
Kaempferol-3-O-galactoside (Trifolin)	2.37E+06	5.15E+05	0.68	2.47E+06	2.72E+05	1.24	4.32E+06	8.49E+05	1.02	3.25E+06	6.03E+05	0.79
Kaempferol-3-O-neohesperidoside*	7.94E+05	1.45E+05	0.23	2.06E+06	1.84E+05	1.03	2.92E+06	3.22E+05	0.69	4.99E+04	2.48E+04	0.01
Kaempferol-3-O-(6''-malonyl)galactoside*	4.40E+06	9.64E+05	1.26	1.41E+06	1.25E+05	0.71	2.53E+06	4.84E+05	0.6	1.05E+07	1.74E+06	2.57
Kaempferol-3-O-(4''-O-acetyl)rhamnoside	1.96E+07	1.52E+06	5.59	2.88E+05	3.55E+04	0.14	1.84E+06	3.34E+05	0.44	1.60E+07	6.57E+06	3.91
Luteolin-7-O-glucuronide	1.26E+07	2.34E+05	3.6	1.51E+07	1.26E+06	7.59	1.95E+07	1.65E+06	4.62	1.66E+07	5.28E+06	4.06
Luteolin-4'-O-glucoside	9.03E+06	7.74E+05	2.58	7.62E+06	4.12E+05	3.83	1.15E+07	1.37E+06	2.73	1.88E+07	2.90E+06	4.6
Chrysoeriol-5-O-glucoside	3.11E+06	1.95E+05	0.89	5.63E+06	3.43E+05	2.83	5.39E+06	6.47E+05	1.28	6.23E+06	5.69E+06	1.52
Luteolin-7-O-glucoside (Cynaroside)	3.24E+06	6.37E+05	0.93	2.30E+06	4.65E+05	1.15	4.93E+06	9.97E+05	1.17	5.06E+06	6.13E+05	1.24

(Continued)

TABLE 1 Continued

Compounds	8M		Ratio (%)	9M		Ratio (%)	10M		Ratio (%)	11M		Ratio (%)
	Average	STDEV		Average	STDEV		Average	STDEV		Average	STDEV	
Cyanidin-3-O-glucoside (Kuromanin)*	2.69E+06	3.62E+06	0.77	3.60E+05	1.24E+05	0.18	2.93E+07	1.76E+07	6.93	8.75E+05	5.97E+05	0.21
Cyanidin-3-O-galactoside*	3.81E+06	5.06E+06	1.09	3.92E+05	3.66E+05	0.2	3.54E+07	1.29E+07	8.38	1.10E+06	9.64E+05	0.27
Catechin-(7,8-bc)-4α-(3,4-dihydroxyphenyl)-dihydro-2-(3H)-one	1.05E+07	3.07E+06	3	2.91E+06	4.70E+05	1.46	8.22E+06	2.29E+06	1.95	2.98E+07	1.61E+07	7.28
Catechin	5.64E+06	9.66E+05	1.61	1.94E+06	1.44E+05	0.98	1.22E+07	5.02E+06	2.88	1.30E+07	3.53E+06	3.17
Epicatechin	4.69E+06	3.52E+05	1.34	1.16E+06	1.48E+05	0.58	9.35E+06	4.10E+06	2.21	6.91E+06	2.97E+06	1.69
Gallocatechin	8.70E+05	9.50E+05	0.25	1.40E+05	7.32E+04	0.07	6.18E+05	9.08E+05	0.15	4.94E+06	1.70E+06	1.21
Myricetin-3-O-glucuronide	1.97E+07	1.03E+07	5.63	2.34E+06	1.17E+06	1.17	1.03E+07	1.07E+07	2.44	3.95E+07	1.78E+07	9.65
Myricetin-3-O-glucoside	2.50E+06	1.66E+06	0.71	2.08E+05	6.68E+04	0.1	9.43E+05	1.11E+06	0.22	1.09E+07	6.90E+06	2.66
Hispidulin (5,7,4'-Trihydroxy-6-methoxyflavone)	9.73E+06	1.92E+06	2.78	2.70E+06	1.20E+06	1.36	4.32E+06	4.31E+04	1.02	9.04E+06	7.37E+06	2.21
Naringenin-7-O-glucoside (Prunin)	2.02E+06	1.56E+05	0.58	7.79E+05	1.15E+05	0.39	3.80E+06	1.94E+05	0.9	5.20E+06	2.50E+06	1.27
Naringenin (5,7,4'-Trihydroxyflavanone)*	1.14E+06	7.91E+04	0.32	2.24E+06	4.57E+05	1.13	1.20E+06	2.73E+04	0.28	9.43E+05	2.66E+05	0.23
Apigenin-6-C-glucoside (Isovitexin)	1.14E+06	2.55E+05	0.33	3.56E+06	3.99E+05	1.79	3.97E+06	2.69E+05	0.94	6.68E+05	2.29E+05	0.16
Tricin (5,7,4'-Trihydroxy-3',5'-dimethoxyflavone)	2.06E+07	3.35E+06	5.89	9.23E+06	1.55E+06	4.64	2.65E+07	2.62E+06	6.28	6.07E+06	8.91E+05	1.48
Total Ratio (%)			**84.76**			**81.95**			**86.37**			**84.43**

FIGURE 6
Heatmap showing correlations between functional genes and flavonoids in *C. paliurus*.

co-expressed at later stages of the process. Additionally, *VviMybC2-L1* and *VviMybC2-L3* are activated during two distinct phases of berry development. Their initial induction occurs towards the end of the green stage, suggesting a potential cooperative role in repressing proanthocyanidin synthesis. These genes are then reactivated during maturation, potentially inhibiting anthocyanin production at this later stage (Muñoz et al., 2019). Furthermore, the role of *bHLH* and *WD40* transcription factors, which often act as co-regulators in the *MYB-bHLH-WD40* complex, is also highlighted in this study. *TT8*, a member of the *bHLH* family, was strongly correlated with proanthocyanidin accumulation, further supporting its function in regulating later stages of the flavonoid biosynthetic pathway, particularly in response to seasonal changes. Studies in *Arabidopsis* have shown that *TT8* plays a critical role in regulating seed coat pigmentation and proanthocyanidin biosynthesis, further validating its importance in flavonoid biosynthesis (Baudry et al., 2004). Another significant finding is the negative correlation between *MYB4* and flavonoid content, particularly in late summer. This suggests that *MYB4* acts as a repressor of flavonoid biosynthesis, likely through the downregulation of *CHS* expression to prevent overaccumulation of flavonoids under high-light conditions. This regulatory function has been well-documented in other species, such as *Arabidopsis*, where *MYB4* negatively regulates flavonoid biosynthesis by repressing the expression of early biosynthetic genes (Jin et al., 2000).

Additionally, transcription factors such as *WRKY33* and *NAC* demonstrated significant correlations with flavonoid levels, particularly kaempferol glycosides. *WRKY33* is known to play a critical role in regulating plant responses to biotic and abiotic stresses, and its positive correlation with flavonoid levels suggests that it may upregulate flavonoid biosynthesis as part of the plant's defense mechanisms. Studies in *Oryza sativa* have similarly highlighted the role of *WRKY33* in enhancing flavonoid production under stress conditions, further supporting its function in *C. paliurus* (Jan et al., 2022). The correlation between *EIN3*, *ABI5*, and anthocyanin content suggests that hormonal signaling pathways,

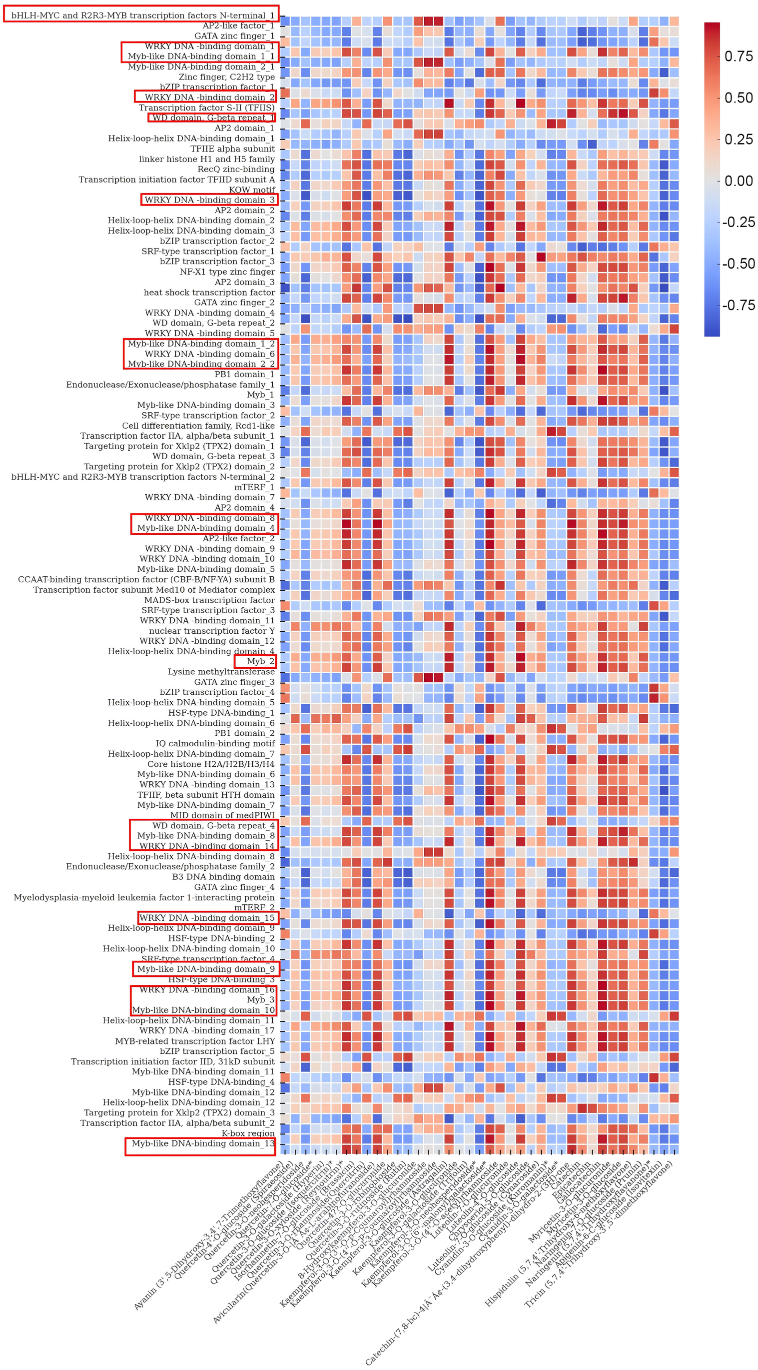


FIGURE 7
Heatmap illustrating the correlations between transcription factors and flavonoids in *C. paliurus*..

particularly ethylene and abscisic acid (ABA), are also involved in the regulation of flavonoid biosynthesis in *C. paliurus*. The positive correlations between these transcription factors and flavonoid content in late summer and early autumn indicate that they may play a role in modulating the plant transition to dormancy by regulating flavonoid production. Similar findings have been reported in *Glycine max*, where ethylene and ABA signaling pathways were shown to regulate flavonoid accumulation in response to environmental stresses (Jan et al., 2022).

4 Discussion

Our study on the seasonal dynamics of flavonoid biosynthesis in *C. paliurus* provides several new insights that advance the understanding of secondary metabolism in this species and contribute to the broader field of plant biology. While previous studies have documented the presence and seasonal variation of flavonoids in other species, our research uniquely combines metabolomic and transcriptomic analyses to elucidate the molecular mechanisms underlying these fluctuations in *C. paliurus*. This comprehensive approach has allowed us to identify not only the seasonal patterns in flavonoid accumulation but also the regulatory networks that drive these changes, thus offering a more complete picture of the plant's adaptive responses. The regulation of flavonoid biosynthesis in *C. paliurus* is intricately controlled by transcription factors, particularly those belonging to the *MYB* and *bHLH* families. These transcription factors are known to interact synergistically, often forming the *MYB-bHLH-WD40* (*MBW*) complex, which plays a crucial role in activating the expression of key structural genes in the flavonoid biosynthetic pathway, such as *DFR* and *ANS* (Xu et al., 2015; Hichri et al., 2010). The *MYB* factors provide specificity to the target genes, while *bHLH* factors enhance the transcriptional activation, ensuring robust flavonoid production in response to environmental and developmental cues. In our study, the upregulation of specific *MYB* and *bHLH* genes during periods of high flavonoid accumulation suggests that these transcription factors are central to the plant adaptive response to seasonal stressors, such as increased UV radiation and oxidative stress in late summer (Dubos et al., 2010). This regulatory mechanism is consistent with findings in other plant species, where the *MYB-bHLH* interaction has been shown to be critical for the production of flavonoids, anthocyanins, and other secondary metabolites. Understanding the specific roles of these transcription factors not only provides insight into the molecular control of flavonoid biosynthesis in *C. paliurus* but also highlights potential targets for genetic manipulation to enhance flavonoid content for agricultural and medicinal purposes.

Understanding the specific roles of these transcription factors not only provides insight into the molecular control of flavonoid biosynthesis in *C. paliurus* but also highlights potential targets for genetic manipulation to enhance flavonoid content for agricultural and medicinal purposes. In line with this, recent studies on *C. paliurus* have contributed to the growing body of knowledge on flavonoid regulation. For instance, the study by Caowen et al. (2023) explored the regulation of flavonoid biosynthesis for leaf coloring, identifying key genes involved in response to environmental factors. This study complements our findings by further illustrating how flavonoid biosynthesis is influenced by environmental conditions such as UV stress, highlighting its adaptive role in *C. paliurus* (Caowen et al., 2023). The whole-genome duplication event in *C. paliurus* also plays a significant role in its adaptive evolution, impacting the regulatory networks governing flavonoid biosynthesis. As Qu et al. (2023) noted, the duplication reshaped the genome, enhancing the plant ability to produce secondary metabolites like flavonoids. This genomic evidence supports our findings, suggesting that genetic evolution has contributed to the plant ability to modulate flavonoid production in response to seasonal changes (Qu et al., 2023). Environmental conditions and genetic similarity were shown to jointly influence the flavonoid variation patterns in *C. paliurus* leaves, as demonstrated by Sun et al. (2024). Their study confirmed that both genetic factors and environmental conditions contribute to flavonoid variation, and their findings align with our results, which highlight the importance of seasonal fluctuations in flavonoid accumulation (Sun et al., 2024).

An innovative aspect of our study is the application of co-expression network analysis to link specific gene modules with the biosynthesis of flavonoid compounds. This approach, which has been used in other plant systems such as *Arabidopsis thaliana*, has allowed us to identify gene networks that are strongly associated with the production of key flavonoids like kaempferol derivatives (Soubeyrand et al., 2021; Langfelder and Horvath, 2008). The tan module, in particular, was found to be correlated with the biosynthesis of these compounds, suggesting that targeted manipulation of this module could enhance flavonoid production in *C. paliurus*. The strong correlations observed between the expression of specific genes and the seasonal variation in flavonoid content in *C. paliurus* highlight the pivotal roles these genes play in the biosynthesis and regulation of flavonoids. The genes identified, particularly those coding for chalcone synthase (*CHS*), cytochrome P450 (*CYP73A100*), and other enzymes involved in flavonoid modification, demonstrate consistent expression patterns that align with the seasonal peaks in flavonoid accumulation. This regularity suggests that these genes are not only integral to the biosynthesis of key flavonoids but also responsive to environmental cues that influence their production. As such, these genes represent valuable targets for molecular breeding programs aimed at enhancing flavonoid content in *C. paliurus*, particularly for the development of varieties with improved medicinal and nutritional properties. By leveraging these genes as molecular markers, it may be possible to select and cultivate *C. paliurus* plants that consistently produce high levels of beneficial flavonoids across different growing seasons, thereby optimizing their potential use in nutraceutical and therapeutic applications. This finding not only advances our understanding of the genetic regulation of flavonoid biosynthesis but also opens up new possibilities for breeding or engineering plants with enhanced levels of these valuable metabolites. The broader implications of our findings are significant for both basic plant science and applied research. The detailed characterization of flavonoid biosynthesis in *C. paliurus* provides a valuable resource for the development of functional foods and medicinal products. Flavonoids are well-known for their health-promoting properties, including antioxidant, anti-inflammatory, and anticancer activities (Gould and Lister, 2006). By identifying the specific genes and regulatory networks that control flavonoid production, our research lays the groundwork for future efforts to enhance these compounds in *C. paliurus* and other economically important plants. Additionally, understanding the seasonal dynamics of flavonoid production could inform optimal harvesting strategies to maximize the yield of these bioactive compounds.

Despite the significant contributions of our study, there are several areas that warrant further investigation. For instance, while we have identified key regulatory genes involved in flavonoid biosynthesis, the specific environmental triggers that modulate their expression remain to be fully understood. Future research could explore the roles of factors such as light intensity, temperature, and water availability in regulating these biosynthetic pathways (Falcone Ferreyra et al., 2012). Moreover, it would be beneficial to extend this research to other parts of the *C. paliurus* plant, such as the bark and roots, to gain a more comprehensive understanding of flavonoid biosynthesis throughout the plant. Finally, given the potential applications of our findings in agriculture and medicine, further studies could explore the feasibility of manipulating the identified gene modules to enhance flavonoid production in other plant species. Our research provides a detailed and comprehensive analysis of the seasonal dynamics of flavonoid biosynthesis in *C. paliurus*. By linking gene expression patterns with flavonoid production, our study not only advances our understanding of plant secondary metabolism but also offers new strategies for enhancing the nutritional and medicinal value of this and other plant species. The innovative methodologies employed, particularly the use of co-expression network analysis, could serve as a model for future studies aimed at unraveling the complex regulatory networks that govern secondary metabolite biosynthesis in plants.

In this study, the Pearson correlation analysis between transcription factor expression and flavonoid content across different months revealed key insights into the role of various transcription factors in the regulation of flavonoid biosynthesis. This analysis underscores the significant seasonal variation in transcription factor activity and its direct impact on flavonoid accumulation, providing a molecular basis for understanding how *C. paliurus* adapts its secondary metabolism to environmental changes. Firstly, members of the *MYB* family, particularly *MYB11, MYB12*, and *MYB111*, exhibited strong positive correlations with flavonols such as quercetin-3-O-glucoside and kaempferol-3-O-glucoside, especially in the late summer and early autumn months. This suggests that these transcription factors play crucial roles in regulating flavonol synthesis during periods of environmental stress, such as increased UV radiation. These findings align with previous research on Arabidopsis thaliana, where these *MYB* transcription factors are known to promote flavonol production under UV light exposure (Stracke et al., 2007). Additionally, *MYB75* (*PAP1*) and *MYB90* (*PAP2*) demonstrated strong correlations with anthocyanin content, particularly cyanidin-3-O-glucoside, suggesting their involvement in anthocyanin biosynthesis during periods of high light intensity, consistent with findings in *Vitis vinifera* (Kobayashi et al., 2004). Furthermore, *bHLH* and *WD40* transcription factors showed strong positive correlations with both flavonols and anthocyanins, indicating their likely involvement in the formation of the *MYB-bHLH-WD40* complex, which regulates the entire flavonoid biosynthesis pathway. This complex has been well-characterized in other species, such as *Petunia hybrida*, where it plays a critical role in anthocyanin biosynthesis (Quattrocchio et al., 2006)). Similarly, *TT8*, a *bHLH* transcription factor, showed a significant positive correlation with proanthocyanidin content during the early autumn period, consistent with its role in *Arabidopsis* seed development, where it regulates proanthocyanidin synthesis (Baudry et al., 2004). Interestingly, *MYB4* exhibited a negative correlation with flavonoid content, particularly in late summer, suggesting that it may act as a repressor of flavonoid biosynthesis by downregulating the expression of chalcone synthase (*CHS*). This aligns with previous studies in *Arabidopsis*, where *MYB4* has been shown to repress *CHS* expression when flavonoid levels are sufficiently high (Jin et al., 2000). Additionally, *MYC2*, a *bHLH* transcription factor involved in stress response, was positively correlated with quercetin glycosides, indicating that it plays a role in enhancing flavonoid biosynthesis as part of the plant defense against environmental stress. Similar regulatory functions have been observed in *Arabidopsis*, where *MYC2* controls jasmonate-induced flavonoid biosynthesis under stress conditions (Dombrecht et al., 2007). *WRKY33*, a *WRKY* transcription factor, showed a strong positive correlation with kaempferol glycosides, especially during periods of environmental stress, indicating its role in upregulating flavonoid biosynthesis as part of the plant defense response. This is consistent with studies in *Oryza sativa*, where *WRKY* has been implicated in stress-induced flavonoid production (Jan et al., 2022). Additionally, members of the *NAC* family of transcription factors were moderately correlated with flavonoid content, suggesting they may coordinate developmental and metabolic processes during seasonal transitions, similar to their role in regulating flavonoid biosynthesis in *Camellia sinensis* (Song et al., 2024). Lastly, *EIN3* (Ethylene-Insensitive 3) and *ABI5* (Abscisic Acid Insensitive 5) demonstrated positive correlations with flavonoid content, particularly with anthocyanins, during late summer and early autumn. This suggests that ethylene and abscisic acid (ABA) signaling pathways are important regulators of flavonoid biosynthesis as the plant transitions towards dormancy. Similar functions for *EIN3* and *ABI5* have been observed in *Cucumic sativus* (Wu et al., 2022) and *Vitis vinifera* (Dong et al., 2020), where these transcription factors regulate flavonoid accumulation in response to environmental stress and developmental cues. These results not only validate previous research but also provide novel insights into the seasonal regulation of flavonoid biosynthesis in *C. paliurus*.

By integrating transcriptomic and metabolomic data, this study presents a comprehensive view of the molecular mechanisms governing flavonoid production in *C. paliurus*, highlighting the critical role of transcription factors in adapting to seasonal changes. These findings provide a valuable foundation for future genetic manipulation and breeding strategies aimed at enhancing flavonoid content in *C. paliurus*, and they extend our understanding of flavonoid regulation beyond model organisms to a commercially and medicinally important species. However, we acknowledge a key limitation in our study, it is that while we identified correlations between the expression of transcription factors and flavonoid levels, we did not perform direct validation of their roles. This gap represents an important area for future investigation. Despite this limitation, our study lays the groundwork for future research that can integrate more robust validation techniques. By exploring additional omics approaches and validating the transcriptional regulation of key genes, future studies will be able to refine our understanding of the molecular networks controlling flavonoid production in *C. paliurus*. This will be crucial for advancing genetic and breeding efforts to optimize flavonoid yields and enhance the medicinal potential of this species.

5 Conclusion

This study explored the seasonal dynamics of flavonoid biosynthesis in *C. paliurus*, integrating transcriptomic and metabolomic analyses to uncover the molecular mechanisms driving these variations. Our findings show that flavonoid content peaks in late summer, driven by the upregulation of key biosynthetic genes and transcription factors, particularly *MYB* and *bHLH* families, which regulate flavonoid production in response to environmental stress. Co-expression network analysis identified gene modules linked to specific flavonoids, such as kaempferol derivatives, highlighting potential targets for genetic enhancement of flavonoid content. The Pearson correlation analysis further confirmed the central role of transcription factors like *MYB11, MYB12, MYB111, MYB75*, and *bHLH* in regulating flavonol and anthocyanin biosynthesis. Seasonal changes in these factors underscore their role in the plant's response to environmental changes, enhancing our understanding of the genetic control mechanisms behind flavonoid production. By integrating transcriptomic and metabolomic data, this study provides a comprehensive view of flavonoid biosynthesis regulation across seasons. These insights not only broaden our understanding of plant secondary metabolism but also suggest new opportunities for optimizing flavonoid production through targeted genetic approaches. This research lays a foundation for future efforts to enhance flavonoid content in *C. paliurus* and other economically and medicinally important plants, with potential applications in functional foods, medicine, agriculture, and biotechnology.

Data availability statement

Raw sequencing data for RNA-seq have been deposited and is available in the BIG Sub system under BioProject accession number CRA005839 (https://ngdc.cncb.ac.cn/gsa).

Author contributions

DC: Conceptualization, Formal Analysis, Investigation, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. YX: Formal Analysis, Investigation, Software, Writing – original draft, Data curation, Resources, Validation. XZ: Data curation, Formal Analysis, Investigation, Methodology, Resources, Software, Validation, Writing – original draft. HS: Data curation, Formal Analysis, Resources, Validation, Writing – original draft. CZ: Data curation, Formal Analysis, Resources, Validation, Writing – original draft. JZ: Conceptualization, Methodology, Resources, Supervision, Writing – original draft. TX: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Software, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (Grant No. 32471847).

Conflict of interest

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

Generative AI statement

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1525226/full#supplementary-material

References

Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* 39, 366–380. doi: 10.1111/j.1365-313X.2004.02138.x

Caowen, S., Shengzuo, F., and Xulan, S. (2023). Flavonoid biosynthesis regulation for leaf coloring of *Cyclocarya paliurus*. *Acta Physiol. Plant* 45, 91. doi: 10.1007/s11738-023-03571-2

Chen, Z., Jian, Y., Wu, Q., Wu, J., Sheng, W., Jiang, S., et al. (2022). *Cyclocarya paliurus* (Batalin) Iljinskaja: Botany, Ethnopharmacology, phytochemistry and pharmacology. *J. Ethnopharmacol* 285, 114912. doi: 10.1016/j.jep.2021.114912

Dombrecht, B., Xue, G. P., Sprague, S. J., Kirkegaard, J. A., Ross, J. J., Reid, J. B., et al. (2007). MYC2 differentially modulates diverse jasmonate-dependent functions in *arabidopsis*. *Plant Cell* 19, 2225–2245. doi: 10.1105/tpc.106.048017

Dong, T., Zheng, T., Fu, W., Guan, L., Jia, H., and Fang, J (2020). The effect of ethylene on the color change and resistance to Botrytis cinerea infection in 'Kyoho' grape fruits. *Foods* 9, 892. doi: 10.3390/foods9070892

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Falcone Ferreyra, M. L., Rius, S. P., and Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3. doi: 10.3389/fpls.2012.00222

Fu, X., Zhou, X., Deng, B., Shang, X., and Fang, S. (2015). Seasonal and genotypic variation of water-soluble polysaccharide content in leaves of *Cyclocarya paliurus*. *South. Forests: J. For. Sci.* 77, 231–236. doi: 10.2989/20702620.2015.1010698

Gould, K., and Lister, C. (2006). "Flavonoid functions in plants," in *Flavonoids: chemistry, biochemistry and applications*. Ed. K.R.M.Ø. M. Andersen (Boca Raton, USA: CRC Press LLC), 397–441.

Hichri, I., Heppel, S. C., Pillet, J., Léon, C., Czemmel, S., Delrot, S., et al. (2010). The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. *Mol. Plant* 3, 509–523. doi: 10.1093/mp/ssp118

Hussain, T., Murtaza, G., Yang, H., Kalhoro, M. S., and Kalhoro, D. H. (2020). Exploiting anti-inflammation effects of flavonoids in chronic inflammatory diseases. *Curr. Pharm. Des.* 26, 2610–2619. doi: 10.2174/1381612826666200408101550

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003

Jan, R., Asaf, S., Lubna,, Asif, S., Kim, E.-G., Jang, Y.-H., et al. (2022). Enhancing the expression of the *osF3H* gene in *oryza sativa* leads to the regulation of multiple biosynthetic pathways and transcriptomic changes that influence insect resistance. *Int. J. Mol. Sci.* 23, 15308. doi: 10.3390/ijms232315308

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., et al. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *EMBO J.* 19, 6150–6161. doi: 10.1093/emboj/19.22.6150

Kobayashi, S., Goto-Yamamoto, N., and Hirochika, H. (2004). Retrotransposon-induced mutations in grape skin color. *Science* 304, 982–982. doi: 10.1126/science.1095011

Kumar, S., and Pandey, A. K. (2013). Chemistry and biological activities of flavonoids: an overview. *ScientificWorldJournal* 2013, 162750. doi: 10.1155/2013/162750

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

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

Lin, Y., Chen, P., Zhou, M., Shang, X., and Fang, S. (2020). Key bioactive substances and their antioxidant activities in *Cyclocarya paliurus* (Batal.) Iljinskaja leaves collected from natural populations. *J. Nanjing Forestry Univ.* 44, 10. doi: 10.3969/j.issn.1000-2006.201901045

Liu, Y., Chen, P., Zhou, M., Wang, T., Fang, S., Shang, X., et al. (2018). Geographic variation in the chemical composition and antioxidant properties of phenolic compounds from *cyclocarya paliurus* (Batal) iljinskaja leaves. *Molecules* 23, 2440. doi: 10.3390/molecules23102440

Liu, Y., Qian, C., Ding, S., Shang, X., Yang, W., and Fang, S. (2016). Effect of light regime and provenance on leaf characteristics, growth and flavonoid accumulation in Cyclocarya paliurus(Batal) Iljinskaja coppices. *Bot. Stud.* 57, 28. doi: 10.1186/s40529-016-0145-7

Luo, N., Ding, W., Wu, J., Qian, D., Li, Z., Qian, Y., et al. (2013). UPLC-Q-TOF/MS coupled with multivariate statistical analysis as a powerful technique for rapidly exploring potential chemical markers to differentiate between radix paeoniae alba and radix paeoniae rubra. *Natural Product Commun.* 8, 1934578X1300800421. doi: 10.1177/1934578X1300800421

Ma, X., Hou, Y., Umar, A. W., Wang, Y., Yu, L., Ahmad, N., et al. (2024). Safflower *ctFLS1*-induced drought tolerance by stimulating the accumulation of flavonols and anthocyanins in *arabidopsis thaliana*. *Int. J. Mol. Sci.* 25, 5546. doi: 10.3390/ijms25105546

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430

Mohammadi, A., Sani, T. A., Ameri, A. A., Imani, M., Golmakani, E., and Kamali, H. (2014). Seasonal variation in the chemical composition, antioxidant activity, and total phenolic content of *Artemisia absinthium* essential oils. *Pharmacognosy Res.* 7, 329–334. doi: 10.4103/0974-8490.158441

Muñoz, C. J., Fanzone, M. L., and Lijavetzky, D. C. (2019). Transcriptional regulation of the anthocyanin biosynthesis pathway in developing grapevine berries in cultivar "Malbec" by putative R2R3 MYB negative regulators. *Scientia Horticulturae*. 257, 108663. doi: 10.1016/j.scienta.2019.108663

Qu, Y., Shang, X., Zeng, Z., Yu, Y., Bian, G., Wang, W., et al. (2023). Whole-genome duplication reshaped adaptive evolution in A relict plant species, *cyclocarya paliurus*. *Genomics Proteomics Bioinf.* 21, 455–469. doi: 10.1016/j.gpb.2023.02.001

Quattrocchio, F., Verweij, W., Kroon, A., Spelt, C., Mol, J., and Koes, R. (2006). PH4 of petunia is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway. *Plant Cell* 18, 1274–1291. doi: 10.1105/tpc.105.034041

Richard, S., Lapointe, G., Rutledge, R. G., and Séguin, A. (2000). Induction of chalcone synthase expression in white spruce by wounding and jasmonate. *Plant Cell Physiol.* 41, 982–987. doi: 10.1093/pcp/pcd017

Seehaus, K., and Tenhaken, R. (1998). Cloning of genes by mRNA differential display induced during the hypersensitive reaction of soybean after inoculation with Pseudomonas syringae pv. glycinea. *Plant Mol. Biol.* 38, 1225–1234. doi: 10.1023/A:1006036827841

Shang, X. L., Wu, Z. F., Yin, Z. Q., Zhang, J., Liu, Z. J., and Fang, S. Z. (2015). Simultaneous determination of flavonoids and triterpenoids in *Cyclocarya paliurus* leaves using high-performance liquid chromatography. *Afr. J. Traditional Complementary Altern. Medicines* 12, 125–134. doi: 10.4314/ajtcam.v12i3.16

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Sheng, X., Chen, H., Wang, J., Zheng, Y., Li, Y., Jin, Z., et al. (2021). Joint transcriptomic and metabolic analysis of flavonoids in *cyclocarya paliurus* leaves. *ACS Omega* 6, 9028–9038. doi: 10.1021/acsomega.1c00059

Song, S., Ran, W., Gao, L., Wang, Y., Lv, W.-Y., Tao, Y., et al. (2024). A functional study reveals CsNAC086 regulated the biosynthesis of flavonols in *Camellia sinensis*. *Planta* 259, 147. doi: 10.1007/s00425-024-04426-x

Soubeyrand, E., Latimer, S., Bernert, A. C., Keene, S. A., Johnson, T. S., Shin, D., et al. (2021). 3-*O*-glycosylation of kaempferol restricts the supply of the benzenoid precursor of ubiquinone (Coenzyme Q) in *Arabidopsis thaliana*. *Phytochemistry* 186, 112738. doi: 10.1016/j.phytochem.2021.112738

Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrtens, F., Niehaus, K., et al. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* 50, 660–677. doi: 10.1111/j.1365-313X.2007.03078.x

Sun, C., Cao, Y., Li, X., Fang, S., Yang, W., and Shang, X. (2024). The impact of genetic similarity and environment on the flavonoids variation pattern of *Cyclocarya paliurus*. *Sci. Rep.* 14, 24187. doi: 10.1038/s41598-024-74957-6

Sun, C., Zhou, Y., Fang, S., and Shang, X. (2021). Ecological gradient analysis and environmental interpretation of *cyclocarya paliurus* communities. *Forests* 12, 146. doi: 10.3390/f12020146

Torres, C. A., Azocar, C., Ramos, P., Pérez-Díaz, R., Sepulveda, G., and Moya-León, M. A. (2020). Photooxidative stress activates a complex multigenic response integrating the phenylpropanoid pathway and ethylene, leading to lignin accumulation in apple (*Malus domestica* Borkh.) fruit. *Horticulture Res.* 7, 22. doi: 10.1038/s41438-020-0244-1

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi: 10.1093/bioinformatics/btp120

Wan, K., Ban, J., Yang, F., Zhang, X., Huang, X., Wang, Y., et al. (2024). Transcriptomic analysis reveals the flavonoid biosynthesis pathway involved in rhizome development in *polygonatum cyrtonema* hua. *Plants* 13, 1524. doi: 10.3390/plants13111524

Wang, T., Cao, H., Du, T., Meng, D., Yang, Q., Li, J., et al. (2024). Integrated transcriptome and metabolome analysis revealed the key role of the flavonoid biosynthesis in olive defense against *alternaria alternata*. *Physiol. Plant* 176, e14529. doi: 10.1111/ppl.14529

Winkel-Shirley, B. (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* 126, 485–493. doi: 10.1104/pp.126.2.485

Wu, P., Kong, Q., Bian, J., Ahammed, G. J., Cui, H., Xu, W., et al. (2022). Unveiling molecular mechanisms of nitric oxide-induced low-temperature tolerance in cucumber by transcriptome profiling. *Int. J. Mol. Sci.* 23, 5615. doi: 10.3390/ijms23105615

Xie, J.-H., Dong, C., Nie, S.-P., Li, F., Wang, Z.-J., Shen, M.-Y., et al. (2015). Extraction, chemical composition and antioxidant activity of flavonoids from *Cyclocarya paliurus* (Batal.) Iljinskaja leaves. *Food Chem.* 186, 97–105. doi: 10.1016/j.foodchem.2014.06.106

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

Xu, P., Li, M., Ma, C., Li, X., Bai, P., Lin, A., et al. (2024). Loss-of-function mutation in anthocyanidin reductase activates the anthocyanin synthesis pathway in strawberry. *Mol. Hortic.* 4, 33. doi: 10.1186/s43897-024-00106-2

Zhang, J., Huang, N., Lu, J.-C., Li, X., Wang, Y.-H., Yang, L.-M., et al. (2020). Water-soluble phenolic compounds and their anti-HIV-1 activities from the leaves of *Cyclocarya paliurus*. *J. Food Drug Anal.* 18, 2. doi: 10.38212/2224-6614.2230

Zhao, W. (2020). Inducement and cultivation of novel red *Cyclocarya paliurus* callus and its unique morphological and metabolic characteristics. *Industrial Crops Prod.* 147, 112266. doi: 10.1016/j.indcrop.2020.112266

Zhao, T., Huang, C., Li, N., Ge, Y., Wang, L., Tang, Y., et al. (2024). Ubiquitin ligase VvPUB26 in grapevine promotes proanthocyanidin synthesis and resistance to powdery mildew. *Plant Physiol.* 195, 2891–2910. doi: 10.1093/plphys/kiae249

Zheng, X., Xiao, H., Chen, J., Zhu, J., Fu, Y., Ouyang, S., et al. (2022). Metabolome and whole-transcriptome analyses reveal the molecular mechanisms underlying hypoglycemic nutrient metabolites biosynthesis in *cyclocarya paliurus* leaves during different harvest stages. *Front. Nutr.* 9. doi: 10.3389/fnut.2022.851569

Zheng, X., Xiao, H., Su, J., Chen, D., Chen, J., Chen, B., et al. (2021). Insights into the evolution and hypoglycemic metabolite biosynthesis of autotetraploid *Cyclocarya paliurus* by combining genomic, transcriptomic and metabolomic analyses. *Industrial Crops Prod.* 173, 114154. doi: 10.1016/j.indcrop.2021.114154

Zheng, X.-R., Zhang, M.-J., Shang, X.-L., Fang, S.-Z., and Chen, F.-M. (2020). Stem Canker on *Cyclocarya paliurus* Is Caused by Botryosphaeria dothidea. *Plant Dis.* 104, 1032–1040. doi: 10.1094/PDIS-11-18-1990-RE

Zhou, M., Lin, Y., Fang, S., Liu, Y., and Shang, X. (2019). Phytochemical content and antioxidant activity in aqueous extracts of *Cyclocarya paliurus* leaves collected from different populations. *PeerJ* 7, e6492. doi: 10.7717/peerj.6492

Zou, R., Wu, S., Jiao, S., and Yu, Z. (2018). Research progress on content variation of main functional compounds in *Cyclocarya paliurus* Liaoning. *J. Tradit Chin. Med.* 8, 1782–1785. doi: 10.13192/j.issn.1000-1719.2018.08.071

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 27 March 2025
DOI 10.3389/fpls.2025.1508454

OPEN ACCESS

EDITED BY
Yongjun Wei,
Zhengzhou University, China

REVIEWED BY
Sharad Vats,
Banasthali University, India
Sebastian John Adams,
University of Mississippi, United States

*CORRESPONDENCE
Aditi Kundu
✉ aditi.kundu@icar.gov.in

RECEIVED 09 October 2024
ACCEPTED 27 January 2025
PUBLISHED 27 March 2025

CITATION
Neel S, Mandal A, Saha S, Das A, Kundu A and Singh A (2025) *Gymnema sylvestre* saponins for potential antifungal action: *in vitro* and *in silico* perspectives.
Front. Plant Sci. 16:1508454.
doi: 10.3389/fpls.2025.1508454

Gymnema sylvestre saponins for potential antifungal action: *in vitro* and *in silico* perspectives

Shila Neel[1,2], Abhishek Mandal[3], Supradip Saha[1], Amrita Das[4], Aditi Kundu[1]* and Anupama Singh[1]

[1]Division of Agricultural Chemicals, ICAR-Indian Agricultural Research Institute, New Delhi, India, [2]The Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi, India, [3]Division of Basic Sciences, ICAR-Indian Institute of Horticultural Research, Bengaluru, India, [4]Division of Plant Pathology, ICAR-Indian Agricultural Research Institute, New Delhi, India

Saponins are responsible for a wide range of biological activities, which is why the present research is focused on the chemical profiling of saponins and other metabolites from *Gymnema sylvestre* leaves for their potential efficacy in managing pathogenic fungi. Leaves of the plant was extracted with chloroform to obtain crude saponin concentrates. Characterizations of the chloroform soluble fraction of the leaves [chloroform extract of *G. sylvestre* (CGS)] in ultra-performance liquid chromatography–quadrupole time of flight–electrospray ionization–tandem mass spectrometry (UPLC-QToF-ESI-MS/MS) displayed 23 metabolites, primarily comprising of saponins and other minor phytocomponents. Among these, two major saponins, gymnemic acid IV and gymnestrogenin, were isolated, purified, and characterized using ^{1}H-NMR, ^{13}C-NMR, and high-resolution mass spectrometry (HRMS). *In vitro* fungistatic efficacy showed the highest effectiveness against *Penicillium digitatum* 6952 (EC_{50} 297.2 µg/mL), followed by *Penicillium expansum* 2995 (360.5 µg/mL) and *Aspergillus flavus* 6678 (369.4 µg/mL). Furthermore, the mechanism of interaction of these metabolites to inhibit cyt P_{450} sterol 1,4-α-demethylase was determined by *in vitro* and *in silico* molecular modeling analysis, explaining the probable reason for the reduction in ergosterol content in the treated fungi. *In silico* analysis suggested the highest binding efficiency of gymnemic acid IV due to the lowest binding energy, specifically interacted through conventional H-bonds, hydrophobic π-alkyl, π-π, and π-sigma interactions. Indeed, the valuable findings of the study would be useful for further development of Gymnema saponin based biopesticidal products.

KEYWORDS

phytochemicals, triterpenoids, ergosterol inhibition, phytopathogen, gymnemic acids, Gymnema saponins

Introduction

Over the past two decades, there has been growing interest in exploring plant-derived bioactive compounds, commonly called phyto-metabolites, for their potential use in pharmaceuticals, nutraceuticals, and sustainable agriculture (Gurnani et al., 2014; Wadkar et al., 2008). These bioactive compounds, including alkaloids, flavonoids, saponins, terpenoids, and phenolics, are produced by plants as part of their defense mechanism against different phytopathogens, pests, and environmental stresses. Additionally, they offer significant potential in managing many crop diseases in field and storage, thereby improving post-harvest management practices (Muddapur et al., 2024). It was estimated that 35%–55% of crop loss occurred due to post-harvest diseases and mishandling (Pillai et al., 2024). Phytometabolites are found to be highly effective in controlling several notorious plant pathogens, particularly those that cause post-harvest damage to fresh produce (Srinivasan and Kumaravel, 2016). Traditionally, the management of these pathogens has relied heavily on the application of synthetic fungicides, which have proven effective but come with significant drawbacks. Continuous use of these chemical fungicides has led to issues such as pathogen resistance and the accumulation of toxic residues in food products and the environment (Devi et al., 2021; Jung et al., 2006). However, plant-based metabolites provide a promising solution, offering natural fungicidal properties with complex modes of action without the associated toxicity, resistance development, and environmental hazards (Khursheed et al., 2022).

Research on plant metabolites revealed that many of these natural products exhibit strong antifungal activity, making them viable candidates for use as biopesticides (Patil et al., 2011). Plants belonging to the Apocynaceae family include approximately 2,000 species across 250 genera, distributed across the tropical and subtropical regions (Tiwari et al., 2014; Muddapur et al., 2024). *Gymnema sylvestre*, commonly called "Gurmar" in local parlance, is a prominent herb in India, distributed widely across Konkan and the Western Ghats, as well as the Deccan peninsula and Western and Northern India (Sharma et al., 2024; Dar et al., 2024; Parveen et al., 2019). The plant contains bioactive saponins, particularly gymnemic acid derivatives, which are considered valuable compounds responsible for bioactivity. *Gymnema*-based remedies have attracted the attention of scientists due to their therapeutic efficacy in the area of alternative and complementary medicine (Romaiyan et al., 2023).

Evidence-based studies documented *G. sylvestre* phytomolecules as powerful phytochemicals capable of combating many diseases and ailments. The leaves contain bioactive triterpenoid saponins. The primary phyto-components of the plant comprise various gymnemic acid saponins and gymnemasides (Arora and Sood, 2017; Vats et al., 2023; Rahman and Husen, 2023; Sagheer et al., 2023). Gymnemic acid analogs were extracted from *G. sylvestre* leaves and analyzed using liquid chromatographic and mass spectrometric methods (Kamble et al., 2013; Chen and Guo, 2017; Dar et al., 2024). Furthermore, a building block strategy was made to identify the oleanane triterpenoids of *G. sylvestre* using ultra-performance liquid chromatography–quadrupole time of flight–mass spectrometry (UPLC-QToF/MS) (Pham et al., 2019). Previous phytochemical investigations on different plant parts led to the isolation of certain triterpenoid saponins, flavonoids, and their glycosides, displaying a broad spectrum of biological activities (Thangavelu et al., 2012; Ahamad et al., 2018). Even the production of gymnemic acid from the suspension culture of *G. sylvestre* was investigated by Chodisetti et al. (2013). Recently, the saponin extraction process from leaves of *Gymnema* using advanced extraction techniques was optimized (Saeed et al., 2022).

Phyto-pharmacological properties of *G. sylvestre* were reported in the literature (Khan et al., 2019). In addition to antidiabetic action, the potential antimicrobial activity of the plant metabolites and their exploitation in the management of fungal organisms were also reported (Satdive et al., 2003; Ibrahim et al., 2017; Suwan et al., 2022). Significant antifungal and antibacterial activities of the saponin fractions extracted from the plant were also reported by Gopiesh Khanna and Kannabiran (2008). Again, the antiviral action of the *Gymnema* metabolites, particularly gymnemagenol, was reported in the literature (Gopiesh Khanna et al., 2011).

Extensive studies are being attempted to isolate and identify various bioactive saponins of *G. sylvestre* leaves for their ethnopharmacological and medicinal properties; however, comprehensive metabolomics profiling and characterizations of the bioactive fraction of the plant are still lacking. Furthermore, limited information is available on the antifungal action of metabolites of *G. sylvestre* such as triterpenoids, flavonoids, and steroids against the important decay-causing fungi such as *Penicillium* sp. and *Aspergillus* sp. Furthermore, a study related to the mechanism of interaction of *Gymnema* metabolites with the fungal proteins and their possible effect on fungal biology is also missing. Therefore, the present investigation emphasized the characterization of bioactive metabolites comprising saponins and other phytochemicals from *G. sylvestre* leaves for potential fungicidal action against fungi (*Penicillium* sp. and *Aspergillus* sp.) causing fungal spoilage under storage of fresh produce. Additionally, the interaction of these metabolites with the target-specific fungal protein, responsible for sterol biosynthesis, was also targeted using molecular modeling analysis.

Material and methods

Plant sample

Fresh leaves of *G. sylvestre* (2.0 kg) were collected from Delhi, India (28.65°N, 77.22°E) in July 2022. Leaves were cleaned with distilled water, shade-dried for 5 days, and ground into a coarse powder. The powdered sample was stored in an airtight container under refrigeration. The moisture content of the sample was 5.9% ± 0.7% (w/w).

Chemicals and instruments

Solvents and media were procured from Merck® India Ltd. (Mumbai, India) and used without further purification. Bioactive metabolites of *G. sylvestre* leaves were purified through silica gel (60–120 mesh)-loaded column chromatography sourced from Merck® India Ltd. Analytical grade solvents (LC-MS grade), HSGF$_{254}$ silica gel-coated Thin Layer Chromatography (TLC) plates, and Preparative-Thin Layer Chromatography (TLC) glass plates (0.4–0.5 mm) pre-coated with silica gel GF$_{254}$ were also obtained from Merck® India Ltd.

Nuclear magnetic resonance (NMR) spectra were obtained using a JEOL 400-MHz NMR spectrometer, with tetramethylsilane (TMS) serving as the internal standard. UPLC–QToF–electrospray ionization–tandem mass spectrometry (UPLC-QToF-ESI-MS/MS) analysis was performed on an Acquity UPLC system, connected to a quadrupole time of flight mass spectrometer (QToF-MS/MS, Xevo G2-XS system, Waters Corporation, Wilmslow, UK).

Extraction of leaves

Conventional solid–liquid extraction

Phytochemical constituents from finely grounded powder (2.0 kg) of *G. sylvestre* leaves were extracted by high-speed homogenization (IKA® India Private Limited, Bengaluru, India) (Dutta et al., 2021). First, the leaf powder was submerged in ethanol (20%) and homogenized for 30 min at room temperature. Then, the hydroalcoholic (20%) extract was partitioned with chloroform (1.5 L) thrice, and the $CHCl_3$ soluble fraction was filtered and evaporated below 40°C using a flash evaporator (Heidolph, Schwabach, Germany) to obtain $CHCl_3$ extract [chloroform extract of *G. sylvestre* (CGS)].

Phytochemical characterizations

Fourier transform infrared spectroscopy

Functional groups of CGS were characterized using Fourier transform infrared spectroscopy (FTIR) (Bruker, Dresden, Germany). The FT-IR spectrum was generated by measuring the transmission percentage against the wave number ranging from 4,000 to 600 cm^{-1}. For analysis, CGS (50 mg) was mixed with 100 mg of potassium bromide and pressed to prepare a pellet. The signals at the characteristic wavenumber indicated specific functional groups of the metabolites present in CGS.

Ultra performance liquid chromatography -quadrapole time of flight- electrospray ionization-mass (UPLC-QToF-ESI-MS)

For metabolomics analysis, a mass spectrometer (Waters Xevo-G2-XS, QToF) paired with an Acquity H Class Plus UPLC was used. The untargeted metabolomics analysis was conducted in positive electrospray ionization mode under the MS^E module with an m/z range of 100–1,500. QToF-ESI-MS analysis was operated using the MassLynx 4.2 software. The system captured full-scan MS data (6 V, low energy) and MS/MS data (ranging from 10.0 to 60.0 V, high energy) in the MS^E mode. Key parameters included a capillary voltage of 3.0 kV, a sampling cone voltage of 30 V, an extraction cone voltage of 4.5 V, a desolvation temperature of 300°C, a source temperature of 150°C, a desolvation gas flow of 800 L/h, and a cone gas flow of 30 L/h. Mass correction was achieved with a lock spray using leucine encephalin (m/z 556.2771 in positive mode) at a flow rate of 10 L/min and a concentration of 1 μg/mL every 10 seconds. Metabolites were separated using an Acquity UPLC BEH C_{18} column (2.1 mm × 100 mm, 1.8 μm) with a gradient elution system of phase A (acetonitrile with 0.1% formic acid) and phase B (water with 0.1% formic acid) at 0.3 mL/min. The gradient profile was as follows: 0–2.5 min/100% A, 2.5–4.5 min/70% A, 4.5–20.0 min/10% A, 20.0–23.0 min/5% A, and 23.0–25.0 min/100% A. The raw data were processed using the UNIFI software version 1.7, following SANTE guidelines, identifying metabolites with a mass error limit of 3 μg/mL (Munjanja, 2017; Dutta et al., 2021).

Isolation of pure compounds

Column chromatography was employed for the separation and isolation of compounds from CGS. This technique, known for its simplicity and effectiveness, is widely used for the isolation of chemical constituents from complex mixtures. CGS (20 g) was subjected to silica gel column chromatography (60–120 mesh size, pre-activated at 110°C) using a gradient of hexane:ethyl acetate (9:1, 8:2, 7:3, 6:4, 1:1, 4:6, 3:7, 2:8, and 0:1 v/v) as the eluent, resulting in 78 fractions. Fractions 35–59, which showed three similar spots on a TLC plate, were combined and re-chromatographed using hexane: ethyl acetate (95:5 v/v) followed by preparative TLC to isolate and crystals of CGS-1 (131 mg) and CGS-2 (57 mg).

Gymnemic acid IV

Lemon-yellowish solid; selected ^1H-NMR (400 MHz, $CDCl_3$): δ 12.16 (1H, s, H-6″), 6.92–6.97 (1H, m, H-1′), 5.57 (1H, s, H-12), 4.49 (1H, t, J = 2.9 Hz, H-16), 4.31–4.36 (1H, m, H-3), 4.46 (1H, d, J = 4.8 Hz, H-5″), 4.52 (1H, d, J = 7.5 Hz, H-1″), 5.44 (1H, d, J = 8.6 Hz, H-21), 4.09 (1H, dd, J = 1.2, 8.2 Hz, H-2″), 4.28 (1H, s, H-28), 3.19 (1H, s, H-23), 4.23 (1H, d, J = 12.2 Hz, H-22), 4.43 (1H, dd, J = 2.5, 11.3 Hz, H-3″), 3.52 (1H, dd, J = 3.5, 11.2 Hz, H-4″), 2.17 (1H, t, J = 6.1 Hz, H-18), 2.33 (1H, d, J = 4.4 Hz, H-19), 2.21–2.26 (1H, m, H-2), 2.14 (1H, d, J = 9.2 Hz, H-11), 1.31 (1H, t, J = 3.2 Hz, H-9), 1.32–1.36 (1H, m, H-7), 1.84–1.89 (1H, m, H-6), 1.06 (3H, s, H_3-24),1.48 (1H, d, J = 4.5 Hz, H-15), 0.74 (3H, s, H_3-27), 1.12 (3H, s, H_3-26), 0.64 (3H, s, H_3-25), 1.21 (1H, t, J = 7.3 Hz, H-1), 0.87 (3H, s, H_3-29), 0.86 (3H, s, H_3-30). For ^{13}C-NMR spectral data, see Table 1. The ESI-MS spectrum showed a protonated molecular ion peak at m/z 765.4424 $[M + H]^+$ ($C_{41}H_{64}O_{13}$).

Gymnestrogenin

White amorphous solid; selected ^1H NMR (400 MHz, $CDCl_3$): δ 5.34 (1H, s, H-12), 4.49 (1H, t, J = 2.7 Hz, H-16), 4.73 (1H, s, H-28), 4.17 (1H, t, J = 3.4 Hz, H-3), 4.46 (1H, t, J = 5.6 Hz, H-21), 3.68 (1H, d, J = 11.2 Hz, H-22), 3.12 (1H, s, H-23), 2.72 (1H, dd, J = 3.2, 7.6 Hz, H-11), 2.43 (1H, t, J = 7.2 Hz, H 18), 2.01 (1H, d, J = 9.4 Hz, H-19), 1.86 (1H, d, J = 6.4 Hz, H-15), 1.58–1.62 (1H, m, H-2), 1.42–1.47 (1H, m, H-6), 1.55 (3H, s, H_3-27), 1.45 (1H, t, J = 8.6 Hz, H-9), 1.14 (1H, t, J = 2.3 Hz, H-7), 1.19 (3H, s, H_3-26), 1.15 (3H, s, H_3-25), 0.82 (1H, t, J = 4.1 Hz, H-1), 0.82 (3H, s, H_3-24), 0.68 (3H, s, H_3-29), 0.57 (3H, s, H_3-30). For ^{13}C-NMR spectral data, see Table 1. The ESI-MS spectrum showed a protonated molecular ion peak at m/z 491.37363 $[M + H]^+$ ($C_{30}H_{50}O_5$).

Inhibition of fungal mycelial growth

Pathogenic fungal strains

Pure cultures of pathogenic fungi, namely, *Aspergillus flavus* 6678, *A. flavus* 5006, *Penicillium expansum* 2995, *P. expansum* 2841, and

TABLE 1 Triterpenoid saponins and phytochemical constituents of *Gymnema sylvestre* leaves (CGS) identified tentatively in UPLC-QToF-ESI-MS/MS.

Peak	R_t (min)	Proposed phytochemicals	Formula	Neutral mass (Da)	$[M + H]^+$/$[M + H]^-$	Mass error (δ, ppm)	MS/MS fragmentation pattern
1.	16.17	Gymnemoside A	$C_{43}H_{66}O_{14}$	806.4452	807.4530	−2.23	788, 646, 590
2.	16.83	Mestanolone	$C_{20}H_{32}O_2$	304.2402	305.2480	1.54	287, 259
3.	17.94	Gymnemasin A	$C_{47}H_{74}O_{17}$	910.4926	911.5004	−0.76	893, 749, 587
4.	18.95	Gymnemoside B	$C_{43}H_{66}O_{14}$	806.4452	807.4530	1.37	645, 483
5.	19.52	Gymnemic acid III	C41H66O13	767.4581	768.4659	0.16	591, 506, 476
6.	20.20	Uscharin	$C_{31}H_{41}NO_8S$	587.2552	588.2591	2.23	231, 388, 569
7.	21.10	Deacylgymnemic acid	C36H58O12	682.3928	683.4011	0.73	508, 438, 368
8.	22.44	Narcissoside	$C_{28}H_{32}O_{16}$	624.1690	625.1768	−0.95	463, 445, 401
9.	22.93	8-Hydroxy gymnamine	C_9H_7NO	145.1617	146.1695	1.67	129, 118
10.	23.83	Madecassic acid	$C_{30}H_{48}O_6$	504.3450	505.3528	0.92	487, 469
11.	23.59	Gymnemic acid IV	$C_{41}H_{64}O_{13}$	764.4346	765.4424	0.62	588, 566, 487
12.	23.76	Ursolic acid	$C_{30}H_{48}O_3$	456.3603	457.3681	−2.37	439, 421
13.	24.10	Gymnestrogenin	$C_{30}H_{50}O_5$	490.3658	491.3736	0.84	475, 456
14.	24.42	Lupeol	$C_{30}H_{50}O$	426.3861	427.3931	−1.87	342, 286, 258
15.	25.95	Hentriacontane	$C_{31}H_{64}$	436.5008	437.5086	0.31	421, 407, 393, 255
16.	26.06	Squalene	$C_{30}H_{50}$	410.3912	411.3990	−0.12	393, 367, 273
17.	26.69	Chelidonine	$C_{20}H_{19}NO_5$	353.1263	354.1341	1.23	336, 321
18.	26.82	Octadecenoic acid	C18H34O2	282.2558	283.2636	−2.42	265, 239
19.	27.37	β-Amyrin	$C_{30}H_{50}O$	426.3861	427.3939	0.51	408, 398
20.	28.01	Eicosenoic acid	$C_{20}H_{38}O_2$	310.5145	311.4584	1.93	270, 214, 286
21.	28.37	β-Sitosterol	$C_{29}H_{50}O$	414.7067	415.7145	0.64	396, 381
22.	28.78	Oleanolic acid	$C_{30}H_{48}O_3$	456.3603	457.3681	−2.25	438, 412, 248
23.	29.02	Stigmasterol	$C_{29}H_{48}O$	412.3705	413.3775	−1.93	273, 245, 173

R_t, retention time (min); Error (ppm), the difference between experimental mass and theoretical mass; a, positive ionization mode; UPLC-QToF-ESI-MS/MS, ultra-performance liquid chromatography–quadrupole time-of-flight–electrospray ionization–tandem mass spectrometry.

Penicillium digitatum 6952, were sourced from the Indian Type Culture Collection, ICAR-Indian Agricultural Research Institute, New Delhi, India. During the assay, each fungus was sub-cultured from a respective actively growing pure culture and used for fungistatic tests.

Fungicidal assay

Fungistatic assessment was conducted as per the method described by Kundu et al. (2013). Briefly, the sample (50.0 mg, CGS) was dissolved in sterilized water (1.0 mL) containing 0.5% Triton X-100 emulsifier under aseptic conditions. The solution was then mixed with the sterilized molten potato dextrose agar media (50.0 mL). The CGS-infused media were poured into two sterilized Petri dishes (90.0-mm diameter) to achieve a test concentration of 1,000 µg/mL. Other test concentrations, ranging from 500 to 62.5 µg/mL, were also prepared accordingly and used against the test fungal pathogens. After allowing the media to solidify for 2 h, a mycelial disk (5-mm diameter) was cut out from the fungal colonies and inoculated at the center of the treated Petri dish. Fluconazole was used as the positive control, while sterilized distilled surfactant (0.5% Triton X-100) water acted as the negative control.

To evaluate the radial mycelial growth of the treated fungi, the colony diameter was measured until full growth was achieved on the control plate after 5–7 days depending on the fungi. The percentage of growth inhibition (I%) for each test concentration along with controls was determined. Furthermore, antifungal data were analyzed using the PoloOne software to determine the effective concentrations (EC, µg/mL) (Zulu et al., 2023).

Inhibition of membrane ergosterol

Ergosterol was extracted from each treated fungus at different test concentrations based on a previous method with minor modifications (Kim and Lee, 2021). In brief, fungal mycelia treated with varying concentrations of CGS were subjected to extraction of membrane ergosterol. The treated mycelia were separated and combined with 10 mL of 3% methanolic KOH and stirred vigorously on a magnetic stirrer followed by heating in a water bath for 5 h. Then, the mixture was centrifuged, and the supernatant was collected. Ergosterol was then extracted by partitioning with heptane (3 × 5 mL).

Estimation of fungal ergosterol was conducted using UPLC equipped with a Photodiode Array (PDA) detector and a mass spectrometer. Separation was achieved using a C_{18} column (2.1 mm × 100 mm, 1.8 µm; Waters, Milford, MA, USA) following a gradient system for the mobile phase, consisting of ACN and H_2O with 0.1% formic acid. The injection volume was 10 µL. The gradient program involved 0%–20% solvent A for the first 5 min, followed by 20%–100% from 5 to 18 min, and 20% solvent A from 18 to 20 min. Ergosterol was detected using the PDA detector at a λ_{max} of 282 nm. A calibration curve was generated using standard ergosterol (99% purity), and the ergosterol content in the samples was quantified. Ergosterol inhibition (%) was also calculated by comparing the ergosterol content with that of the control (untreated) (Pandey et al., 2021).

Molecular modeling and docking

The identified metabolites of CGS were subjected to molecular docking analysis to investigate their interactions with the target protein, cytochrome P_{450} sterol 14-α-demethylase. The amino acid sequence for this fungal protein was obtained from the National Center for Biotechnology Information (NCBI) database. Additionally, the NCBI BLAST tool and Protein Data Bank (PDB) database were utilized to find appropriate templates for designing the secondary structures of the specified amino acid sequences. Subsequently, Modeller (version 9.24: r11614) was employed to model the homology protein structures, which were saved in.pdb format (Laskowski et al., 1993). The accuracy of the modeled receptor protein was evaluated using the PROCHECK software (default version Linux 64: 3.5.4). In this study, "ligands" refers to the 3D molecular structures of the identified metabolites that were used for molecular docking. The 3D structures of these compounds were prepared using Chem Draw Ultra 11.0 and saved in.sdf format. *In silico* molecular docking simulations were performed using the SeeSAR v10.3.1 software. The structure of the receptor protein was generated, and the vacant active site residues were identified. A comprehensive set of molecules was selected, generating both 3D and 2D frameworks. The interactions between the docked receptor and ligands were visualized using the Discovery Studio v4.1 visualizer (Saxena et al., 2010).

Statistical analysis

Statistical analysis was performed using the SAS® Proprietary Software version 9.4 (TS1M1), licensed to ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India. Results were considered not significant when $p < 0.05$.

Results

Saponins and others bioactive metabolites

Metabolites of CGS were determined in UPLC-QToF-ESI-MS/MS, indicating the occurrence of several major and minor components. UPLC chromatogram (Figure 1) of CGS displayed the detection of peaks, fully separated within 30 min of run time.

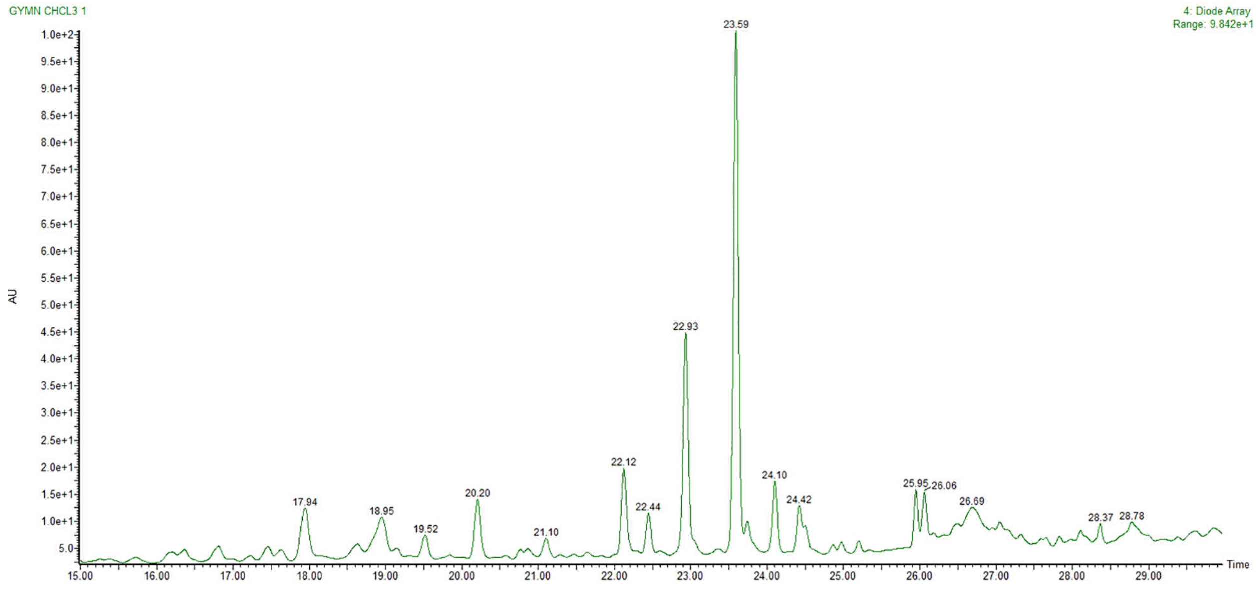


FIGURE 1
UPLC chromatogram of CGS as analyzed and separated through RP-18 column. UPLC, ultra-performance liquid chromatography; CGS, chloroform extract of *Gymnema sylvestre*.

Twenty-three metabolites were identified tentatively from CGS based on their exact molecular ion peaks (Table 1).

Gymnemoside A ($C_{43}H_{66}O_{14}$), mestanolone ($C_{20}H_{32}O_2$), gymnemasin A ($C_{47}H_{74}O_{17}$), and gymnemoside B ($C_{43}H_{66}O_{14}$) were eluted from the C_{18} column at the retention time (R_t) of 16.17–18.95 min with their respective accurate mass $[M + H]^+$ peaks at m/z 807.4530, 305.2480, 911.5004, and 807.4530 amu. Similarly, gymnemic acid III ($C_{41}H_{66}O_{13}$), uscharin ($C_{31}H_{41}NO_8S$), deacylgymnemic acid (C36H58O12), narcissoside ($C_{28}H_{32}O_{16}$), 8-hydroxy gymnamine (C_9H_7NO), and madecassic acid ($C_{30}H_{48}O_6$) were also separated and detected at the R_t of 19.52–23.83 min with their accurate mass $[M + H]^+$ peaks at m/z 768.4659, 588.2591, 683.4011, 625.1768, 146.1695, and 505.3528 amu, respectively. Likewise, gymnemic acid IV ($C_{41}H_{64}O_{13}$), ursolic acid ($C_{30}H_{48}O_3$), gymnestrogenin ($C_{30}H_{50}O_5$), lupeol ($C_{30}H_{50}O$), hentriacontane ($C_{31}H_{64}$), squalene ($C_{30}H_{50}$), and chelidonine ($C_{20}H_{19}NO_5$) were also identified with their corresponding accurate mass $[M + H]^+$ peaks at m/z 765.4424, 457.3681, 491.3736, 427.3931, 437.5086, 411.3990, and 354.1341 amu, respectively. Later, fatty acids, sterols, and hydrocarbons such as octadecenoic acid (C18H34O2), β-amyrin ($C_{30}H_{50}O$), eicosenoic acid ($C_{20}H_{38}O_2$), β-sitosterol ($C_{29}H_{50}O$), oleanolic acid ($C_{30}H_{48}O_3$), and stigmasterol ($C_{29}H_{48}O$) were also tentatively characterized from their respective accurate mass $[M + H]^+$ peaks at m/z 282.2558, 426.3861, 310.5145, 414.7067, 456.3603, and 412.3705 amu, respectively.

Among the metabolites of CGS identified in UPLC-QToF-MS/MS, two compounds (CGS-1 and CGS-2) were isolated through column chromatography and characterized spectroscopically using ^{1}H-NMR, ^{13}C-NMR, and high-resolution mass spectrometry (HRMS). The purified crystals of CGS-1 appeared lemon-yellowish in color. The ^{1}H-NMR spectrum of CGS-1 displayed signals as doublet at δ 5.44 ppm with the J value of 8.6 Hz and at δ 4.23 ppm with the J value of 12.2 Hz, which corresponded to the respective proton at H-21 and H-22, respectively (Figure 2). An additional signal as singlet appeared at δ 5.34 ppm, attributed to the proton at H-12. Other signals appeared as triplet at δ 4.49 ppm, which were assigned to the proton at H-16. Again, characteristic signals were observed in the ^{13}C-NMR spectrum of CGS-1 (Table 2). ^{13}C-NMR spectrum exhibited δ 121.8 ppm (C-12) and δ 142.4 ppm (C-13), indicating the occurrence of conjugated and aromatic carbons. The chemical shift values for the glucopyranosiduronic acid unit were evident in the spectrum, with the signals appearing at δ 106.2 ppm (C1″), δ 75.3 ppm (C2″), δ 69.7 ppm (C3″), δ 72.4 ppm (C4″), δ 76.1 ppm (C5″), and δ 172.6 ppm (C6″). Furthermore, HRMS analysis gave a sharp peak at m/z 765.4424, corresponding to its adduct $[M + H]^+$ with the empirical formula $C_{41}H_{64}O_{13}$ (Figure 3). Based on these features, CGS-1 was identified as gymnemic acid IV (Table 2).

CGS-2 was also purified from CGS using column chromatography and recrystallized, which appeared as a white amorphous crystal. The ^{1}H-NMR spectrum of CGS-2 gave a signal as singlet at δ 5.34 ppm, attributed to the H-12 proton. Other signals were detected as triplet at δ 4.46 ppm with a J value of 5.6 Hz and at δ 3.68 ppm as doublet with a J value of 11.2 Hz, corresponding to the H-21 and H-22 protons, respectively. Another signal appeared as singlet at δ 4.73 ppm, attributed to the proton at H-28 of CGS-2. Similarly, the ^{13}C-NMR spectrum of CGS-2 exhibited signals at δ 124.2 ppm (C-12) and δ 141.5 ppm (C-13), suggesting the presence of allylic and aromatic carbons. Carbons at δ 52.6 ppm (C-5) and δ 56.8 ppm (C-18) reflected the presence of aliphatic hydrocarbons. Additional signals at δ 75.2 ppm (C-21) and δ 74.6 ppm (C-22) corresponded to the carbon attached with hydroxyl functionalities. HRMS analysis of the purified molecule (CGS-2) gave $[M + H]^+$ peak at m/z 491.37363 with the corresponding empirical formula $C_{30}H_{50}O_5$ (Figure 3). With these characteristic features, the molecule CGS-2 was identified as gymnestrogenin (Figure 4).

Fungal growth inhibition

The antifungal effectiveness of CGS against selected fungal pathogens such as *A. flavus* 6678, *A. flavus* 5006, *P. expansum* 2995, *P. expansum* 2841, and *P. digitatum* 6952 revealed broad-spectrum activity. Among the tested fungi, strains of *P. digitatum* 6952 were highly susceptible to CGS, exhibiting more than 50% mycelial growth inhibition. Maximum growth inhibition (%) was recorded against *P. digitatum* 6952, *P. expansum* 2995, and *A. flavus* 6678. Additionally, significant growth inhibition was also noticed

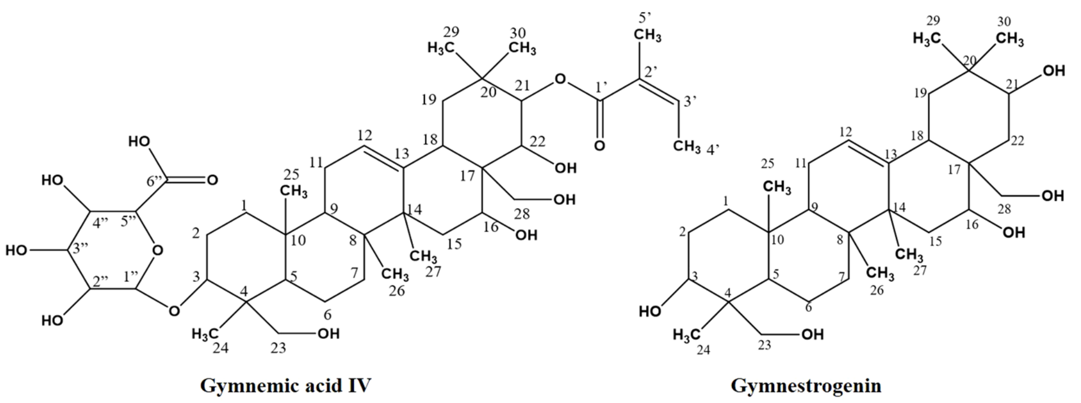


FIGURE 2
Isolation of compounds (CGS-1 and CGS-2) from chloroform extract of *Gymnema sylvestre* leaf.

TABLE 2 ^{13}C-NMR spectral data for gymnemic acid IV and gymnestrogenin (δ in $CDCl_3$) from the leaves of *Gymnema sylvestre*.

Position (C/H)	CGS-1		CGS-2	
	δ ^{1}H (ppm)	δ ^{13}C (ppm)	δ ^{1}H (ppm)	δ ^{13}C (ppm)
1	1.21 (t, J = 7.3 Hz)	27.9	0.82 (t, J = 4.1 Hz)	39.7
2	2.21–2.26 (m)	26.3	1.58–1.62 (m)	29.3
3	4.31–4.36 (m)	75.1	4.17 (t, J = 3.4 Hz)	72.7
4	2.49 (d, J = 2.6 Hz)	42.6	2.14 (d, J = 8.6 Hz)	44.1
5	–	52.4	–	52.8
6	1.84–1.89 (m)	17.2	1.42–1.47 (m)	19.5
7	1.32–1.36 (m)	29.7	1.14 (t, J = 2.3 Hz)	33.4
8	–	53.6	–	43.9
9	1.31 (t, J = 3.2 Hz)	48.5	1.45 (t, J = 8.6 Hz)	48.7
10	–	28.6		39.2
11	2.14 (d, J = 9.2 Hz)	24.3	2.72 (dd, J = 3.2, 7.6 Hz)	27.32
12	5.34 (s)	121.8	5.34 (s)	124.2
13	–	142.4		141.5
14	–	44.8	2.04 (s)	46.2
15	1.48 (d, J = 4.5 Hz)	34.3	1.86 (d, J = 6.4 Hz)	32.4
16	4.49 (t, J = 2.7 Hz)	64.9	5.17 (t, J = 3.7 Hz)	64.52
17	–	47.6	–	47.9
18	2.17 (t, J = 6.1 Hz)	58.1	2.43 (t, J = 7.2 Hz)	56.8
19	2.33 (d, J = 4.4 Hz)	53.6	2.01 (d, J = 9.4 Hz)	54.3
20	–	36.8	–	29.8
21	5.44 (d, J = 8.6 Hz)	78.1	4.46 (t, J = 5.6 Hz)	75.2
22	4.23 (d, J = 12.2 Hz)	74.4	3.68 (d, J = 11.2 Hz)	74.6
23	3.19 (s)	68.7	3.12 (s)	68.2
24	1.06 (s)	15.4	0.82 (s)	14.3
25	0.64 (s)	17.3	1.15 (s)	19.8
26	1.12 (s)	18.3	1.19 (s)	23.6
27	0.74 (s)	27.9	1.55 (s)	26.2
28	4.28 (s)	61.5	4.73 (s)	63.5
29	0.87 (s)	28.3	0.68 (s)	23.3
30	0.86 (s)	20.3	0.57 (s)	28.5
1’	–	171.6		
2’	1.47 (s)	131.6		
3’	6.92–6.97 (m)	146.8		
4’	1.89 (d, J = 8.9 Hz)	16.3		
5’	1.86 (s)	17.9		
1″	4.52 (d, J = 7.5 Hz)	106.2		
2″	4.09 (dd, J = 1.2, 8.2 Hz)	75.3		
3″	4.53 (dd, J = 2.5, 11.3 Hz)	69.7		

(Continued)

TABLE 2 Continued

Position (C/H)	CGS-1		CGS-2	
	δ ^{1}H (ppm)	δ ^{13}C (ppm)	δ ^{1}H (ppm)	δ ^{13}C (ppm)
4″	3.52 (dd, J = 4.5, 9.2 Hz)	72.4		
5″	4.46 (d, J = 4.8 Hz)	76.1		
6″	12.16 (s)	172.6		

s, singlet; d, doublet; t, triplet; m, multiplet; dd, double doublet; br, broad; J, coupling constant; Hz, hertz.

against *P. expansum* 2841 and *A. flavus* 5006. Antifungal action of the bioactive saponin-rich CGS showed a higher response in terms of activity, which increased with the concentration. Fungistatic efficacy of CGS showed EC_{50} 297.2, 360.5, and 369.4 µg/mL against *P. digitatum* 6952, *P. expansum* 2995, and *A. flavus* 6678, respectively (Table 3). The order of fungistatic action followed *P. digitatum* 6952 > *P. expansum* 2995 > *A. flavus* 6678 > *P. expansum* 2841 > *A. flavus* 5006. However, fluconazole (positive control) was more inhibitory against *P. digitatum* 6952 and *P. expansum* 2995 with EC_{50} 67.3 and 79.0 µg/mL, respectively.

Furthermore, the influence of the bioactive metabolites (saponins) of CGS on the fungal species was also assessed by determining the membrane ergosterol content of the treated fungal pathogens. Ergosterol is a functional sterol that generally imparts strength to the fungal membrane structure, providing stability and integrity. In this context, the efficacy of CGS was determined using UPLC-MS, which indicated complete inhibition of ergosterol content at 2,000 µg/mL concentration of CGS. More than 50% inhibition in ergosterol content (57.61% ± 2.08%) was recorded in *P. digitatum* 6952 at the concentration of 125 µg/mL (Table 4).

Interaction with target protein

Molecular docking analysis was extensively exploited recently to explain the functional interaction of the metabolites with target-specific enzymes, facilitating possible inhibition of the enzymes. Metabolites identified in CGS were used for the analysis of their potential inhibitory action toward cyt P_{450} sterol 1,4-α-demethylase. The results revealed that the saponins exhibited promising docking scores ranging from −21.7 to −6.3 kcal/mol (Table 5). Based on the poses generated through molecular docking, the potential molecules engaging in protein–ligand complexes are mentioned in decreasing order: gymnemic acid IV > gymnemoside A > gymnestrogenin > gymnemic acid III > gymnemoside B > gymnemasin A > lupeol. The lower binding affinity range indicated favorable interactions between the molecules (ligands) and cyt P_{450} sterol 1,4-α-demethylase (target protein). Among the tested compounds, gymnemic acid IV was identified as the most efficient binder with the lowest binding energy (–21.7 kJ/mol), suggesting better stability of the ligand–protein complex. Furthermore, the ligand efficiency (LE) and lipophilic ligand efficiency (LLE) were also deduced to determine their ability to bind the sterol biosynthesizing protein based on their respective lipophilic property. Herein, the steroidal and triterpenic structures of the saponins may have helped to make the strong ligand–protein complex, facilitating the possibility of blocking the target protein. Based on the binding affinity and associated free energy, the promising molecules are mentioned as gymnemic acid IV (−21.7 kJ/mol), gymnemoside A (−19.5 kJ/mol), and gymnestrogenin (−16.5 kJ/mol). Furthermore, gymnemic acid IV showed interactions with specific amino acid residues such as HIS259, VAL434, PHE255, PHE83, LEU321, MOL451, and LEU100 at the target sites of the enzyme (Figure 5). The stability of the ligand–protein complex could be attributed to low-distance conventional hydrogen bonds and hydrophobic pi-alkyl interactions.

TABLE 3 Growth inhibition (EC_{50}, µg/mL) of five different fungal strains of storage pathogens treated with chloroform extract of *Gymnema sylvestre* leaves.

Pathogen	$^a EC_{50}$ (µg/mL)	95% Confidence limit (µg/mL)		^{b}Slope ± SE	^{c}Intercept ± SE	$^d(\chi^2)$
		Lower	Upper			
Aspergillus flavus 6678	369.4	279.1	517.6	1.59 ± 0.015	−0.57 ± 0.23	5.40
A. flavus 5006	433.9	341.4	584.9	1.14 ± 0.14	−0.57 ± 0.20	1.30
Penicillium digitatum 6952	297.2	212.3	369.8	1.14 ± 0.14	−0.48 ± 0.20	0.50
Penicillium expansum 2995	360.5	288.5	465.7	1.18 ± 0.14	−0.50 ± 0.20	1.73
Penicillium expansum 2841	427.6	284.2	785.0	1.22 ± 0.14	−0.52 ± 0.21	3.93

[a] EC_{50}, effective concentration (µg/mL) at which 50% mycelial growth inhibition recorded.
[b] Slope response of regression equation ± standard error.
[c] Intercept regression equation ± standard error.
[d] χ^2, chi-squared value.

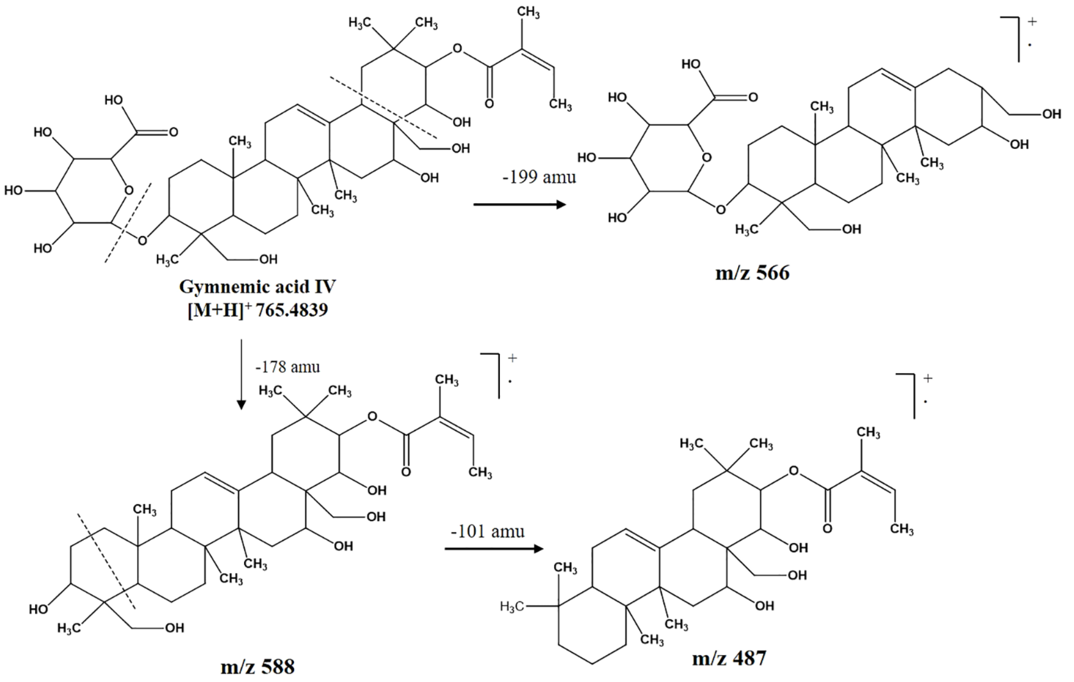


FIGURE 3
Mass fragmentation pattern of CGS-1 (gymnemic acid IV) as analyzed in UPLC-QToF-MS/MS. UPLC-QToF-MS/MS, ultra-performance liquid chromatography–quadrupole time-of-flight–tandem mass spectrometry.

Discussion

Metabolomics profiling and analysis of *G. sylvestre* leaves have shown the dominant composition of triterpenoid saponins, along with other phyto-molecules. Saponins of the plant have been studied extensively through either comprehensive profiling or stepwise isolation, purification, and characterization techniques (Alkefai et al., 2019). In the current investigation, saponins with diverse structural components along with other minor metabolites have been identified using chromatographic and spectroscopic techniques. UPLC-QToF-ESI-MS/MS-based tentative characterizations following mass accuracy along with fragmentation patterns indicated the identification

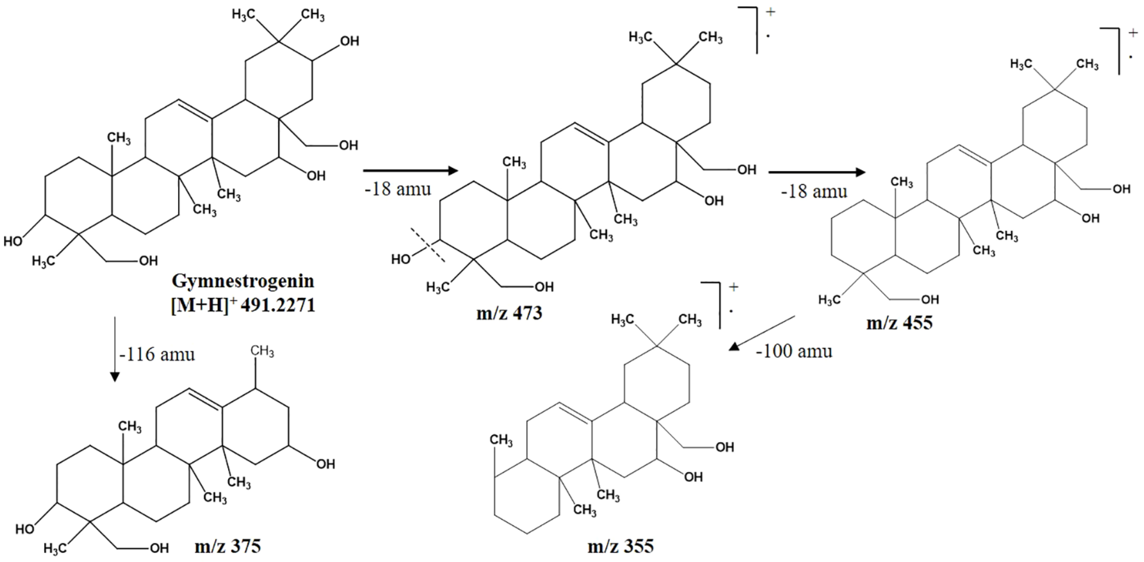


FIGURE 4
Mass fragmentation pattern of CGS-2 (gymnestrogenin) as analyzed in UPLC-QToF-MS/MS. UPLC-QToF-MS/MS, ultra-performance liquid chromatography–quadrupole time-of-flight–tandem mass spectrometry.

TABLE 4 Ergosterol inhibition (%) with the treatment of chloroform soluble fraction at different concentrations against selected fungal pathogens.

Conc. (µg/mL)	Ergosterol inhibition *(%)				
	Penicillium digitatum 6952	*Penicillium expansum* 2995	*Aspergillus flavus* 6678	*P. expansum* 2841	*A. flavus* 5006
62.5	39.35 ± 2.74	35.85 ± 1.24	22.72 ± 2.53	28.84 ± 2.09	31.01 ± 3.48
125	59.61 ± 2.08	57.38 ± 3.13	31.62 ± 3.69	45.31 ± 1.76	53.74 ± 2.61
250	66.73 ± 2.18	64.77 ± 2.74	58.50 ± 2.61	54.36 ± 2.69	60.81 ± 1.87
500	82.99 ± 2.42	79.92 ± 3.02	69.33 ± 2.42	68.18 ± 3.86	65.46 ± 2.27
1,000	89.43 ± 3.96	87.31 ± 2.65	80.32 ± 1.81	76.19 ± 2.31	69.28 ± 1.97
2,000	100 ± 0.00	100 ± 0.00	100 ± 0.00	100 ± 0.00	100 ± 0.00

*Results are shown as mean ± SD.

of 23 metabolites, including gymnemic acid IV (765.4424), deacylgymnemic acid (683.4011), gymnestrogenin (491.3736), and hentriacontane (437.5086).

Metabolomics investigation of the methanolic extract of *G. sylvestre* stem bark has revealed ample triterpenoids, sterols, and flavonoids (Ditchou et al., 2024; Neel et al., 2025). Similarly, the presence of arylated gymnemic acids and *Gymnema* saponins has been confirmed in the aqueous extract of *G. sylvestre* leaves (Alkefai et al., 2018). Likewise, two triterpenoid saponins, namely, gymnemosides W_1 and W_2, have been isolated and characterized from the plant (Zhu et al., 2008). In the present study, triterpenoid saponins have been primarily characterized along with other constituents. Furthermore, gymnemic acid IV and gymnestrogenin have also been isolated and characterized from the saponin-rich extracts using chromatographic and spectroscopic techniques. These findings have been consistent with those reported by Renga et al. (2015), suggesting the dominant composition of oleanane saponins in *G. sylvestre*.

CGS has been found to be highly effective against storage fungi, which could be attributed to the biofunctional metabolites, mainly gymnemic acid derivatives, particularly gymnemic acid IV, gymnemoside A, and gymnestrogenin. Saponins and other valuable metabolites of the plant have been recognized for multidimensional biofunctional properties. Even the semi-nonpolar extracts ($CHCl_3$ soluble fraction) of *G. sylvestre* aerial parts have been found effective for antimicrobial activity as compared to the other solvent extracts (Chodisetti et al., 2013), indicating the presence of the most prominent components in the $CHCl_3$ soluble fraction.

In the current investigation, the fungistatic action of *G. sylvestre* leaves has been demonstrated against *P. digitatum* 6952, *P. expansum* 2995, *A. flavus* 6678, *P. expansum* 2841, and *A. flavus* 5006. Herein, *Gymnema* extract has been tested against the two different strains of *P. expansum* along with different species and other fungi. Previous reports suggested that variable responses of the different strains of the same species could be due to their difference in the virulence level (Kundu et al., 2022; Neel et al., 2023). In the current study, treatment of CGS has been found to be the most effective in arresting the growth of *P. digitatum* 6952, exhibiting EC_{50} 297.2 µg/mL. Furthermore, CGS was also inhibitory to the other pathogens, indicating broad-spectrum activity. Future research has been planned to utilize

TABLE 5 Molecular docking scores and energies associated with the enzyme complex containing major components of *Gymnema sylvestre* leaves.

Identified compounds	Log P	Binding affinity range (nM)	∆G	LE	LLE
Gymnemic acid IV	2.30	19,560.06 < KI < 1,943,407.61	−21.7	–	–
Gymnemoside A	3.02	40,371.02 < KI < 4,011,099.55	−19.5	–	–
Gymnestrogenin	4.13	134,723.46 < KI < 13,642,372.71	−16.5	–	–
Gymnemic acid III	2.01	15,920.64 < KI < 1,581,810.32	−15.1	–	–
Gymnemoside B	4.54	3,751,403.48 < KI < 372,724,116.5	−8.1	–	–
Gymnemasin A	1.71	1,013,571.01 < KI < 100,704,272.7	−6.3	–	–
Lupeol	2.53	2,332,832.65 < KI < 136,148,241.2	−2.4	–	–

LE, ligand efficiency; LLE, lipophilic ligand efficiency.

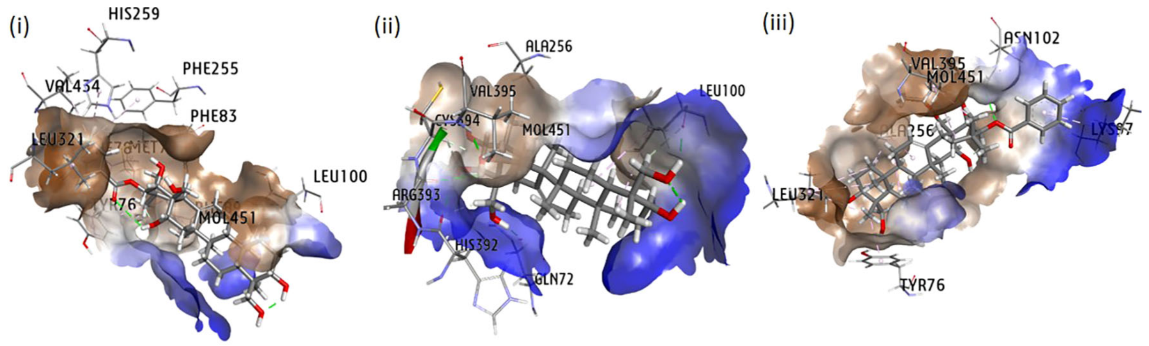


FIGURE 5
Interaction in 3D poses of *Gymnema sylvestre* metabolites with protein squalene epoxidase (i) gymnemic acid IV, (ii) gymnemoside A, and (iii) gymnestrogenin.

the most bioactive *Gymnema* chloroform fraction for the development of suitable coating formulation to make the process economical. Nevertheless, the purified compounds gymnemic acid IV and gymnestrogenin have not been further tested against the fungi. However, there could be a possibility of better antifungal action with these isolated compounds. In line with our study, ethanolic and ethyl acetate extracts of the plant have been reported to possess antimicrobial activity against various pathogenic bacteria (Satdive et al., 2003; Arora and Sood, 2017). As far as our literature survey could ascertain, the potential of chloroform soluble fraction of *G. sylvestre* to control fungal pathogens has not been reported previously. Thus, the present research could be considered the first report on the fungistatic action of the saponins of *G. sylvestre* leaves against spoilage-causing fungal pathogens.

In order to explain the efficacy of *Gymnema* metabolites in arresting fungal growth, membrane ergosterol content in the treated fungi has been studied. Ergosterol is an important component of fungal membranes, providing integrity and strength to the membrane. Hampering ergosterol biosynthesis could lead to less deposition on the membrane, hence the possibility of disintegration of fungal membrane damage (Kumar et al., 2023). In the present study, complete inhibition of ergosterol production has been noticed in the fungi at 0.2% concentration of CGS. Notably, the treated fungal pathogens displayed a significant reduction in membrane ergosterol content with the saponin-rich metabolites of *G. sylvestre* leaves; therefore, membrane integrity has been severely compromised. Inhibition of ergosterol production in the treated fungi has been consistent; hence, these findings have also been confirmed through an *in silico* analysis.

In silico molecular modeling of *Gymnema* metabolites to inhibit/block the target-specific protein, cyt P_{450} sterol 1,4-α-demethylase, responsible for ergosterol production in fungi, has been explained, suggesting positive interaction of the metabolites over the native ligand. In order to decipher the probable mechanism of inhibition by these molecules, stable conformations of these metabolites forming low energy bonds within the amino acid residues at the specific binding pocket of the protein have been studied. Strong interactions between amino acid residues and metabolites were confirmed through the docking scores and corresponding associated binding energies. Major metabolites have been found to be energetically favorable and qualitatively aligned with the *in vitro* analysis. Discrete attempts have been made to understand the possibilities of triterpenoids from *G. sylvestre* to inhibit α-glycosidase (Parveen et al., 2019).

Ligand–receptor-based molecular docking analysis of the identified metabolites of *G. sylvestre* leaves with the fungal ergosterol biosynthesizing enzyme cyt P_{450} sterol 1,4-α-demethylase has been found to be significant enough to consider the disruption of fungal ergosterol production, thereby damaging fungal membrane. The interaction complex of the compound gymnemic acid IV with the target protein has been found to be highly stable. Furthermore, multiple amino acid residues at site 1 of the enzyme have been identified in the interaction, forming a thermodynamically stable complex. Herein, gymnemic acid IV has been found to be the most effective with the lowest binding energy and desired LE and LLE parameters to block the enzyme with conventional H-bonds, hydrophobic π-alkyl, π-π, and π-sigma interactions. Additionally, gymnemoside A and gymnestrogenin have also been recognized next in order with respect to binding energy and stability to inhibit the protein. The interaction of gymnemic acid IV, gymnemoside A, and gymnestrogenin with cyt P_{450} sterol 1,4-α-demethylase for the first time has been reported to predict the inhibition mechanism of such saponins from *G. sylvestre* leaves.

Conclusion

In summary, comprehensive profiling and characterizations of metabolites of *G. sylvestre* leaves have been conducted to identify 23 metabolites that showed significant fungistatic action to inhibit the growth of fungal pathogens *P. digitatum* 6952, *P. expansum* 2995, *A. flavus* 6678, *P. expansum* 2841, and *A. flavus* 5006, responsible for causing rot during post-harvest operations of agricultural produce. The saponin-rich fraction comprising abundant gymnemic acid IV along with other metabolites has been found to be highly inhibitory toward the fungi and also affects their ergosterol production. Molecular docking analysis revealed enough evidence regarding the possible interactions of these metabolites with the ergosterol biosynthesizing fungal protein, which could be responsible for the inhibition of ergosterol production in the treated fungi. These encouraging scientific findings could be of high importance for the exploitation of *G. sylvestre* under sustainable product development in crop protection.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

SN: Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft. AM: Data curation, Formal Analysis, Validation, Writing – original draft. SS: Methodology, Resources, Writing – review & editing. AD: Validation, Writing – review & editing. AK: Conceptualization, Supervision, Visualization, Writing – review & editing. AS: Resources, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors wish to thank the Director, Indian Council of Agricultural Research-Indian Agricultural Research Institute (ICAR-IARI), New Delhi, India.

Conflict of interest

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

Generative AI statement

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

Publisher's note

References

Ahamad, J., Sh, M., Mohammed Ameen, E. T., and Anwer, A. (2018). A critical review of the potential pharmacological and phytochemical properties of *Gymnema sylvestre* R. Br. *J. Glob. Trends Pharm. Sci.* 9, 5869–5886. Available online at: https://www.jgtps.com/index.php.

Alkefai, N. H., Ahamad, J., Amin, S., Sharma, M., and Mir, S. R. (2018). Arylatedgymnemic acids from *Gymnema sylvestre* R. Br. as potential α-glucosidase inhibitors. *Phytochem. Lett.* 25, 196–202. doi: 10.1016/j.phytol.2018.04.021

Alkefai, N. H. A., Amin, S., Sharma, M., Ahamad, J., and Mir, S. R. (2019). New olean-15-ene type gymnemic acids from *Gymnema sylvestre* (Retz.) R. Br. and their antihyperglycemic activity through α-glucosidase inhibition. *Phytochem. Lett.* 32, 83–89. doi: 10.1016/j.phytol.2019.05.005

Arora, D. S., and Sood, H. (2017). *In vitro* antimicrobial potential of extracts and phytoconstituents from *Gymnema sylvestre* R. Br. leaves and their biosafety evaluation. *AMB Express* 7, 1–13. doi: 10.1186/s13568-017-0416-z

Chen, G., and Guo, M. (2017). Rapid screening for α-glucosidase inhibitors from *Gymnema sylvestre* by affinity ultrafiltration–HPLC-MS. *Front. Pharmacol.* 8, 228. doi: 10.3389/fphar.2017.00228

Chodisetti, B., Rao, K., and Giri, A. (2013). Phytochemical analysis of *Gymnema sylvestre* and evaluation of its antimicrobial activity. *Nat. Prod. Res.* 27, 583–587. doi: 10.1080/14786419.2012.676548

Dar, M. I., Sultan, A., Abass, S., Dev, K., Parveen, R., Ahmad, S., et al. (2024). Exploring the anti-diabetic mechanism of selective phytochemicals identified from *Gymnema sylvestre* using TLC-UPLC-MS, complemented by in silico studies. *Phytomed. plus* 4, 100606. doi: 10.1016/j.phyplu.2024.100606

Devi, M. B., Kelapure, N. N., Kadam, M. L., Salve, R. V., Wadatkar, H. V., and Bisarya, D. (2021). Phytochemical composition, traditional uses, and health benefits of *Gymnema sylvestre* (Gudmar) leaves powder. *Bioscene* 21, 884–905.

Ditchou, Y. O. N., Leutcha, P. B., Miaffo, D., Mamoudou, H., Ali, M. S., à Ngnoung, G. A. A., et al. (2024). *In vitro* and in silico assessment of antidiabetic and antioxidant potencies of secondary metabolites from *Gymnema sylvestre*. *Biomed. Pharmacother.* 177, 117043. doi: 10.1016/j.biopha.2024.117043

Dutta, S., Kundu, A., Dutta, A., Saha, S., and Banerjee, K. (2021). A comprehensive chemical profiling of phytochemicals from *Trachyspermum ammi* and encapsulation for sustained release. *LWT* 147, 111577. doi: 10.1016/j.lwt.2021.111577

Gopiesh Khanna, V., and Kannabiran, K. (2008). Antimicrobial activity of saponin fractions of the leaves of *Gymnema sylvestre* and *Eclipta prostrata*. *World J. Microbiol. Biotechnol.* 24, 2737–2740. doi: 10.1007/s11274-008-9758-7

Gopiesh Khanna, V., Kannabiran, K., Sarath Babu, V., and Sahul Hameed, A. S. (2011). Inhibition of fish nodavirus by gymnemagenol extracted from Gymnema sylvestre. *J. Ocean Univ China* 10, 402–408. doi: 10.1007/s11802-011-1841-2

Gurnani, N., Mehta, D., Gupta, M., and Mehta, B. K. (2014). Natural products: Source of potential drugs. *Afr. J. Basic. Appl. Sci.* 6, 171–186. doi: 10.5829/idosi.ajbas.2014.6.6.21983

Ibrahim, A., Babandi, A., Sani, A. H., Wudil, A. M., Murtala, Y., Umar, I. A., et al. (2017). HPLC profile, *in vitro* alpha-amylase, alpha-glucosidase inhibitory and antioxidant activities of *Gymnema sylvestre* ethyl acetate leaf extract. *Bayero J. Pure Appl. Sci.* 10 (1), 72–80. doi: 10.19044/esj.2017.v13n36p218

Jung, M., Park, M., Lee, H. C., Kang, Y.-H., Kang, E. S., and Kim, S. K. (2006). Antidiabetic agents from medicinal plants. *Curr. Med. Chem.* 13, 1203–1218. doi: 10.2174/092986706776360860

Kamble, B., Gupta, A., Patil, D., Janrao, S., Khatal, L., and Duraiswamy, B. (2013). Quantitative estimation of gymnemagenin in *Gymnema sylvestre* extract and its marketed formulations using the HPLC–ESI–MS/MS method. *Phytochem. Anal.* 24, 135–140. doi: 10.1002/pca.v24.2

Khan, F., Sarker, M. M. R., Ming, L. C., Mohamed, I. N., Zhao, C., Sheikh, B. Y., et al. (2019). Comprehensive review on phytochemicals, pharmacological, and clinical potentials of *Gymnema sylvestre*. *Front. Pharmacol.* 10, 1223. doi: 10.3389/fphar.2019.01223

Khursheed, A., Rather, M. A., Jain, V., Rasool, S., Nazir, R., Malik, N. A., et al. (2022). Plant based natural products as potential ecofriendly and safer biopesticides: A comprehensive overview of their advantages over conventional pesticides, limitations and regulatory aspects. *Microbial. Pathogenesis* 173, 105854. doi: 10.1016/j.micpath.2022.105854

Kim, G., and Lee, S. E. (2021). Antifungal and antiaflatoxigenic properties of naphthoquinones toward *Aspergillus flavus* and their mode of inhibitory action on aflatoxin biosynthesis. *Food Control* 119, 107506. doi: 10.1016/j.foodcont.2020.107506

Kumar, R., Mandal, A., Saha, S., Dutta, A., Chawla, G., Das, A., et al. (2023). Zanthoxylum alatum fruits: process optimization for tambulin-rich valuable phytocompounds, antifungal action coupled with molecular modeling analysis. *Biomass Conv. Bioref.* 14, 30467–30484. doi: 10.1007/s13399-023-04848-9

Kundu, A., Mandal, A., Dutta, A., Saha, S., Raina, A. P., Kumar, R., et al. (2022). Nanoemulsification of Kaempferia galanga essential oil: Characterizations and molecular interactions explaining fungal growth suppression. *Process Biochem.* 121, 228–239. doi: 10.1016/j.procbio.2022.07.008

Kundu, A., Saha, S., Walia, S., Shakil, N. A., Kumar, J., and Annapurna, K. (2013). Cadinene sesquiterpenes from Eupatorium adenophorum and their antifungal activity. *J. Environ. Sci. Health, B.* 48 (6), 516–522.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. doi: 10.1107/S0021889892009944

Muddapur, U. M., Manjunath, S., Alqahtani, Y. S., Shaikh, I. A., Khan, A. A., Mannasaheb, B. A., et al. (2024). Exploring bioactive phytochemicals in *Gymnema sylvestre*: Biomedical uses and computational investigations. *Separations* 11, 50. doi: 10.3390/separations11020050

Munjanja, B. K. (2017). "Liquid chromatography mass spectrometry (LC-MS)," in *Multiresidue Methods for the Analysis of Pesticide Residues in Food* (Boca Raton: CRC Press), 197–234.

Neel, S., Mandal, A., Dutta, A., Saha, S., Das, A., Chawla, G., et al. (2023). Response surface methodology guided process optimizations, modeling and biofunctional analysis of phytochemicals from *Nigella sativa* seeds as a potential antifungal agent. *Ind. Crop Prod.* 199, 116695. doi: 10.1016/j.indcrop.2023.116695

Neel, S., Suman, S., Barik, A., Mandal, A., Saha, S., Basak, B. B., et al. (2025). Optimization of extraction and isolation of *Gymnema sylvestre* bioactive metabolites for potential antifungal activity. *Biomass Convers. Biorefin.* 1–17. doi: 10.1007/s13399-024-06478-1

Pandey, S., Giri, V. P., Kumari, M., Tripathi, A., Gupta, S. C., and Mishra, A. (2021). Comparative study of the development and characterization of ecofriendly oil and water nanoemulsions for improving antifungal activity. *ACS Agric. Sci. Technol.* 1, 640–654. doi: 10.1021/acsagscitech.1c00141

Parveen, S., Ansari, M. H. R., Parveen, R., Khan, W., Ahmad, S., and Husain, S. A. (2019). Chromatography-based metabolomics and in silico screening of *Gymnema sylvestre* leaf extract for its antidiabetic potential. *Evid.-Based Complement. Altern. Med.* 1, 7523159. doi: 10.1155/2019/7523159

Patil, R., Patil, R., Ahirwar, B., and Ahirwar, D. (2011). Current status of Indian medicinal plants with antidiabetic potential: A review. *Asian Pac. J. Trop. Biomed.* 1, S291–S298. doi: 10.1016/S2221-1691(11)60175-5

Pham, H. T. T., Kim, H. W., Han, S., Ryu, B., Doan, T. P., An, J. P., et al. (2019). Development of a building block strategy to target the classification, identification, and metabolite profiling of pleanane triterpenoids in *Gymnema sylvestre* using UHPLC-qTOF/MS. *J. Natural products* 82, 3249–3266. doi: 10.1021/acs.jnatprod.9b00328

Pillai, A. R., Eapen, A. S., Zhang, W., and Roy, S. (2024). Polysaccharide-based edible biopolymer-based coatings for fruit preservation: A review. *Foods* 13, 1529. doi: 10.3390/foods13101529

Rahman, S., and Husen, A. (2023). "Phytochemical constituents, antidiabetic, and other activities of Gurmar (*Gymnema sylvestre* R. Br.)," in *Antidiabetic Medicinal Plants and Herbal Treatments*. Boca Raton. 159–168.

Renga, B., Festa, C., De Marino, S., Di Micco, S., D'Auria, M. V., Bifulco, G., et al. (2015). Molecular decodification of gymnemic acids from *Gymnema sylvestre*. *Steroids* 96, 121–131. doi: 10.1016/j.steroids.2015.01.024

Romaiyan, A., Persaud, S. J., and Jones, P. M. (2023). Identification of potential plant-derived pancreatic beta-cell-directed agents using a new custom-designed screening method: *Gymnema sylvestre* as an example. *Molecules* 29, 194. doi: 10.3390/molecules29010194

Saeed, R., Ahmed, D., and Mushtaq, M. (2022). Ultrasound-aided enzyme-assisted efficient extraction of bioactive compounds from *Gymnema sylvestre* and optimization as per response surface methodology. *Sustain. Chem. Pharm.* 29, 100818. doi: 10.1016/j.scp.2022.100818

Sagheer, M. S., Aslam, M. R., Jabbar, S., Ahmad, K., Aslam, N., and Saeed, M. M. (2023). Current updates on pharmacological profile, chemical constituents, and traditional uses of *Gymnema sylvestre* (Gurmar). *RADS J. Pharm. Allied Health Sci.* 1, 68–79. doi: 10.1016/S0367-326X(03)00154-0

Satdive, R. K., Abhilash, P., and Fulzele, D. P. (2003). Antimicrobial activity of *Gymnema sylvestre* leaf extract. *Fitoterapia* 74, 699–701. doi: 10.1016/S0367-326X(03)00154-0

Saxena, A. K., Saxena, S., and Chaudhaery, S. S. (2010). Molecular modeling and docking studies on heat shock protein 90 (Hsp90) inhibitors. *SAR QSAR Environ. Res.* 21, 1–20. doi: 10.1080/10629360903560504

Sharma, D., Yusuf, M., and Asif, M. (2024). *Gymnema sylvestre*: Phytochemistry, pharmacology, and economical perspectives. *J. Adv. Pharmacogn.* 4, 79–90.

Srinivasan, K., and Kumaravel, S. (2016). Unraveling the potential phytochemical compounds of *Gymnema sylvestre* through GC-MS study. *Int. J. Pharm. Pharm. Sci.* 8, 1–4.

Suwan, N., Baison, W., and Chuajedton, A. (2022). Purification of *Gymnema inodorum* leaf extract and its antifungal potential against *Colletotrichum gloeosporioides*. *Proc. Natl. Acad. Sci. India Sect. B. Biol. Sci.* 92, 667–677. doi: 10.1007/s40011-022-01364-0

Thangavelu, D., Thangavelu, T., Vembu, T., Venkatachalam, K., Selladurai, E., and Jegadeesan, M. (2012). Pharmacognostic and phytochemical studies on *Gymnema sylvestre* R. Br. hairy variant. *Int. J. Pharm. Phytopharmacol. Res.* 2, 143–147.

Tiwari, P., Mishra, B. N., and Sangwan, N. S. (2014). Phytochemical and pharmacological properties of *Gymnema sylvestre*: An important medicinal plant. *Biomed. Res. Int.* 2014, 830285. doi: 10.1155/2014/830285

Vats, S., Dey, A., Bhandari, N., Kumari, K., and Kaushal, C. (2023). *Gymnema sylvestre* R. Br.: Phytochemicals and medicinal properties. *Medicinal Aromatic Plants India* 2, 125–151.

Wadkar, K. A., Magdum, C. S., Patil, S. S., and Naikwade, N. S. (2008). Antidiabetic potential and Indian medicinal plants. *J. Herb. Med. Toxicol.* 2, 45–50.

Zhu, X. M., Xie, P., Di, Y. T., Peng, S. L., Ding, L. S., and Wang, M. K. (2008). Two new triterpenoid saponins from *Gymnema sylvestre*. *J. Integr. Plant Biol.* 50, 589–592. doi: 10.1111/j.1744-7909.2008.00661.x

Zulu, L., Gao, H., Zhu, Y., Wu, H., Xie, Y., Liu, X., et al. (2023). Antifungal effects of seven plant essential oils against *Penicillium digitatum*. *Chem. Biol. Technol. Agric.* 10, 82. doi: 10.1186/s40538-023-00434-3

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 24 April 2025
DOI 10.3389/fpls.2025.1568548

OPEN ACCESS

EDITED BY
Xin Fang,
Chinese Academy of Sciences (CAS), China

REVIEWED BY
Dr. Pawan Kumar,
Agricultural Research Organization
(ARO), Israel
Tianyue An,
Binzhou Medical University, China

*CORRESPONDENCE
Kim L. Johnson
k.johnson@latrobe.edu.au
Matthew T. Welling
m.welling@latrobe.edu.au

RECEIVED 30 January 2025
ACCEPTED 28 March 2025
PUBLISHED 24 April 2025

CITATION
Senevirathne GI, Gendall AR, Johnson KL and Welling MT (2025) Understanding the role of oxylipins in *Cannabis* to enhance cannabinoid production. *Front. Plant Sci.* 16:1568548. doi: 10.3389/fpls.2025.1568548

Understanding the role of oxylipins in *Cannabis* to enhance cannabinoid production

Gayathree I. Senevirathne[1,2], Anthony R. Gendall[1,2,3,4], Kim L. Johnson[1,2,3*] and Matthew T. Welling[1,2,3*]

[1]Australian Research Council Research Hub for Medicinal Agriculture, Department of Ecological Plant and Animal Sciences, School of Agriculture, Biomedicine and Environment, La Trobe University, Bundoora, VIC, Australia, [2]La Trobe Institute of Sustainable Agriculture and Food, Department of Ecological Plant and Animal Sciences, School of Agriculture, Biomedicine and Environment, La Trobe University, Bundoora, VIC, Australia, [3]Australian Research Council Research Hub for Protected Cropping, Department of Ecological Plant and Animal Sciences, School of Agriculture, Biomedicine and Environment, La Trobe University, Bundoora, VIC, Australia, [4]Australian Research Council Research Hub for Sustainable Crop Protection, Department of Ecological Plant and Animal Sciences, School of Agriculture, Biomedicine and Environment, La Trobe University, Bundoora, VIC, Australia

Phytocannabinoids are medically important specialized defense compounds that are sparsely distributed among plants, yet *Cannabis sativa* can synthesize unprecedented amounts of these compounds within highly specialized surface cell factories known as glandular trichomes. The control mechanisms that allow for this high level of productivity are poorly understood at the molecular level, although increasing evidence supports the role of oxylipin metabolism in phytocannabinoid production. Oxylipins are a large class of lipid-based oxygenated biological signaling molecules. Although some oxylipins are known to participate in plant defense, roles for the majority of the ca. 600 plant oxylipins are largely unknown. In this review, we examine oxylipin gene expression within glandular trichomes and identify key oxylipin genes that determine the fate of common lipid precursors. Mechanisms by which oxylipins may be interacting with phytocannabinoid metabolism, as well as specialized plant metabolism more broadly, are discussed and a model summarizing these contributions proposed.

KEYWORDS

glandular trichomes, green leaf volatiles, jasmonates, lipoxygenase, oxylipins, specialized metabolites

1 Introduction

Cannabis sativa L. (*Cannabis*) is a chemically complex plant that has a repertoire of more than 500 diverse molecules that span several classes of specialized metabolites, including a large class of medically important isoprenylated resorcinol aliphatic polyketides known as phytocannabinoids (PCs). PCs serve as the active pharmaceutical ingredients in medicines used as adjunct therapies for drug-resistant epilepsies (Epidiolex) and multiple sclerosis (Sativex), as well as nausea, pain, and loss of appetite (Dronabinol)

(Lichtman et al., 2018; Thiele et al., 2018). In plants, PCs play a critical role in defense responses, including protection from UV radiation, herbivorous insects, and microbial pathogens, but the exact role of PCs in *Cannabis* is poorly understood (Tanney et al., 2021). While many of the enzymes responsible for producing these metabolites have been elucidated, the control mechanisms governing their synthesis remain underexplored (Stout et al., 2012; Welling et al., 2023). Increasing evidence suggests that PC production may be mediated by oxylipins (Bailey, 2019; Booth et al., 2020; Burgel et al., 2020; Apicella et al., 2022; Garrido et al., 2022; Welling et al., 2023; Table 1).

Oxylipins are structurally and functionally diverse, lipid-based biological signaling molecules that participate in plant immunity (Blée, 2002; Christensen et al., 2016; Ameye et al., 2018; Wasternack and Feussner, 2018; Toporkova et al., 2024; Figure 1). Oxylipins play vital roles in regulating biological functions in plants including growth and development, innate and induced defense mechanisms as response to diverse environmental elicitors and as secondary messengers activating defense compounds (McGarry et al., 2024; Wang et al., 2024; Yokoyama et al., 2025). They are produced by the activity of lipoxygenases (LOXs) and various enzymes of the cytochrome P450 CYP74 family, such as allene oxide synthase (AOS) and hydroperoxide lyase (HPL) (Song and Brash, 1991; Howe et al., 2000; Wasternack and Feussner, 2018). Jasmonates (JAs), including jasmonate-isoleucine (JA-Ile), jasmonic acid, and its methylated form methyl jasmonate (MeJA) are oxylipin-derived phytohormones formed from the LOX-AOS cascade (Wasternack and Hause, 2013; Wasternack and Feussner, 2018; Ghorbel et al., 2021).

JAs induce specialized metabolism in diverse plant lineages, including *Solanum lycopersicum* (tomato) and *Lavandula* spp. (lavender), while also regulating the development of glandular trichomes and flowers where PCs are concentrated (Boughton et al., 2005; Chen et al., 2018; Dong et al., 2022a; Kianersi et al., 2022; Yan et al., 2022). More recently, JAs have been shown to enhance the levels of PCs, although the intricate mechanism(s) driving this increase in PC production is not fully understood (Table 1). There is also an emerging hypothesis that green leaf volatiles (GLVs) formed via the LOX-HPL branch of oxylipin metabolism may provide substrate for the synthesis of hexanoic acid, the polyketide starter unit for PC production, although only limited evidence has been reported supporting this interaction (Stout et al., 2012; Welling et al., 2023). There has been considerable gene expansion in the *LOX* family in *Cannabis* compared with other species, and expression of some *LOXs* have been localized to the glandular trichomes where PCs are synthesized, suggesting that these may have acquired specialized roles in trichome biology and PC metabolism (Livingston et al., 2020; Borrego et al., 2023; Fayaz et al., 2023; Welling et al., 2023). Collectively, these data serve as evidence for the distinct yet incompletely understood role of the oxylipin pathway in PC production. In this review, we use a combination of phylogenetic and transcriptomic analyses to investigate the mechanisms by which oxylipins may influence PC production in *Cannabis* and propose a model highlighting these potential interactions.

2 *Cannabis:* a versatile and underutilized crop

2.1 Origin, classification and uses

Cannabis is a predominantly dioecious, herbaceous annual plant in the Cannabaceae family that originated in central Asia and now has a broad distribution (Kovalchuk et al., 2020). This genus comprises only one species but exhibits high polymorphism, with subspecies: *C. sativa* ssp. *sativa*, *C. sativa* ssp. *indica*, and *C. sativa* ssp. *rudelaris* characterized by varying chemical and morphological features (Small, 2015; Hazekamp et al., 2016). Although the plant is known for its intoxicating properties, it has been used over millennia for medicine, fiber, and oil extraction (Omare et al., 2021; Shakil et al., 2021). Current applications also extend to phytoremediation, functional foods, and as ornamental plants (Frassinetti et al., 2018; Wu et al., 2021; Hesami et al., 2022b). Classification of *Cannabis* has often been based on its end use, with recreational drug types typically referred to as 'marijuana' and fiber or seed types 'industrial hemp'. Plants can also be classified into chemical phenotypes (chemotypes) based on their proportions of Δ^9-tetrahydrocannabinol (Δ^9-THC) or cannabidiol (CBD) PCs, or on the level of Δ^9-THC, the principal intoxicant found in *Cannabis* (Jin et al., 2021; Hesami et al., 2022a; Salamone et al., 2022).

Despite the multiple end uses of *Cannabis*, research on this plant has been constrained over several decades due to its narcotic status, which has limited germplasm exchange and the application of modern genetic improvement strategies (Welling et al., 2016). Recent reforms to the cultivation and use of *Cannabis* by many legislators have enabled strong growth in licit *Cannabis* markets and these are projected to increase rapidly in the coming decades (de Brito Siqueira et al., 2023). Similarly, legislative reform has allowed for an increase in *Cannabis* research, with recent efforts focusing on establishing genomic resources for *Cannabis* and improving understanding on factors driving PC yield, such as flowering behavior as well as trichome initiation and hormonal regulation (Livingston et al., 2020; Hurgobin et al., 2021; Livingston et al., 2021; Apicella et al., 2022; Toth et al., 2022; Sands et al., 2023; Steel et al., 2023; Dowling et al., 2024; Huang et al., 2024).

Abbreviations: AOS, Allene oxide synthase; CBC, Cannabichromene; CBCA, Cannabichromenic acid; CBD, Cannabidiol; CBDA, Cannabidiolic acid; CBDAS, Cannabidiolic acid synthase; CBGA, Cannabigerolic acid; CBGAS, Cannabigerolic acid synthase; ECS, Endocannabinoid system; GLV, Green leaf volatiles; GPP, Geranyl pyrophosphate; GPPS, Geranyl pyrophosphate synthase; HPL, Hydroperoxide lyase; JA/JAs, Jasmonate/Jasmonates; JA-Ile, Jasmonic acid isoleucine; LOX, Lipoxygenase; MeJA, Methyl jasmonate; MEP, Methylerythritol 4-phosphate; OA, Olivetolic acid; OAC, Olivetolic acid cyclase; *cis*-(+)-OPDA, *cis*-(+)-12-oxo-phytodienoic acid; PC, Phytocannabinoid; PUFA, Polyunsaturated fatty acid; ROS, Reactive oxygen species; Δ^9-THC, Δ^9-Tetrahydrocannabinol; Δ^9-THCA, Δ^9-Tetrahydrocannabinolic acid; TKS, Tetraketide synthase.

TABLE 1 Effect of JAs on PC production and glandular trichome density in *Cannabis*.

Cannabis variety	MeJA/JA treatment	Effect on PC production/trichome density	Changes in gene expression	Reference
Medicinal (Skunk)	100 µM jasmonic acid/0.3 mM MeJA applied to cell cultures	No detectable level of Δ^9-THC	No detectable change in *THCAS*	Flores-Sanchez et al., 2009
Industrial hemp (Santhica)	0.03, 0.1, 0.3, 1, and 3 mM jasmonic acid applied to 15-day old seedlings	Not reported	Higher expression of *LOX2* and *4-coumarate-CoA ligase-like 7*	Behr et al., 2018
Industrial hemp (Cherry, Cherry Blossom, Canada)	44.9 mg/L MeJA applied as foliar and roots as separate treatments for 7weeks old plants	A significant increase in CBDA (>2-fold) in foliar applications in all three varieties A decrease in Δ^9-THC in and increase in total PC content	Not reported	Bailey, 2019
Medicinal (White Tangy Haze)	0, 100, 500, and 1000 µM total plant spray MeJA after 2 weeks of flower development	A significant increase inΔ^9-THC in 1000 µM at week 1 (>0.1-fold) and week 4 (>0.07-fold)	Not reported	Apicella et al., 2022
Medicinal (Beatriz)	0, 0.1, 1.0, and 10 mM MeJA, 8 applications (once a week) after shifting to reproductive photoperiod	An increase in Δ^9-THCA (>0.08-fold) and CBDA (>0.15-fold) in response to 0.1 mM A decrease in Δ^9-THCA and CBDA in response to 10 mM	No significant change in the expression of *AOC-1*, *AOC-2*, *PT4*, *THCAS*, and *CBDAS*	Garrido et al., 2022
Not reported	150 µM MeJA applied to 4-week-old plants	Not reported	Significant increase in expression of *CsLOX1*, *CsLOX2*, *CsLOX5*, *CsLOX8*, *CsLOX10*, *CsLOX11*, *CsLOX12*, *CsLOX3*, *CsLOX17* 2 hours and CsLOX1, *CsLOX2*, *CsLOX5*, *CsLOX10*, *CsLOX11*, *CsLOX12*, CsLOX16, 24 hours post application	Fayaz et al., 2023
Industrial hemp (MW6-15)	0, 1, 7.5, and 15 mM MeJA, foliar spray on 14th days after transitioning into reproductive photoperiod, after initiation of terminal flowering and 7 and 14-days post first application	A significant increase (~1-fold) in total cannabinoids, Δ^9-THCA, THCVA CBDA, CBDVA, and CBCA (> 1-fold) in response to 15 mM	*JMT* responsive to 7.5 and 15 mM MeJA in the leaf and inflorescence. *LOX-L* and *HPL* highly expressed in the leaf and inflorescence	Welling et al., 2023
Hemp (Hot Blond)	100, 200, and 400 µM, foliar spray on 0 and 21 days after transitioning to reproductive photoperiod	A significant increase in glandular trichome density in calyx, A significant increase in CBD (>0.1-fold in both leaves and inflorescence) and Δ^9-THC (>0.3-fold in leaves, >0.1-fold in inflorescence) in response to 100 mM	Not reported	Hahm et al., 2024
CANN97	1 µM MeJA applied 5 times at 3-day intervals to 18-day old plants	A significant increase in leaf sessile glandular trichome density and non-glandular trichome density	Significant increase in the expression of *CsJAZ1, CsJAZ2, CsJAZ3, CsMYC1, CsMYC4, CsCOI1*	Huang et al., 2024
High CBD	100 µM jasmonic acid as foliar spray, for consecutive 8 weeks after shifting to flowering photoperiod	No significant effect on overall glandular trichome density in bracts A decrease in Δ^9-THCA, and CBDA but significant increase (> 0.2-fold) in Δ^9-THC and CBD after 7 and 8 weeks in flowering	Not reported	Oultram et al., 2024
Not reported	Adventitious roots incubated in 0, 50, 100, 150, and 200 µM MeJA for 2 days	A significant increase (>0.5-fold) in CBD in 200 µM	Not reported	Wang Y, et al., 2024

2.2 PCs: important medicinal compounds

Extensive research has been conducted on the pharmacology of the two major PCs Δ^9-THC and CBD, both of which have modulatory effects on the human endocannabinoid system (ECS). Δ^9-THC is a partial agonist of type 1 G protein-coupled cannabinoid receptors (Tagen and Klumpers, 2022). These receptors have important roles in regulating memory, mood, sleep, appetite, inflammation, and pain sensation, and their activation has been shown to have analgesic, neuroprotective, and anti-nausea effects (Mechoulam and Parker, 2013). In contrast, CBD is also thought to modulate the ECS via a less understood mechanism involving other

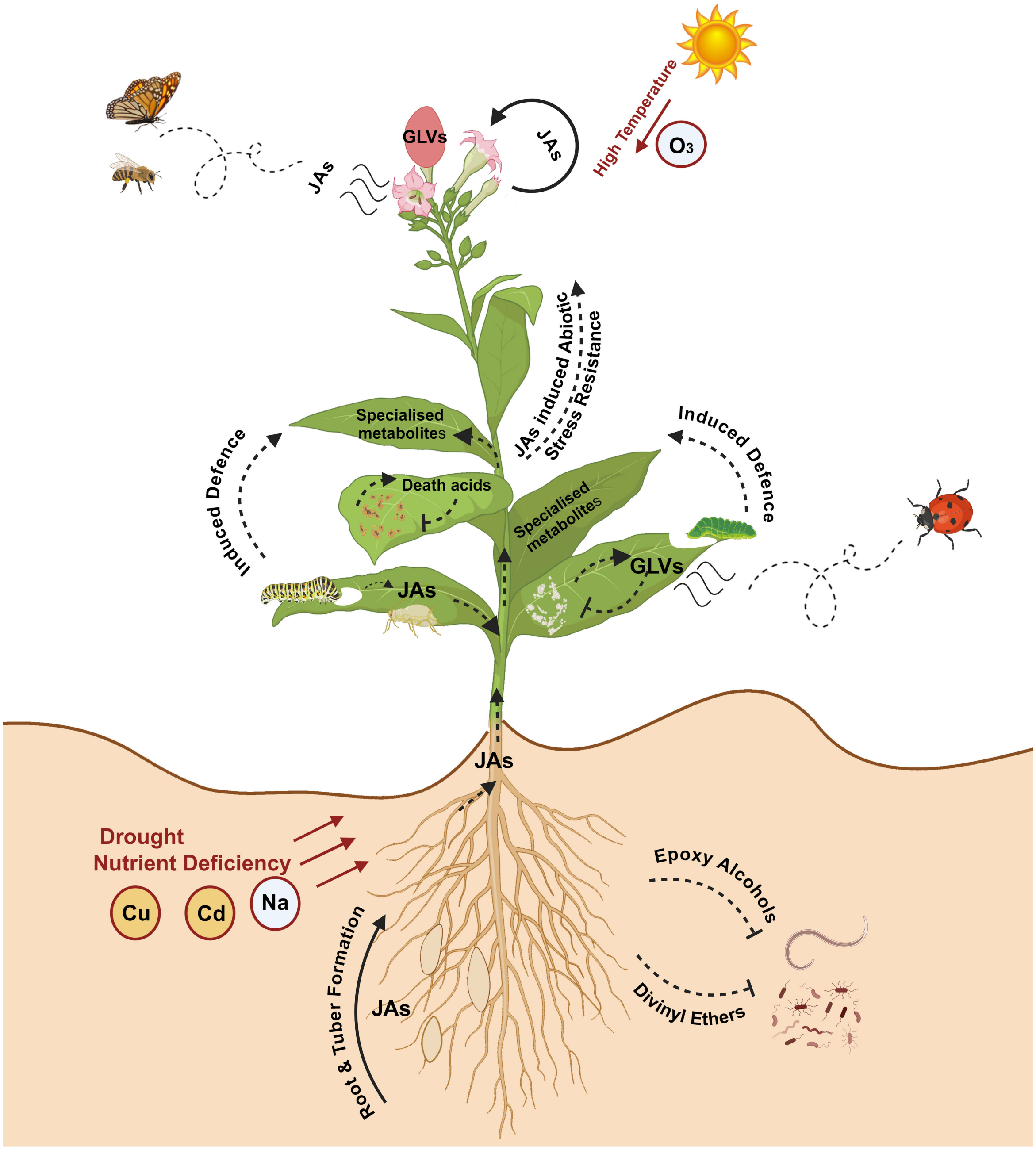


FIGURE 1
Role of oxylipins in plant growth, development, biotic and abiotic stress tolerance. Oxylipins play multifaceted roles in regulating plant growth, development, and responses to biotic and abiotic stresses. Jasmonates (JAs) mediate developmental processes, including lateral root formation, tuber formation, flowering, and trichome development. Several JAs are also known to attract pollinators and influence plant reproductive strategies. JAs activate defense genes in plants to protect against biotic and abiotic stress (nutrient deficiency, heavy metals [copper (Cu), cadmium (Cd)] and ozone (O_3)). Green leaf volatiles (GLVs) contribute to fruit ripening and act as direct defense molecules against pathogenic microbes. GLVs also induce other defense responses, such as attracting natural predators and promoting the production of specialized metabolites. Oxylipins like epoxy alcohols and divinyl ethers serve as direct defense molecules against various pests and pathogens. Image created in BioRender.com.

ECS targets and has been reported to counteract the intoxicating properties of Δ^9-THC (Devinsky et al., 2014). Medicines containing Δ^9-THC and CBD are currently used in the treatment of pediatric epilepsies (e.g., Dravet syndrome and Lennox-Gastaut syndrome) and to alleviate side-effects of chemotherapy, with many other PCs, such as the varinoid PCs cannabidivarin (CBDV; GWP42006) and tetrahydrocannabivarin (THCV; GWP42004), at various levels of drug development (Chanda et al., 2019; Urits et al., 2019; Britch et al., 2021). Other non-intoxicating and low abundant or minor PCs, including cannabichromene (CBC), cannabigerol (CBG) and their

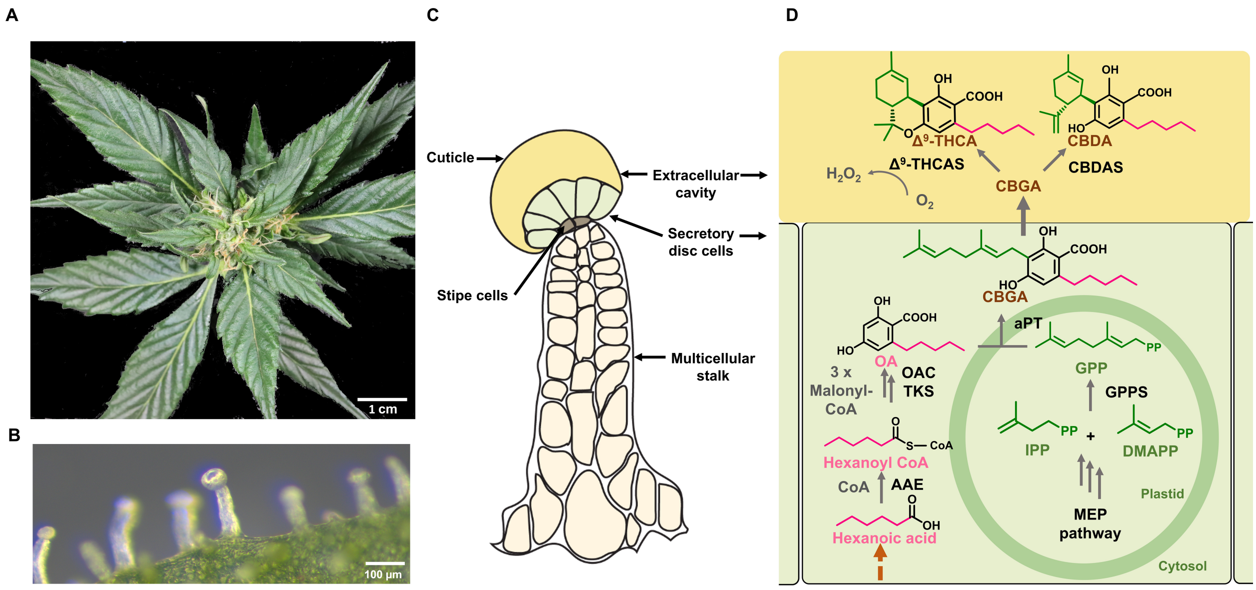


FIGURE 2
Spatial localization of the phytocannabinoid (PC) pathway in *Cannabis*. **(A)** Image of *Cannabis* female inflorescence. **(B)** Microscopic view of a capitate stalked glandular trichome on a perigonal bract. **(C)** Schematic representation of a capitate stalked glandular trichome, the major site of PC biosynthesis and storage. This consists of multicellular stalk (cream), stipe cells at the base of the trichome head (grey), secretory disc cells (light green) and large extracellular storage cavity (yellow). **(D)** The PC biosynthesis pathway in *Cannabis*. The methylerythritol phosphate (MEP) pathway takes place in non-photosynthetic chloroplasts of glandular trichome disc cells, while the polyketide pathway is localized in the cytosol. Geranyl pyrophosphate (GPP) is formed from the MEP pathway. The polyketide pathway requires an activated fatty acid starter unit (e.g., hexanoyl-CoA) and three molecules of malonyl-CoA and gives rise to olivetolic acid (OA). Prenylation with GPP by an aromatic prenyltransferase forms CBGA. This is exported to the extracellular cavity where cannabinoid synthases perform oxidative cyclization of the isoprenoid moiety of CBGA, forming the major acid PCs; CBDA and Δ^9-THCA. In the final reaction, H_2O_2 is produced as a biproduct when CBGA is converted to PCs in the presence of O_2. Solid dark grey arrows represent known steps in pathways while the brown dashed arrows represent steps of the pathway that have not been resolved. OA, olevetolic acid; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl pyrophosphate; CBGA, cannabigerolic acid; CBDA, cannabidiolic acid; Δ^9-THCA, Δ^9-tetrahydrocannabinolic acid; AAE, acyl activating enzyme; TKS, tetraketide synthase; OA, olivetolic acid cyclase; GPPS, geranyl pyrophosphate synthase; aPT, aromatic prenyltransferase; CBGAS, cannabigerolic acid synthase, CBDAS, cannabidiolic acid synthase; D_9-THCAS, D9-tetrahydrocannabinolic acid synthase.

carboxylated analogues, cannabichromenic acid (CBCA) and cannabigerolic acid (CBGA), are also considered promising therapeutics (Borrelli et al., 2013; Anderson et al., 2021; Cabrera et al., 2021).

2.3 Glandular trichomes are multicellular structures that synthesize, secrete, and store enormous amounts of PCs

Cannabis produces glandular as well as non-glandular trichomes on the surface of leaves and other parts of the plant body (Livingston et al., 2020; Xie et al., 2023). However, it is the capitate stalked glandular trichomes on modified leaves (perigonal bracts) surrounding female flowers that are the primary site for PC biosynthesis (Figure 2A). These trichomes are multicellular organs, consisting of a stalk, a disc of secretory cells, and a large globular head containing an extracellular cavity formed by cell wall delamination that stores the resinous compounds (Hammond and Mahlberg, 1973; Livingston et al., 2020; Figures 2B, C). These specific trichome morphotypes can produce substantial amounts of PCs, with these compounds contributing up to 40% of their dry weight (w/w) (Livingston et al., 2020). Capitate stalked glandular trichome initiation and development are coordinated with the onset of flowering and likely influenced by several phytohormones, including gibberellins, cytokinins, and JAs, as have been established in numerous plant species (Pattanaik et al., 2014; Li et al., 2021; Huang et al., 2024).

The biosynthesis of PCs traverses across several organelles within the secretory disc cells of glandular trichomes (Figure 2D). The precursors of PCs, olivetolic acid (OA), and geranyl pyrophosphate (GPP) are synthesized by the polyketide pathway in the cytosol and the methylerythritol 4-phosphate (MEP) pathway in plastids, while the final step occurs in the extracellular cavity (Livingston et al., 2022; Figures 2C, D). OA is formed by a type III polyketide synthase (tetraketide synthase; TKS), which produces a tetraketide intermediate, while the accessory protein, OA cyclase (OAC) catalyzes a C2-C7 aldol condensation (Gagne et al., 2012). Together, these enzymes form OA using three molecules of malonyl-CoA and an activated fatty acid, hexanoyl-CoA (Stout et al., 2012). Prenylation of OA with GPP produce the first PC, CBGA, a step catalyzed by an aromatic prenyltransferase, CBGA synthase (CBGAS) (Luo et al., 2019; Sands et al., 2023). CBGA is then transported to the extracellular cavity and converted to Δ^9- THCA and CBDA by Δ^9- THCA and CBDA synthases, respectively (Taura et al., 1996; Sirikantaramas et al., 2005).

This compartmentalization is thought to be attributed to the cytotoxicity of PCs and their bi-product H_2O_2 which is formed at a molar ratio of 1:1 during the oxidative cyclisation of the isoprenoid residue of CBGA (Sirikantaramas et al., 2004, 2005; Figure 2D). The acidic PCs are then non-enzymatically decarboxylated over time or when subjected to heat, forming the bioactive neutral PCs, such as Δ^9-THC and CBD (Sirikantaramas and Taura, 2017).

3 The origin of PC fatty acid precursors is unresolved

A key unanswered question in PC biosynthesis is the origin of the fatty acid hexanoic acid, which, once activated, serves as the C6 carbon starter unit for OA (Welling et al., 2023; Figure 2D). Several metabolic pathways could synthesize hexanoic acid within the secretory disc cells of capitate stalked trichomes. In the trichome glands of *Petunia*, C6-C8 straight-chain acyl acids are formed from branched-chain amino acid catabolism and α-ketoacid elongation (Kroumova et al., 1994). This pathway requires the activity of four proteins, yet transcripts for these genes have not been reported in *Cannabis* trichomes (Kroumova et al., 1994; Marks et al., 2009). A second pathway involves *de novo* fatty acid synthesis. This is supported by trichome-specific expressed sequence tags for a *3-oxoacyl- [acyl-carrier-protein (ACP)] reductase* (*fabG*), as well as the association of this gene with changes in PC composition in a genome-wide association study (Marks et al., 2009; Welling et al., 2020). While C6-specific thioesterases required to terminate fatty acid elongation have not been recovered from trichome cells, this enzyme could reduce the β-keto group following the first two condensation reactions of acetyl CoA, forming hexanoic acid, and, instead of leaving the condensing enzyme, participate in additional reactions with malonyl CoA (Horper and Marner, 1996; Stout et al., 2012). Contrary to this hypothesis is the functional validation of a cytosolic-localized hexanoic acid-specific AAE (hexanoyl-CoA synthetase) which is highly expressed in *Cannabis* trichome secretory disc cells (Stout et al., 2012).

A third hypothetical pathway could involve an oxylipin-based origin (Stout et al., 2012; Welling et al., 2023). This would involve the degradation of polyunsaturated fatty acids (PUFAs), such as linoleic or linolenic acid, and the activity of LOX to form hydroperoxy PUFAs that can be cleaved into C6 aldehydes and progenitors of hexanoic acid (Nakashima et al., 2013; Welling et al., 2023; Figure 3). Genes encoding desaturases, plastid-localized LOXs, and HPL that participate in these reactions are highly expressed in *Cannabis* trichomes (Stout et al., 2012; Balcke et al., 2017), while the secondary cleavage product of LOX/HPL, α-oxo acid 12-oxo-(10*E*)-dodecenoic acid ((10*E*)-traumatin) (Figure 3), has recently been correlated with PC content in *Cannabis* inflorescences following treatment with the oxylipin phytohormone MeJA (Welling et al., 2023; Table 1). C18 PUFAs and their oxylipin derivatives are also highly abundant in type VI stalked glandular trichomes of *Solanum* spp. and genes encoding acyl-hydrolyzing GDSL-type lipases that can release free C18 fatty acid substrates from galactolipids have been identified in *Cannabis* quantitative trait loci (QTL) for PC content (Cropano et al., 2022). The following sections provide a detailed discussion of oxylipins in *Cannabis* and highlight key areas of research that could be used to understand the interaction between oxylipins and PC production.

4 Oxylipin pathway and metabolism

4.1 Key metabolic branch points determine the fate of oxylipin molecules and their biological functions

In plants, oxylipins are generated through the oxidation of PUFAs. PUFA oxidation can occur enzymatically by dioxygenases or nonenzymatic by the activity of free radicals (e.g., singlet oxygen ($^1O_2$)) (Brash, 1999; Zoeller et al., 2012; Figure 3). The resulting hydroperoxides are highly reactive and act as key substrates in oxylipin metabolism, with the fate of oxylipin molecules decided by several enzymatic branch points including (1) the previously described HPL branch that forms C6 aldehydes, alcohols and acyl esters (collectively known as GLVs) concomitant with 12-oxo-(9*Z*)-dodecenoic acids ((9*Z*)-traumatin), (2) the JAs-forming AOS pathway, (3) the divinyl ether synthase (DES) pathway that forms divinyl ethers, and (4) the epoxy alcohol synthase/peroxygenase pathways that produce epoxy hydroxy PUFAs (Wasternack and Hause, 2013; Christensen et al., 2015, 2016; Figure 3). Oxylipins within each branch can have distinct roles (Schlotzhauer et al., 1996; Figure 3). For example, divinyl ethers and epoxy alcohols act as defense compounds in roots and leaves (Blée, 1998; Hamberg, 1999; Weber et al., 1999; Hamberg, 2004; Sanadhya et al., 2021; Gorina et al., 2022; Toporkova et al., 2024; Figure 1). The HPL and AOS branches are among the most extensively researched, with AOS-derived JAs having broad roles as signaling molecules, in reproduction, plant growth, and development (Scala et al., 2013; Ghasemi Pirbalouti et al., 2014; Ameye et al., 2018; Raza et al., 2021; Figures 1, 3).

4.2 LOX serves as the critical juncture dictating oxylipin fate

LOXs, such as linoleate 13*S*-lipoxygenase (EC 1.13.11.12), serve as primary enzymes that mediate the fate of lipid precursors (Figure 3). These are monomeric non-heme iron (Fe)-containing dioxygenases widely distributed in plants that convert PUFAs into hydroperoxides in a two-step reaction, involving the reduction of Fe^{3+} to Fe^{2+} by proton-coupled electron transfer, followed by oxygen insertion (Egmond et al., 1973; Goldsmith et al., 2002). LOXs are categorized into two main functional classes based on the

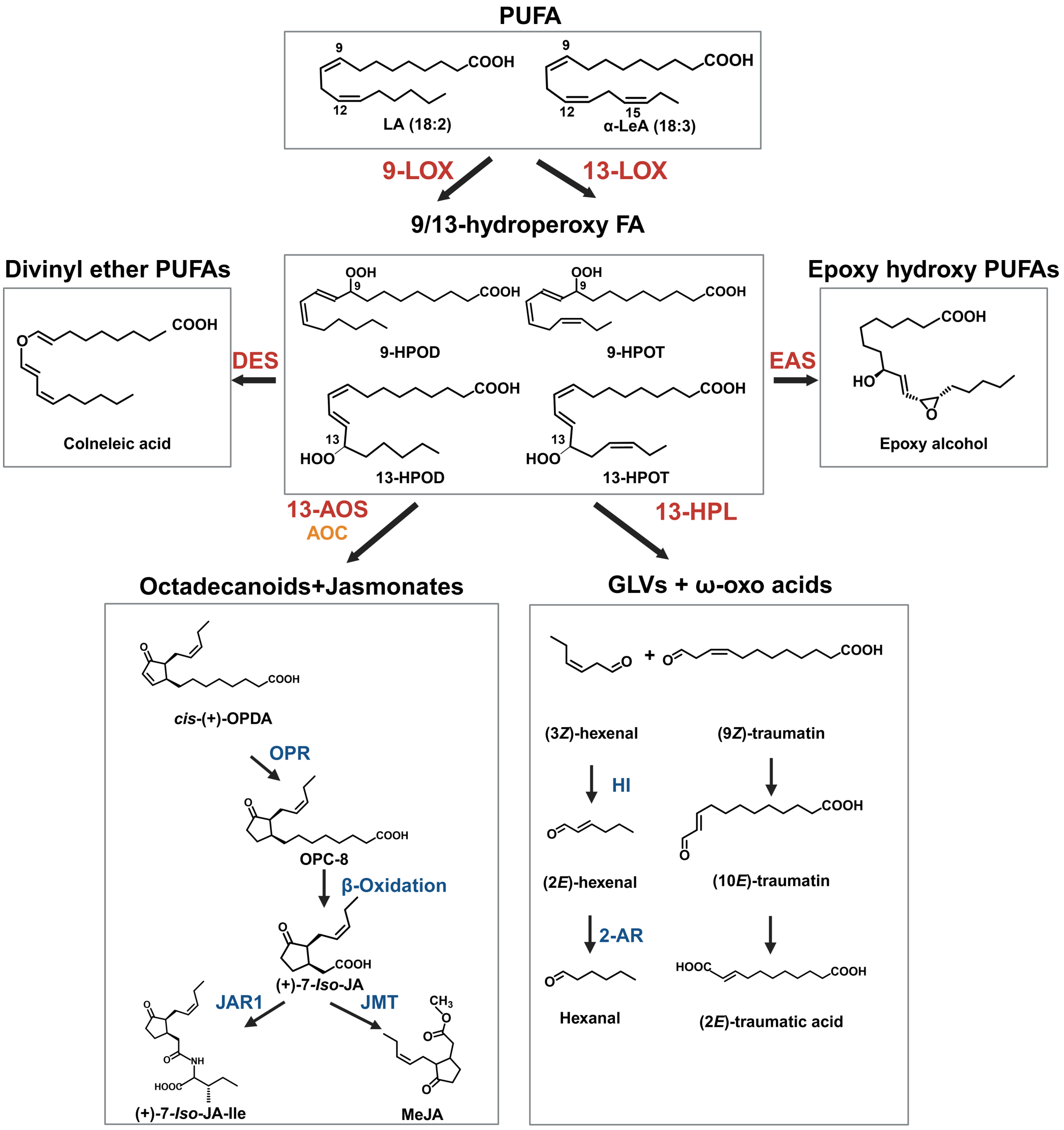


FIGURE 3
Key oxylipins in the oxylipin pathway. The oxylipin pathway starts with the hydroperoxidation of polyunsaturated fatty acids (PUFA) by lipoxygenase (LOX) enzymes. The position of oxygenation is either at the 9th or 13th carbon from the carbonyl end of the fatty acid. The resulting polyunsaturated hydroperoxides include 9-/13-hydroperoxy linoleic acid (9/13-HPOD) and 9-/13-hydroperoxy linolenic acid (9/13-HPOT). This reaction represents a key branchpoint for oxylipin biosynthesis and provides substrate for several enzymatic pathways, including allene oxide synthase (AOS), hydroperoxide lyase (HPL), divinyl ether synthase (DES), epoxy alcohol synthase (EAS) and peroxygenase (POX) pathways. The AOS and allene oxide cyclase (AOC) cascade leads to the formation of jasmonates (JAs), while the HPL branch leads to the production of green leaf volatiles (GLVs) and 12-oxo acids/traumatins. The other three pathways produce several groups of molecules that act in plant defense and signaling, including epoxy alcohols and divinyl ethers (e.g., colneleic acid). Abbreviations for compounds: α-LeA, α-linolenic acid; LA, Linoleic acid, 9-HPOD, (9*S*)-hydroperoxy-(10*E*,12*Z*)-octadecadienoic acid; 9-HPOT, (9*S*)-hydroperoxy-(10*E*,12*Z*,15*Z*)-octadecatrienoic acid; 13-HPOD, (13*S*)-hydroperoxy-(9*Z*,11*E*)-octadecadienoic acid; 13-HPOT, (13*S*)-hydroperoxy-(9*Z*,11*E*,15*Z*)- octadecatrienoic acid; epoxy alcohol, 13(*S*)-epoxy-9(*S*)-hydroxy-10(*E*)-octadecenoic acid; *cis*-(+)-OPDA, *cis*-(+)-12-oxo-phytodienoic acid; OPC-8, 3-oxo-2-(2-pentenyl)-cyclopentane-1-octanoic acid; (+)-7-*iso*-JA, jasmonic acid; (+)-7-*iso*-JA-Ile, jasmonic acid isoleucine conjugate;. (9*Z*)-traumatin, 12-oxo-(9*Z*)-dodecenoic acid; (10*E*)-traumatin, 12-oxo-(10*E*)-dodecenoic acid, (2*E*)-traumatic acid; (2*E*)-dodecanedioic acid; JAR1, jasmonoyl amino acid conjugate synthase; OPR, 12-oxo-phytodienoic acid reductase; AOC, allene oxide cyclase, HI, hexenal isomerase. 2-AR, 2-alkyl reductase.

carbon position they catalyze. For the 9-LOX and 13-LOX this is at the 9th and 13th position from the carbonyl carbon in the fatty acid, respectively, and can also be distinguished as type I and type II based on subcellular localization (Figure 3). In addition, different LOX paralogs with the same regiospecificity can control different branches of the oxylipin pathway but are spatially separated. For example, in *Zea mays*, the JA-producing 13-LOX ZmLOX8 is localized to the chloroplast, while a GLV-producing 13-LOX ZmLOX10 is localized to non-photosynthetic plastids (Christensen et al., 2013). While often associated with defense process, impaired *LOX* function has resulted in changes in plant reproduction, fruit ripening and root development (Kolomiets et al., 2001; Fortes et al., 2004; Liu and Han, 2010; Caldelari et al., 2011; Vogt et al., 2013; Liu et al., 2020).

TABLE 2 LOC IDs of *Cannabis* 21 LOXs and different nomenclature used in other studies.

Nomenclature[1]	LOC ID	AA length	Major domains			Annotation	Nomenclature of genes used by different authors	
			Protein family IPR000907	PLAT/ LH2 (IPR001024)	C Terminal (IPR013819)		Fayaz et al. (2023)	Welling et al. (2023)
CsLOX1	LOC115719608	860	21-860	18-161	163-860	probable linoleate 9S-lipoxygenase 5	CsLOX8	
CsLOX2	LOC115718693	955	116-955	113-256	258-955	probable linoleate 9S-lipoxygenase 5	CsLOX4	
CsLOX3	LOC115724062	873	32-873	29-172	174-873	probable linoleate 9S-lipoxygenase 5	CsLOX18	
CsLOX4	LOC115709296	859	18-859	15-160	162-869	probable linoleate 9S-lipoxygenase 5	CsLOX1	
CsLOX5	LOC115720291	848	34-848	5-149	151-848	probable linoleate 9S-lipoxygenase 5	CsLOX5	
CsLOX6	LOC115721132	859	42-859	18-160	162-859	probable linoleate 9S-lipoxygenase 5	CsLOX6	
CsLOX7	LOC115719336	871	29-871	21-163	165-871	probable linoleate 9S-lipoxygenase 5	CsLOX7	
CsLOX8	LOC115722276	869	51-869	19-161	163-869	probable linoleate 9S-lipoxygenase 5	CsLOX21	
CsLOX9	LOC115721268	875	52-875	25-167	169-875	probable linoleate 9S-lipoxygenase 5	CsLOX3	
CsLOX10	LOC115722275	868	45-868	18-160	162-868	probable linoleate 9S-lipoxygenase 5	CsLOX20	
CsLOX11	LOC115723988	868	45-868	18-160	162-868	probable linoleate 9S-lipoxygenase 5	CsLOX19	
CsLOX12	LOC115718785	707	1-707	1-51	54-707	linoleate 9S-lipoxygenase 1	CsLOX2	
CsLOX13	LOC115712696	935	129-935	54-236	238-935	lipoxygenase 6, chloroplastic	CsLOX17	LOX
CsLOX14	LOC115707105	931	125-931	97-236	238-931	linoleate 13S-lipoxygenase 3-1, chloroplastic	CsLOX1	
CsLOX15	LOC115719612	928	114-928	80-227	229-928	linoleate 13S-lipoxygenase 2-1, chloroplastic	CsLOX13	LOX-L
CsLOX16	LOC115719614	926	105-926	78-219	221-926	lipoxygenase 2, chloroplastic	CsLOX14	
CsLOX17	LOC115720530	926	105-926	78-219	221-926	lipoxygenase 2, chloroplastic	CsLOX15	
CsLOX18	LOC115719613	929	121-929	92-232	234-929	linoleate 13S-lipoxygenase 2-1, chloroplastic	CsLOX9	
CsLOX19	LOC115719615	922	105-922	92-228	230-922	linoleate 13S-lipoxygenase 2-1, chloroplastic	CsLOX12	
CsLOX20	LOC115719616	716	22-716	–	23-716	linoleate 13S-lipoxygenase 2-1, chloroplastic	CsLOX11	
CsLOX21	LOC115719617	906	83-906	75-212	214-906	linoleate 13S-lipoxygenase 2-1, chloroplastic	CsLOX10	

[1] Borrego et al., 2023 LOX nomenclature.

5 Phylogenetic insights and structural features of *Cannabis* LOX proteins

5.1 Phylogenetic analysis of the *Cannabis* LOX protein family reveals distinct clades that may have specific metabolic functions

To explore the diversity of LOX gene family in *Cannabis*, a phylogenetic tree was constructed using protein sequences predicted from the *Cannabis* var. CBDRx genome [NCBI GenBank Accession: GCA_900626175.2] (Supplementary Table S1). Of the twenty-one *Cannabis* LOXs, all had the PLAT/LH2 domain (IPR001024) at the N terminus except CsLOX20 (Table 2, Supplementary Table S1). They also contained the C-terminal domain IPR013819 and were classified as LOX super family (IPR000907) members, which is consistent with authentic LOX proteins (Chen et al., 2015; Song et al., 2016). Twelve were clustered with the type-I cytosolic-localized 9-LOXs, while nine were clustered with the type-II plastidial-localized 13-LOX sub-class (Figure 4A). Within the 13-LOX sub-group, CsLOX13 and CsLOX14 form a clade with LOX orthologs associated with JA synthesis, while CsLOX15-CsLOX17 form a distinct clade that shows homology with AtLOX2, a dual functioning ortholog involved in both JA and GLV synthesis (Chen et al., 2004; Allmann et al., 2010; Chauvin et al., 2013; Christensen et al., 2013; Shen et al., 2014; Mochizuki et al., 2016; He et al., 2020; Figure 4A).

5.2 Enhanced expression of *Cannabis LOX* in trichomes and floral tissues

We hypothesized that *LOX*s which have high expression in PC producing tissues, such as trichomes, perigonal bracts and inflorescences may have important roles in PC production. For example, some 9-LOX isoforms are exclusively expressed in roots where they participate in the synthesis of defense compounds (e.g., etherolenic acid, colnelenic acid) against root pathogens (Grechkin et al., 1997; Sanadhya et al., 2021). *CsLOX7*, *CsLOX15, CsLOX16* and *CsLOX17* consistently showed high levels of expression in the PC-producing tissues across two medicinal cannabis chemotypes, while *CsLOX15* was predominantly expressed in the trichomes (Figures 4B, C), indicating its potential involvement in trichome biochemistry. The recent heterologous expression of *CsLOX15* in *Nicotiana attenuata* and *N. benthamiana* plants indicate that this paralog encodes a functional enzyme, although activity was assessed using a single substrate, linoleic acid (Fayaz et al., 2023). Elucidating the molecular context of *CsLOX15* in its native environment will be critical to disentangle its precise role within *Cannabis* glandular trichomes and subsequent impact on PC production.

5.3 CsLOX15 structure - conserved residues in catalytic region

To further understand the structure and function of *CsLOX15* and other *Cannabis* LOXs, we compared the predicted structures of these *Cannabis* proteins with the crystal structure from the *Glycine max* (soybean) GmLOX1. As expected, many of the functional residues were conserved across the CsLOXs, with minor exceptions (Meyer et al., 2008; Offenbacher et al., 2017; Ruddat et al., 2004; Li et al., 2018) (Supplementary Table S2, Supplementary Figure S1). The predicted 3D structure of CsLOX15 shows the β-sheet rich N-terminal domain as well as the α-helix rich long carboxylic domain within the catalytic site, which facilitates electron transfer and the insertion of oxygen (Figures 5A, B, Supplementary Figure S1). The predicted 3D structure of CsLOX15, however, indicates that the side chain of W585 near the Fe ligand interacts with residue H584 (GmLOX1: H499) (Figures 5C, D). This may reduce space around Fe ligand by distorting the octahedral geometry, which could impact electron transfer (Steczko et al., 1992; Minor et al., 1996; Tomchick et al., 2001; Liavonchanka and Feussner, 2006; Figure 5A). We also observed that the N780 residue (GmLOX1: N694 involved in Fe binding) was positioned further away from the Fe ligand which could further weaken contact and increase flexibility (Figures 5C, D), Modification of N694H in GmLOX1 reduced flexibility due to stronger H bonding in the coordination sphere and resulted in reduced catalytic activity (Holman et al., 1998; Schenk et al., 2003; Segraves et al., 2006). The application of electron paramagnetic resonance (EPR) and magnetic circular dichroism (MCD) hydrogen deuterium exchange mass spectrometry will further aid in resolving the 3D structure and catalytic center of LOXs in non-model species such as *Cannabis* (Holman et al., 1998; Hu et al., 2019; Gaffney, 2020).

Analysis of the *in silico* predicted 3D structure of CsLOX15 also indicates changes in the substrate entry pocket which could potentially impact substrate specificity. Strong H bonds from residues H248-E256-N534 at the entrance of the substrate binding site of GmLOX1 have been partially lost in CsLOX15 (Figures 5E, F) (Skrzypczak-Jankun et al., 2001). Replacement of these with hydrophobic residues leucine and threonine as seen in CsLOX15 may alter the H bond network and expand the PUFA entry point (Youn et al., 2006). Replacement of V340 with the bulkier T259 residue (GmLOX1) also appears to have expanded the entry pocket (Figures 5E, F). Substrate orientation may also be affected in CsLOX15. For example, there was less distance between A641 (GmLOX1: T556) and R797 (GmLOX1: R707) within the substrate cavity which could potentially affect pocket volume, substrate entry and orientation (Hornung et al., 1999; Hughes et al., 2001; Hershelman et al., 2019; Figures 5G, H). Understanding the impact of these changes is complicated by the fact that there are no crystal structures of plant LOXs bound with linoleic acid or α-linolenic acid (Knapp and Klinman, 2003; Newie et al., 2016). Using *in-silico* tools such as Autodock to model the interaction between fatty acid and LOX proteins would provide deeper insights into substrate-enzyme interactions (Naimuzzaman et al., 2025).

Variations in the predicted tertiary structure of CsLOX15 compared to GmLOX1 suggest potential for differences in catalytic activity that might allow CsLOX15 to catalyze reactions utilizing numerous fatty acid substrates. As *Cannabis* glandular trichome disc cells are proposed to have non-photosynthetic chloroplasts, the role of CsLOX15 may facilitate the mobilization of storage lipids to provide carbon for producing defense compounds (Livingston et al., 2022). To meet the carbon demand, trichomes reliant on the supply of

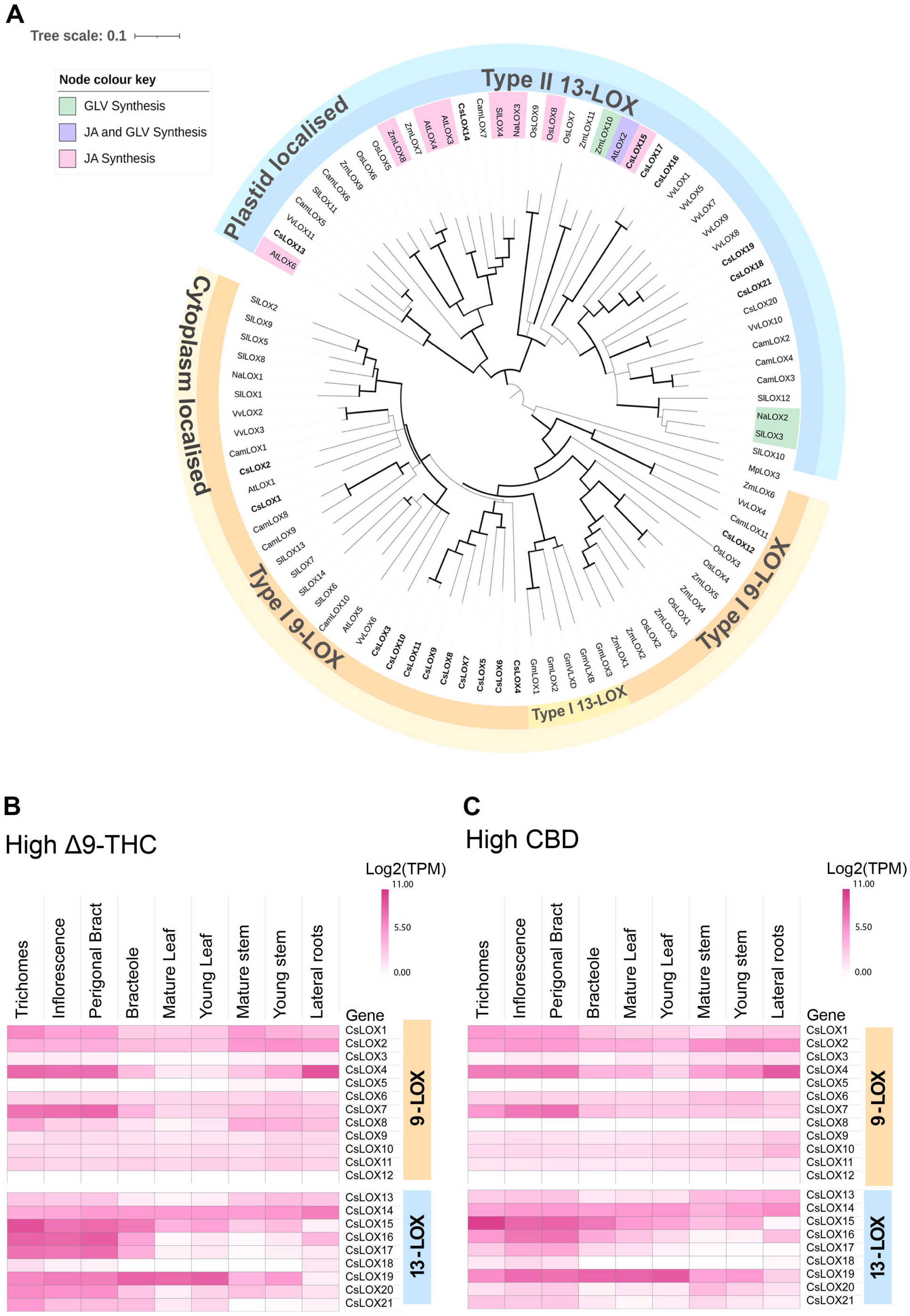


FIGURE 4
Phylogenetic relationship of LOX proteins and tissue series gene expression in *Cannabis*. **(A)** Phylogenetic relationship of lipoxygenase (LOX) proteins in *Cannabis* to several plant species, including *Arabidopsis thaliana* (AtLOX), *Camelia sinensis* (CamLOX), *Nicotiana attenuata* (NaLOX), *Oryza sativa* (OsLOX) *Solanum lycopersicum* (SlLOX), *Vitis vinifera* (VvLOX) and *Zea mays* (ZmLOX). Branches with Bootstrap values above 80 shown by greater line thickness (black). *Cannabis* LOX proteins are shown in bold. Functionally validated genes for corresponding LOXs are shaded as indicated by the key. Absolute expression (Log2 (mean TPM) of *LOX* genes in different tissue types of **(B)** high Δ^9-THC and **(C)** high CBD chemotypes. RNA-Seq data reported by Jost et al. (2025). Protein sequences of *Arabidopsis thaliana* (At), *Cannabis sativa* (Cs), *Camelia sinensis* (Cam), *Nicotiana attenuata* (Na), *Glycine max* wild type (Gm) *Oryza sativa* (Os), *Solanum lycopersicum* (Sl) *Zea mays* (Zm) *and Vitis vinifera* (Vv) were used for phylogenetic analysis and PLAT/LH2 and lipoxygenase (LOX) domains identified using InterPro[1] (v.98.0). Non-*Cannabis* sequences without both domains were removed from the analysis. Ninety-one higher plant LOX proteins were used for phylogenetic analysis. The *Marchantia polymorpha* (common liverwort) LOX3 protein served as an outgroup. The alignment was trimmed of spurious sequences and only residues of the Interpro lipoxygenase super family (IPR000907) domain was compared. The multiple sequence alignment was generated using clustal omega (ver. 1.2.2) in Geneious prime (ver. 2023.0.4), and the phylogenetic tree was generated using the neighbor-joining method (Sievers and Higgins, 2021) and visualized using Interactive Tree Of Life (iTOL; v. 6.8.1) (Letunic and Bork, 2007).

---

[1] https://www.ebi.ac.uk/interpro/.

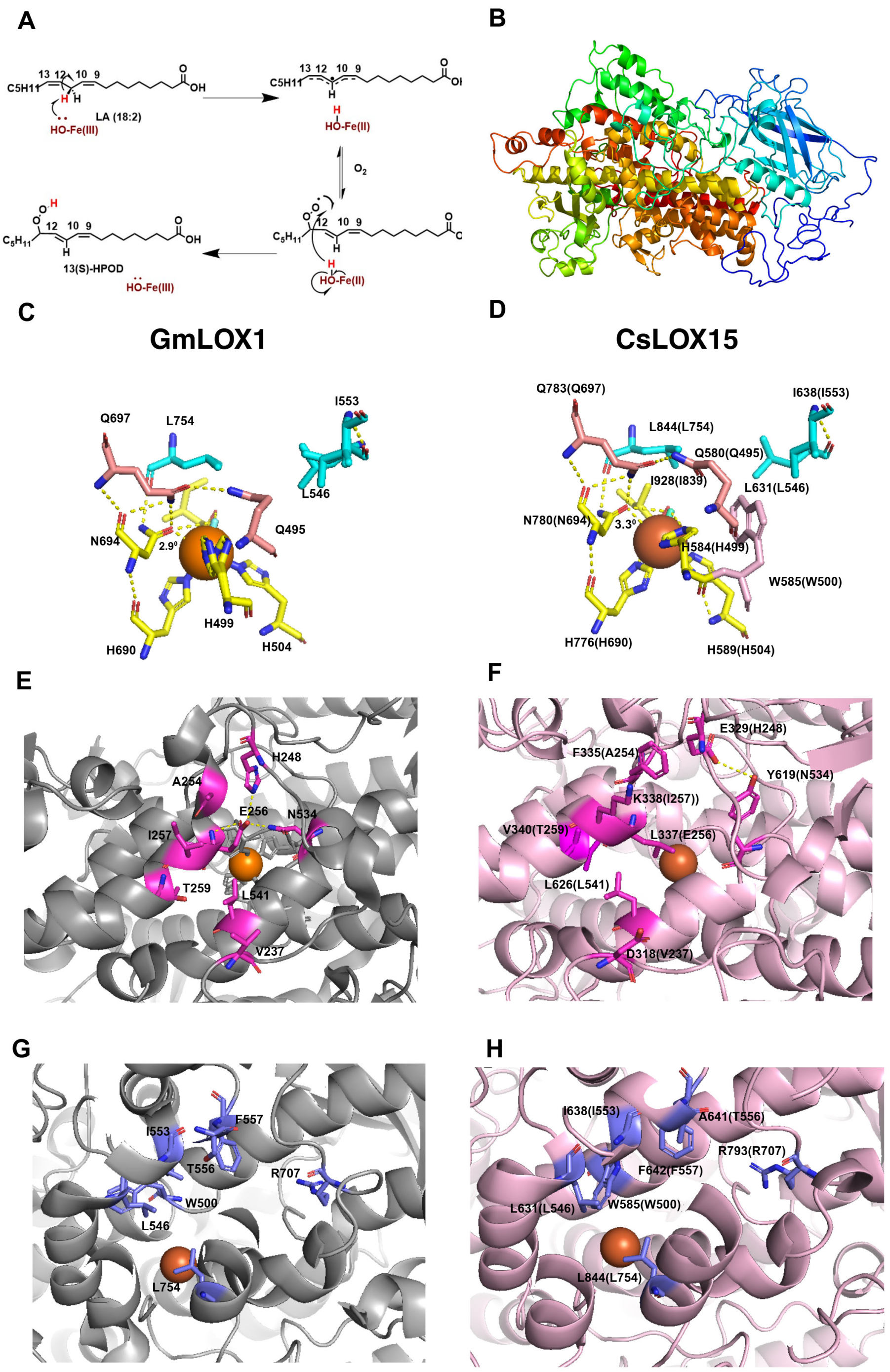


FIGURE 5
Structural features of GmLOX1 and CsLOX15. **(A)** Lipoxygenase (LOX) reaction mechanism, 13-LOX convert polyunsaturated fatty acids i.e. linoleic acid (LA) into (13*S*)-hydroperoxy-(9*Z*,11*E*)-octadecadienoic acid ((13*S*)-HPOD) by a two-step reaction involving the reduction of Fe^{3+} to Fe^{2+} by proton coupled electron transfer and then insertion of oxygen. **(B)** *In silico* 3D structure of CsLOX15 showing β- sheet rich N-terminal domain and α- helix rich C-terminal domain. **(C)** Coordination geometry of the catalytic site and Fe ligand of WT GmLOX1, showing important residues involved in iron coordination (H499, H504, H690, N694, I839, OH/H_2O, Q495, Q697, N694) and substrate binding (I553, L546, L754). **(D)** *In silico* coordination geometry of CsLOX15, around the Fe ligand (superimposed to (5 A^0)) showing iron coordination, substrate binding residues, and an additional side chain residue W585. **(E)** Amino acid residues in substrate entry of GmLOX1 facilitating H binding network. **(F)** Probable amino acid residues in CsLOX15 substrate entry. **(G)** Key amino acid residues within GmLOX1 substrate cavity. **(H)** Key amino acid residues within CsLOX15 substrate cavity. *Orange sphere*, LOX Fe ligand; *blue rod*, water molecule; *yellow dashed lines*, polar contacts. The secondary structure was predicted using Phyre2 (V 2.0) (normal mode) with 100% confidence (Kelley et al., 2015) and the validation of the model was carried out using PROCHECK program in SAVES v6.1[2] (Supplementary Figure 2). The 3D-structure was modelled using PYMOL (Ver. 3.3) by superimposing with GmLOX1 (PDB entry 1F8N).

---

[2] https://saves.mbi.ucla.edu.

photosynthates from source organs, may be partially dependent on locally stored non-structural carbohydrates to maintain carbon supply (Huang et al., 2020). Hence, trichome-specific LOXs could be providing an alternate carbon source through hydrolysis of storage lipids.

6 The cytochrome P450 enzymatic step may represent an important branch point to study the effects of oxylipins on PC production

Both HPL and AOS are homologous cytochrome P450 enzymes that do not require molecular oxygen or the reducing power of NADPH reductase (Li et al., 2008). Structural analysis of the *A. thaliana* AOS protein shows that one point mutation in the catalytic site converts AOS to HPL (Li et al., 2008). While both enzymes compete for the same substrate (e.g. hydroperoxy fatty acids), competition can be controlled through temporal regulation, with changes in the time of expression observed in HPL- or AOS-specific LOXs in *N. attenuata*, or through spatial discrimination, as described previously in *Z. mays* and tomato (Froehlich et al., 2001; Allmann et al., 2010). Silencing *HPL* in soybean, rice, and potato impaired GLV production while increasing JAs synthesis, suggesting crosstalk between these pathways (Vancanneyt et al., 2001; Tong et al., 2012; Wang et al., 2020). As *Cannabis* is predicted to contain only a single copy of *HPL* (CYP74B) and *AOS* (CYP74A) (Borrego et al., 2023), silencing *CsHPL* or *CsAOS2* would be beneficial to determine compensatory mechanisms from loss of function of either gene and subsequent impact on PC production. To elucidate sequential steps up and downstream of these CYP74 enzymes, protein-protein interactions could also be explored by chemical cross-linking coupled with mass spectrometry (MS)-cleavable tag analysis (Piersimoni et al., 2021).

7 JAs and other oxylipins in PC production

7.1 The precise mechanism(s) by which JAs mediate PC production remains unresolved

There are several mechanisms by which JAs such as MeJA may increase PC content (Table 1). For example, through glandular trichome induction, increasing inflorescence compactness by reducing internode length, and/or the spacing of repeating phytomer units that host these specialized structures (Spitzer-Rimon et al., 2019; Welling et al., 2023; Huang et al., 2024). The complexity of *Cannabis* floral architecture makes measuring JA responses challenging. Consequently, there have only been limited attempts to quantify trichome density on floral leaves (e.g., perigonal bracts/calyxes), and there are mixed reports of the effectiveness of MeJA to induce trichomes on the foliar fan leaves (Welling et al., 2023; Hahm et al., 2024; Huang et al., 2024). Promisingly, a nearly two-fold increase in trichome density of calyxes was observed in 100 μM MeJA treated plants (Hahm et al., 2024). However, trichome imaging was limited to five calyxes per plant, which may not be representative, and density measurements were calculated based on a 6 x 4 mm leaf surface area (Hahm et al., 2024). The application of deep learning models and high spatial resolution imaging technologies could be used to accurately measure changes in floral architecture following exogenous JA exposure (Huang et al., 2024; Matsumoto et al., 2024).

JAs may also directly increase PC content by inducing PC gene expression and synthesis. To date, no significant changes in the expression of PC-related genes have been reported following a MeJA treatment (Table 1). However, spatiotemporal analysis of PC and precursor pathway genes within target tissues, such as the trichome secretory cells, following JA exposure are lacking (Garrido et al., 2022; Welling et al., 2023). JA related molecular responses are often associated with JA pathway genes (*LOX, AOS, AOC, JAZ* (*JASMONATE-ZIM-DOMAIN*), *COI1* (*CORONATINE INSENSITIVE 1*) interacting with transcription factors like MYC2 and basic helix–loop–helix (bHLH)148, and a similar model is predicted for *Cannabis* (Schuurink and Tissier, 2020; Wang et al., 2020; Han et al., 2022; Song et al., 2022; Xie et al., 2023; Huang et al., 2024). The regulatory functions of JA signaling are typically initiated by the synthesis of JA-Ile, which promotes formation of the SCF^{COI1}–JAZ co-receptor complex (Hickman et al., 2017). Degradation of JAZ repressor proteins by SCF^{COI1} results in the release of JAZ-mediated transcription factors, including MYC2, which activates the expression of JA-responsive genes by targeting their promoters (Hickman et al., 2017).

In *Cannabis*, the application of exogenous MeJA has increased the expression of JA-responsive genes like *JAZ* and *COI1* and the key regulator of these genes, *CsMYC4* (Huang et al., 2024). This suggests that *CsMYC4* may have a similar function to MYC orthologs in tomato and *L. angustifolia*, both of which influence glandular trichome size and density (Xu et al., 2018; Dong et al., 2022b; Huang et al., 2024). Members of the WRKY gene family, which have roles in glandular trichome formation in *Artemisia annua*, have also been shown to be responsive to MeJA in *Cannabis* (Xie et al., 2021; Rashid et al., 2023), indicating that a much larger network of transcription factors may be contributing to JA-mediated trichome formation. To fully understand these complexes, it will be necessary to comprehensively reconstruct the gene regulatory networks that form in response to JA exposure in *Cannabis*.

7.2 Mechanisms by which oxylipins may be interacting with PC production

Despite significant progress towards understanding the influence of oxylipin metabolism on PC production (Table 1), there are multiple avenues by which these lipid molecules may be driving metabolite production (Figure 6). The following interactions are proposed: (i) JAs could induce PC production through transcription regulatory mechanism, either indirectly through JA-induced responses, or directly by inducing the biosynthesis pathways. Similarly, HPL derived oxo acids and GLVs could interact with JAs and induce JA mediated responses or be directly involved as signal molecules inducing PCs, (ii) oxylipin-induced cleavage of fatty acids may form aldehydes that act as a direct carbon source for polyketide and PC synthesis, and

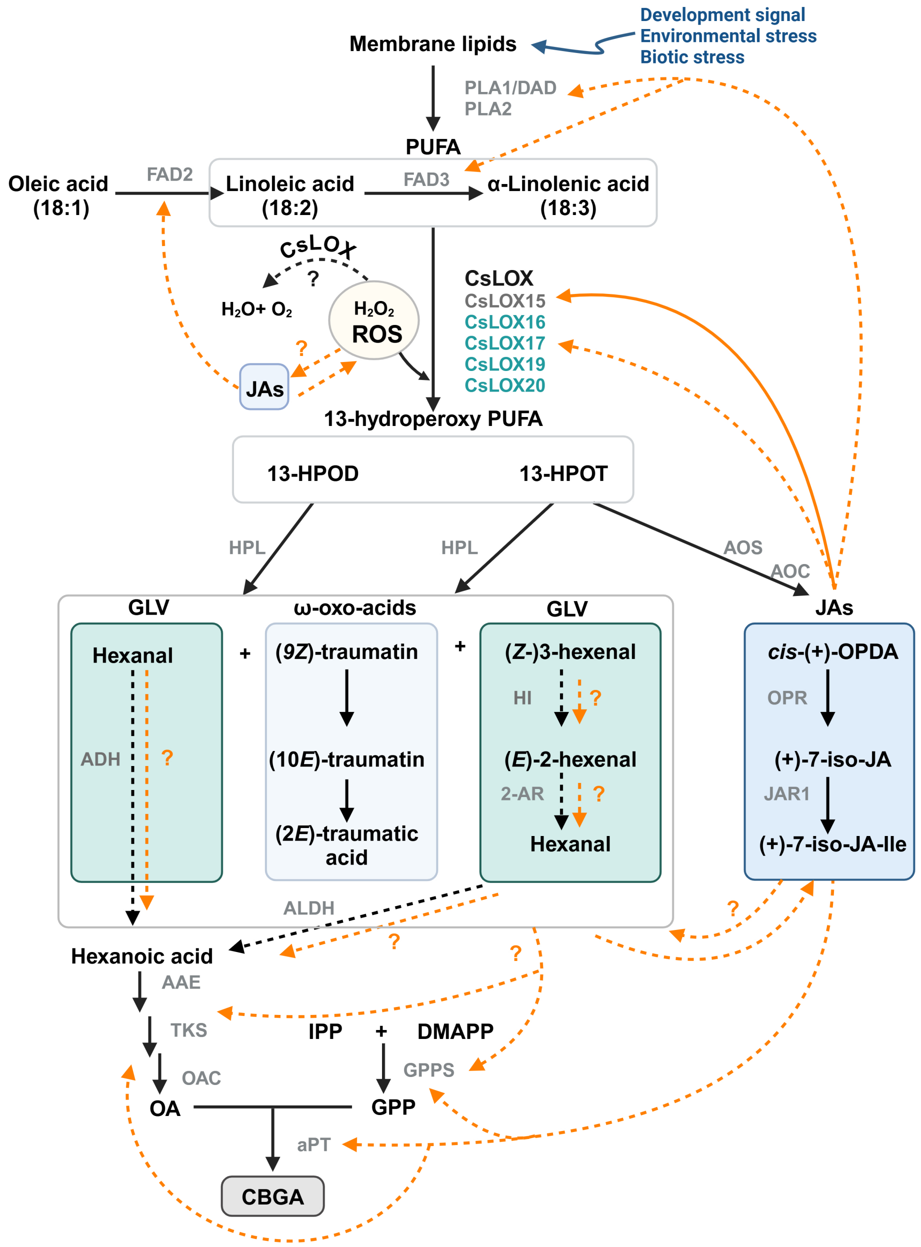


FIGURE 6

Proposed model for the interactions between oxylipin and phytocannabinoid metabolism in *Cannabis*. LOX-AOS and LOX-HPL branches of the oxylipin pathway may be interacting with phytocannabinoid biosynthesis. Potential mechanisms could include JA-mediated transcriptional regulation of phytocannabinoid biosynthesis genes or by providing a carbon source via the breakdown of PUFA. *Solid orange lines* represent known interactions; *orange dashed lines* indicate other possible interactions. Trichome-specific *Cannabis* LOX candidates for GLV and JA synthesis are shown (Table 1); *Black dashed lines* represent enzymatic reaction not yet fully explored in *Cannabis*. CsLOX15 (dark grey) has been functionally characterized, while other CsLOXs (turquoise) remain uncharacterized. Image created in BioRender.com PUFA, poly unsaturated fatty acids; 13-HPOD, (13*S*)-hydroperoxy-(9*Z*,11*E*)-octadecadienoic acid; 13-HPOT, (13*S*)-hydroperoxy-(9*Z*,11*E*,15*Z*)- octadecatrienoic acid; *cis*-(+)-OPDA, *cis*-(+)-12-oxo-phytodienoic acid; (+)-7-*iso*-JA, jasmonic acid; (+)-7-*iso*-JA-Ile, jasmonic acid isoleucine conjugate; (9*Z*)-traumatin, 12-oxo-(9*Z*)-dodecenoic acid; (10*E*)-traumatin, 12-oxo-(10*E*)-dodecenoic acid, (2*E*)-traumatic acid; (2*E*)-dodecanedioic acid; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; OA, olivetolic acid; GPP, geranyl pyrophosphate; DAD, defective in anther dehiscence protein, a phospholipase (PLA1); FAD2/3, fatty acid desaturase 2/3; CsLOX, *Cannabis* lipoxygenase; HPL, hydroperoxide lyase; AOS, allene oxide synthase; AOC, allene oxide cyclase; OPR, 12-oxo-phytodienoic acid reductase; JAR1, jasmonoyl amino acid conjugate synthase; HI, hexenal isomerase; 2-AR, 2-alkynal reductase; ALDH, aldehyde dehydrogenase; AAE, acyl activating enzyme; TKS, tetraketide synthase; OAC, olivetolic acid cyclase; GPPS, geranyl pyrophosphate synthase; aPT, aromatic prenyltransferase.

(iii) oxylipins could be involved in the detoxification of reactive oxygen species (ROS) in the microenvironment of glandular trichomes (Balcke et al., 2017), as LOXs consume H_2O_2, the biproduct of PC synthesis (Taura et al., 1996; Sirikantaramas et al., 2005). ROS could also be generated in response to JAs and in return JAs can modulate ROS homeostasis (Denness et al., 2011; Corpas et al., 2015; Demiwal et al., 2024; Zhao et al., 2025). The interplay between ROS and JAs may lead to complex signaling pathways that could activate specialized metabolism in *Cannabis*. Additionally, JA signaling is closely aligned with the regulation of other phytohormones, including ethylene and salicylic acid, all of which can affect *LOX* expression and PC production (Yamamoto et al., 2020; Li et al., 2021; Jeyasri et al., 2023; Rodrigues Magalhães et al., 2023).

8 Conclusion

Despite being a multi-use crop with huge potential as a source for food, fiber and medicine, research on *Cannabis* has been constrained for more than 50 years. There exists a substantial knowledge gap on molecular drivers responsible for the diverse array of chemicals synthesized by *Cannabis* glandular trichomes. *Cannabis* has an extensive number of LOXs, and several are highly expressed in trichomes and are responsive to JAs (Fayaz et al., 2023). We identified the trichome-specific *CsLOX15*, which together with *CsLOX16* and *CsLOX17*, forms a distinct phylogenetic clade. The use of chemical inhibitors and gene silencing approaches, such as virus-induced gene-silencing (Schachtsiek et al., 2019; Alter et al., 2022), targeting these LOXs and other key branch points of oxylipin biosynthesis will be essential in understanding the contribution of oxylipins to PC production. Structure-function analysis of candidate LOXs by site-directed mutagenesis may also prove useful in understanding sequence variation among trichome-specific LOXs (Farmer et al., 1994; Li et al., 2020).

Determining the influence of oxylipin products among different trichomes types or tissues would also be interesting avenues for future research. However, due to high reactivity and volatile nature of these compounds, challenges exist in examining these in their native form. The use of spatial omics analyses of glandular trichomes and adjacent tissues may also help to resolve the contextual molecular components of these highly specialized cell types and aid in understanding of the molecular interactions between oxylipin and PC pathways. To trace the fate of oxylipin products, it will be necessary to use isotopic tracers to label PUFA precursors. Disentangling these interactions could provide important knowledge on the partitioning of carbon resources between primary and specialized metabolism and offer new opportunities for the biotechnological enhancement of *Cannabis*. These innovations could ultimately facilitate the development of elite chemical phenotypes for both industrial and medicinal applications.

Author contributions

GS: Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. AG: Conceptualization, Funding acquisition, Supervision, Writing – review & editing. KJ: Conceptualization, Funding acquisition, Supervision, Writing – review & editing. MW: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Australian Research Council (ARC) Research Hub for Medicinal Agriculture (IH180100006). Cann Group Limited are an Industry Partner Organization of IH180100006. GIS is supported by a La Trobe University ARC Research Hub for Medicinal Agriculture Graduate Research Scholarship and a La Trobe University Full Fee Research Scholarship. Work in ARGs laboratory is also supported by the Australian Research Council Research Hub for Sustainable Crop Production (IH19100022).

Acknowledgments

Authors like to acknowledge Prof. Matthew Lewsey, Dr. Oliver Berkowitz, Dr. Ricarda Jost, and Dr. Muluneh Oli for providing RNA sequence data for the *LOX* gene tissue expression analysis.

Conflict of interest

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

Generative AI statement

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1568548/full#supplementary-material

References

Allmann, S., Halitschke, R., Schuurink, R. C., and Baldwin, I. T. (2010). Oxylipin channelling in *Nicotiana attenuata*: lipoxygenase 2 supplies substrates for green leaf volatile production. *Plant Cell Environ.* 33, 2028–2040. doi: 10.1111/j.1365-3040.2010.02203.x

Alter, H., Peer, R., Dombrovsky, A., Flaishman, M., and Spitzer-Rimon, B. (2022). Tobacco rattle vrus as a tool for rapid reverse-genetics screens and analysis of gene function in *Cannabis sativa* L. *Plants* 11, 327. doi: 10.3390/plants11030327

Ameye, M., Allmann, S., Verwaeren, J., Smagghe, G., Haesaert, G., Schuurink, R. C., et al. (2018). Green leaf volatile production by plants: a meta-analysis. *New Phytol.* 220, 666–683. doi: 10.1111/nph.14671

Anderson, L. L., Ametovski, A., Luo, J. L., Everett-Morgan, D., McGregor, I. S., Banister, S. D., et al. (2021). Cannabichromene, related phytocannabinoids, and 5-fluoro-cannabichromene have anticonvulsant properties in a mouse model of Dravet syndrome. *ACS Chem. Neurosci.* 12, 330–339. doi: 10.1021/acschemneuro.0c00677

Apicella, P. V., Sands, L. B., Ma, Y., and Berkowitz, G. A. (2022). Delineating genetic regulation of cannabinoid biosynthesis during female flower development in. *Cannabis sativa. Plant Direct* 6, e412. doi: 10.1002/pld3.412

Bailey, R. (2019). *The effect of elicitor stimulation on cannabinoid production by industrial hemp (Cannabis sativa) varieties in a hydroponic system* (University Honors College, Middle Tennessee State University). Available online at: https://jewlscholar.mtsu.edu/server/api/core/bitstreams/932cdea6-01bd-4bea-ab88-85c5d1ce4ca9/content (Accessed 25 Dec. 2024).

Balcke, G. U., Bennewitz, S., Bergau, N., Athmer, B., Henning, A., Majovsky, P., et al. (2017). Multi-omics of tomato glandular trichomes reveals distinct features of central carbon metabolism supporting high productivity of specialized metabolites. *Plant Cell* 29, 960–983. doi: 10.1105/tpc.17.00060

Behr, M., Lutts, S., Hausman, J. F., and Guerriero, G. (2018). Jasmonic acid to boost secondary growth in hemp hypocotyl. *Planta* 248, 1029–1036. doi: 10.1007/s00425-018-2951-5

Blée, E. (1998). Biosynthesis of phytooxylipins: the peroxygenase pathway. *Lipid/Fett* 100, 121–127. doi: 10.1002/(SICI)1521-4133(19985)100:4/5%3C121::AID-LIPI121%3E3.0.CO;2-4

Blée, E. (2002). Impact of phyto-oxylipins in plant defense. *Trends Plant Sci.* 7, 315–322. doi: 10.1016/S1360-1385(02)02290-2

Booth, J. K., Yuen, M. M. S., Jancsik, S., Madilao, L. L., Page, J. E., and Bohlmann, J. (2020). Terpene synthases and terpene variation in *Cannabis sativa*. *Plant Physiol.* 184, 130–147. doi: 10.1104/pp.20.00593

Borrego, E. J., Robertson, M., Taylor, J., Schultzhaus, Z., and Espinoza, E. M. (2023). Oxylipin biosynthetic gene families of *Cannabis sativa*. *PLoS One* 18, e0272893. doi: 10.1371/journal.pone.0272893

Borrelli, F., Fasolino, I., Romano, B., Capasso, R., Maiello, F., Coppola, D., et al. (2013). Beneficial effect of the non-psychotropic plant cannabinoid cannabigerol on experimental inflammatory bowel disease. *Biochem. Pharmacol.* 85, 1306–1316. doi: 10.1016/j.bcp.2013.01.017

Boughton, A. J., Hoover, K., and Felton, G. W. (2005). Methyl jasmonate application induces increased densities of glandular trichomes on tomato, *Lycopersicon esculentum*. *J. Chem. Ecol.* 31, 2211–2216. doi: 10.1007/s10886-005-6228-7

Brash, A. R. (1999). Lipoxygenases: occurrence, functions, catalysis, and acquisition of substrate. *J. Biol. Chem.* 274, 23679–23682. doi: 10.1074/jbc.274.34.23679

Britch, S. C., Babalonis, S., and Walsh, S. L. (2021). Cannabidiol: pharmacology and therapeutic targets. *Psychopharmacol.* 238, 9–28. doi: 10.1007/s00213-020-05712-8

Burgel, L., Hartung, J., Schibano, D., and Graeff-Hönninger, S. (2020). Impact of different phytohormones on morphology, yield and cannabinoid content of *Cannabis sativa* L. *Plants* 9, 725. doi: 10.3390/plants9060725

Cabrera, C. L. R., Keir-Rudman, S., Horniman, N., Clarkson, N., and Page, C. (2021). The anti-inflammatory effects of cannabidiol and cannabigerol alone, and in combination. *Pulm. Pharmacol. Ther.* 69, 102047. doi: 10.1016/j.pupt.2021.102047

Caldelari, D., Wang, G., Farmer, E. E., and Dong, X. (2011). Arabidopsis *lox3 lox4* double mutants are male sterile and defective in global proliferative arrest. *Plant Mol. Biol.* 75, 25–33. doi: 10.1007/s11103-010-9701-9

Chanda, D., Neumann, D., and Glatz, J. F. (2019). The endocannabinoid system: overview of an emerging multi-faceted therapeutic target. *Prostaglandins Leukot. Essent. Fatty Acids* 140, 51–56. doi: 10.1016/j.plefa.2018.11.016

Chauvin, A., Caldelari, D., Wolfender, J. L., and Farmer, E. E. (2013). Four 13-lipoxygenases contribute to rapid jasmonate synthesis in wounded *Arabidopsis thaliana* leaves: a role for lipoxygenase 6 in responses to long-distance wound signals. *New Phytol.* 197, 566–575. doi: 10.1111/nph.12029

Chen, G., Hackett, R., Walker, D., Taylor, A., Lin, Z., and Grierson, D. (2004). Identification of a specific isoform of tomato lipoxygenase (TomloxC) involved in the generation of fatty acid-derived flavor compounds. *Plant Physiol.* 136, 2641–2651. doi: 10.1104/pp.104.041608

Chen, G., Klinkhamer, P. G. L., Escobar-Bravo, R., and Leiss, K. A. (2018). Type VI glandular trichome density and their derived volatiles are differently induced by jasmonic acid in developing and fully developed tomato leaves: implications for thrips resistance. *Plant Sci.* 276, 87–98. doi: 10.1016/j.plantsci.2018.08.007

Chen, Z., Chen, X., Yan, H., Li, W., Li, Y., Cai, R., et al. (2015). The lipoxygenase gene family in poplar: identification, classification, and expression in response to MeJA treatment. *PLoS One* 10, e0125526. doi: 10.1371/journal.pone.0125526

Christensen, S. A., Huffaker, A., Hunter, C. T., Alborn, H. T., and Schmelz, E. A. (2016). A maize death acid, 10-oxo-11-phytoenoic acid, is the predominant cyclopentenone signal present during multiple stress and developmental conditions. *Plant Signal. Behav.* 11, e1120395. doi: 10.1080/15592324.2015.1120395

Christensen, S. A., Huffaker, A., Kaplan, F., Sims, J., Ziemann, S., Doehlemann, G., et al. (2015). Maize death acids, 9-lipoxygenase-derived cyclopente(a)nones, display activity as cytotoxic phytoalexins and transcriptional mediators. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11407–11412. doi: 10.1073/pnas.1511131112

Christensen, S. A., Nemchenko, A., Borrego, E., Murray, I., Sobhy, I. S., Bosak, L., et al. (2013). The maize lipoxygenase, *ZmLOX10*, mediates green leaf volatile, jasmonate and herbivore-induced plant volatile production for defense against insect attack. *Plant J.* 74, 59–73. doi: 10.1111/tpj.12101

Corpas, F. J., Gupta, D. K., and Palma, J. M. (2015). "Production sites of reactive oxygen species (ROS) in organelles from plant cells," in *Reactive Oxygen Species and Oxidative Damage in Plants Under Stress*. Eds. D. K. Gupta, J. M. Palma and F. J. Corpas (Springer, Cham), 1–22. doi: 10.1007/978-3-319-20421-5_1

Cropano, C., Carrera, D.Á., George, G. M., Katsir, L., Vogt, M. M., Ruckle, M. E., et al. (2022). *Quantitative trait loci (QTLs) associated with a high-varin trait in Cannabis*. WO2022180532A1. Available online at: https://patents.google.com/patent/WO2022180532A1/en?assignee=puregene&oq=puregene (Accessed November 10, 2024).

de Brito Siqueira, A. L., Cremasco, P. V., Bahú, J. O., da Silva, A. P., de Andrade, L. R. M., González, P. G., et al. (2023). Phytocannabinoids: Pharmacological effects, biomedical applications, and worldwide prospection. *J. Tradit. Complement. Med.* 13, 575–587. doi: 10.1016/j.jtcme.2023.08.006

Demiwal, P., Nabi, S. U., Mir, J. I., Verma, M. K., Yadav, S. R., Roy, P., et al. (2024). Methyl jasmonate improves resistance in scab-susceptible Red Delicious apple by altering ROS homeostasis and enhancing phenylpropanoid biosynthesis. *Plant Physiol. Biochem.* 207, 108371. doi: 10.1016/j.plaphy.2024.108371

Denness, L., McKenna, J. F., Segonzac, C., Wormit, A., Madhou, P., Bennett, M., et al. (2011). Cell wall damage-induced lignin biosynthesis is regulated by a reactive oxygen species-and jasmonic acid-dependent process in Arabidopsis. *Plant Physiol.* 156, 1364–1374. doi: 10.1104/pp.111.175737

Devinsky, O., Cilio, M. R., Cross, H., Fernandez-Ruiz, J., French, J., Hill, C., et al. (2014). Cannabidiol: pharmacology and potential therapeutic role in epilepsy and other neuropsychiatric disorders. *Epilepsia* 55, 791–802. doi: 10.1111/epi.12631

Dong, Y., Li, J., Zhang, W., Bai, H., Li, H., and Shi, L. (2022a). Exogenous application of methyl jasmonate affects the emissions of volatile compounds in lavender (*Lavandula angustifolia*). *Plant Physiol. Biochem.* 185, 25–34. doi: 10.1016/j.plaphy.2022.05.022

Dong, Y., Zhang, W., Li, J., Wang, D., Bai, H., Li, H., et al. (2022b). The transcription factor LaMYC4 from lavender regulates volatile Terpenoid biosynthesis. *BMC Plant Biol.* 22, 289. doi: 10.1186/s12870-022-03660-3

Dowling, C. A., Shi, J., Toth, J. A., Quade, M. A., Smart, L. B., McCabe, P. F., et al. (2024). A *FLOWERING LOCUS T* ortholog is associated with photoperiod-insensitive flowering in hemp (*Cannabis sativa* L.). *Plant J.* 119, 383–403. doi: 10.1111/tpj.16769

Egmond, M., Veldink, G. A., Vliegenthart, J. F. G., and Boldingh, J. (1973). C-11 H-abstraction from linoieic acid, the rate-limiting step in lipoxygenase catalysis. *Biophys. Res. Commun.* 54, 1178–1184. doi: 10.1016/0006-291X(73)90816-4

Fayaz, M., Kundan, M., Gani, U., Sharma, P., Wajid, M. A., Katoch, K., et al. (2023). Identification of Lipoxygenase gene repertoire of *Cannabis sativa* and functional characterization of *CsLOX13* gene. *Plant Sci.* 334, 111780. doi: 10.1016/j.plantsci.2023.111780

Farmer, E. E., Caldelari, D., Pearce, G., Walker-Simmons, M., and Ryan, C. A. (1994). Diethyldithiocarbamic acid inhibits the octadecanoid signaling pathway for the wound induction of proteinase inhibitors in tomato leaves. *Plant Physiol.* 106 (1), 337–342. doi: 10.1104/pp.106.1.337

Flores-Sanchez, I. J., Pec, J., Fei, J., Choi, Y. H., Dusek, J., Verpoorte, R., et al (2009). Elicitation studies in cell suspension cultures of Cannabis sativa L. *J. Biotechnol.* 143 (2), 157–168. doi: 10.1016/j.jbiotec.2009.05.006

Fortes, A. M., Coronado, M. J., Testillano, P. S., Risueño, M. D. C., and Pais, M. S. (2004). Expression of lipoxygenase during organogenic nodule formation from hop internodes. *J. Histochem. Cytochem.* 52, 227–241. doi: 10.1177/002215540405200211

Frassinetti, S., Moccia, E., Caltavuturo, L., Gabriele, M., Longo, V., Bellani, L., et al. (2018). Nutraceutical potential of hemp (*Cannabis sativa* L.) seeds and sprouts. *Food Chem.* 262, 56–66. doi: 10.1016/j.foodchem.2018.04.078

Froehlich, J. E., Itoh, A., and Howe, G. A. (2001). Tomato allene oxide synthase and fatty acid hydroperoxide lyase, two cytochrome P450s involved in oxylipin metabolism, are targeted to different membranes of chloroplast envelope. *Plant Physiol.* 125, 306–317. doi: 10.1104/pp.125.1.306

Gaffney, B. J. (2020). EPR spectroscopic studies of lipoxygenases. *Chem. Asian J.* 15, 42–50. doi: 10.1002/asia.201901461

Gagne, S. J., Stout, J. M., Liu, E., Boubakir, Z., Clark, S. M., and Page, J. E. (2012). Identification of olivetolic acid cyclase from *Cannabis sativa* reveals a unique catalytic route to plant polyketides. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12811–12816. doi: 10.1073/pnas.1200330109

Garrido, J., Rico, S., Corral, C., Sanchez, C., Vidal, N., Martinez-Quesada, J. J., et al. (2022). Exogenous application of stress-related signaling molecules affect growth and cannabinoid accumulation in medical cannabis (*Cannabis sativa* L.). *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1082554

Ghasemi Pirbalouti, A., Sajjadi, S. E., and Parang, K. (2014). A review (research and patents) on jasmonic acid and its derivatives. *Arch. Pharm. (Weinheim).* 347, 229–239. doi: 10.1002/ardp.201300287

Ghorbel, M., Brini, F., Sharma, A., and Landi, M. (2021). Role of jasmonic acid in plants: the molecular point of view. *Plant Cell Rep.* 40, 1471–1494. doi: 10.1007/s00299-021-02687-4

Goldsmith, C. R., Jonas, R. T., and Stack, T. D. P. (2002). C– H bond activation by a ferric methoxide complex: modeling the rate-determining step in the mechanism of lipoxygenase. *J. Am. Chem. Soc* 124, 83–96. doi: 10.1021/ja016451g

Gorina, S. S., Mukhtarova, L. S., Iljina, T. M., Toporkova, Y. Y., and Grechkin, A. N. (2022). Detection of divinyl ether synthase CYP74H2 biosynthesizing (11*Z*)-etheroleic and (1′*Z*)-colnelenic acids in asparagus (*Asparagus officinalis* L.). *Phytochemistry* 200, 113212. doi: 10.1016/j.phytochem.2022.113212

Grechkin, A. N., Ilyasov, A. V., and Hamberg, M. (1997). On the mechanism of biosynthesis of divinyl ether oxylipins by enzyme from garlic bulbs. *Eur. J. Biochem.* 245, 137–142. doi: 10.1111/j.1432-1033.1997.00137.x

Hahm, S., Lee, Y., Lee, K., and Park, J. (2024). Optimization of cannabinoid production in hemp through methyl jasmonate application in a vertical farming system. *Horticulturae* 10, 1165. doi: 10.3390/horticulturae10111165

Hamberg, M. (1999). An epoxy alcohol synthase pathway in higher plants: biosynthesis of antifungal trihydroxy oxylipins in leaves of potato. *Lipids* 34, 1131–1142. doi: 10.1007/s11745-999-0464-7

Hamberg, M. (2004). Isolation and structures of two divinyl ether fatty acids from *Clematis vitalba*. *Lipids* 39, 565–569. doi: 10.1007/s11745-004-1264-9

Hammond, C. T., and Mahlberg, P. G. (1973). Morphology of glandular hairs of *Cannabis sativa* from scanning electron microscopy. *Am. J. Bot.* 60, 524–528. doi: 10.1002/j.1537-2197.1973.tb05953.x

Han, G., Li, Y., Yang, Z., Wang, C., Zhang, Y., and Wang, B. (2022). Molecular mechanisms of plant trichome development. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.910228

Hazekamp, A., Tejkalová, K., and Papadimitriou, S. (2016). Cannabis: from cultivar to chemovar II—a metabolomics approach to cannabis classification. *Cannabis Cannabinoid Res.* 1, 202–215. doi: 10.1089/can.2016.0017

He, Y., Borrego, E. J., Gorman, Z., Huang, P. C., and Kolomiets, M. V. (2020). Relative contribution of LOX10, green leaf volatiles and JA to wound-induced local and systemic oxylipin and hormone signature in *Zea mays* (maize). *Phytochemistry* 174, 112334. doi: 10.1016/j.phytochem.2020.112334

Hershelman, D., Kahler, K. M., Price, M. J., Lu, I., Fu, Y., Plumeri, P. A., et al. (2019). Oxygenation reactions catalyzed by the F557V mutant of soybean lipoxygenase-1: evidence for two orientations of substrate binding. *Arch. Biochem. Biophys.* 674, 108082. doi: 10.1016/j.abb.2019.108082

Hesami, M., Pepe, M., Baiton, A., and Jones, A. M. P. (2022a). Current status and future prospects in cannabinoid production through *in vitro* culture and synthetic biology. *Biotechnol. Adv.* 108074. doi: 10.1016/j.bioteChadv.2022.108074

Hesami, M., Pepe, M., Baiton, A., Salami, S. A., and Jones, A. M. P. (2022b). New insight into ornamental applications of cannabis: Perspectives and challenges. *Plants* 11 2383. doi: 10.3390/plants11182383

Hickman, R., Van Verk, M. C., Van Dijken, A. J., Mendes, M. P., Vroegop-Vos, I. A., Caarls, L., et al. (2017). Architecture and dynamics of the jasmonic acid gene regulatory network. *Plant Cell.* 29, 2086–2105. doi: 10.1105/tpc.16.00958

Holman, T. R., Zhou, J., and Solomon, E. I. (1998). Spectroscopic and functional characterization of a ligand coordination mutant of soybean lipoxygenase-1: first coordination sphere analogue of human 15-lipoxygenase. *J. Am. Chem. Soc.* 120, 12564–12572. doi: 10.1021/ja982844c

Hornung, E., Walther, M., Kühn, H., and Feussner, I. (1999). Conversion of cucumber linoleate 13-lipoxygenase to a 9-lipoxygenating species by site-directed mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4192–4197. doi: 10.1073/pnas.96.7.4192

Horper, W., and Marner, F. J. (1996). Biosynthesis of primin and miconidin and its derivatives. *Phytochemistry* 41, 451–456. doi: 10.1016/0031-9422(95)00590-0

Howe, G. A., Lee, G. I., Itoh, A., Li, L., and DeRocher, A. E. (2000). Cytochrome P450-dependent metabolism of oxylipins in tomato. Cloning and expression of allene oxide synthase and fatty acid hydroperoxide lyase. *Plant Physiol.* 123, 711–724. doi: 10.1104/pp.123.2.711

Hu, S., Offenbacher, A. R., Thompson, E. M., Gee, C. L., Wilcoxen, J., Carr, C. A., et al. (2019). Biophysical characterization of a disabled double mutant of soybean lipoxygenase: the "undoing" of precise substrate positioning relative to metal cofactor and an identified dynamical network. *J. Am. Chem. Soc.* 141, 1555–1567. doi: 10.1021/jacs.8b10992

Huang, X., Chen, W., Zhao, Y., Chen, J., Ouyang, Y., Li, M., et al. (2024). Deep learning-based quantification and transcriptomic profiling reveal a methyl jasmonate-mediated glandular trichome formation pathway in *Cannabis sativa*. *Plant J.* 118, 1155–1173. doi: 10.1111/tpj.16663

Huang, J., Rücker, A., Schmidt, A., Gleixner, G., Gershenzon, J., Trumbore, S., et al. (2020). Production of constitutive and induced secondary metabolites is coordinated with growth and storage in Norway spruce saplings. *Tree Physiol.* 40, 928–942. doi: 10.1093/treephys/tpaa040

Hughes, R., Lawson, D. M., Hornostaj, A. R., Fairhurst, S. A., and Casey, R. (2001). Mutagenesis and modelling of linoleate-binding to pea seed lipoxygenase. *Eur. J. Biochem.* 268, 1030–1040. doi: 10.1046/j.1432-1327.2001.01964.x

Hurgobin, B., Tamiru-Oli, M., Welling, M. T., Doblin, M. S., Bacic, A., Whelan, J., et al. (2021). Recent advances in *Cannabis sativa* genomics research. *New Phytol.* 230, 73–89. doi: 10.1111/nph.17140

Jeyasri, R., Muthuramalingam, P., Karthick, K., Shin, H., Choi, S. H., and Ramesh, M. (2023). Methyl jasmonate and salicylic acid as powerful elicitors for enhancing the production of secondary metabolites in medicinal plants: an updated review. *Plant Cell Tissue Organ Cult.* 153, 447–458. doi: 10.1007/s11240-023-02485-8

Jin, D., Henry, P., Shan, J., and Chen, J. (2021). Identification of chemotypic markers in three chemotype categories of cannabis using secondary metabolites profiled in inflorescences, leaves, stem bark, and roots. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.699530

Jost, R., Berkowitz, O., Pegg, A., Hurgobin, B., Tamiru-Oli, M., Welling, M. T., et al. (2025). Sink strength, nutrient allocation, cannabinoid yield, and associated transcript profiles vary in two drug-type *Cannabis* chemovars. *J. Exp. Bot.* 76, 152–174. doi: 10.1093/jxb/erae367

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Kianersi, F., Azarm, D. A., Pour-Aboughadareh, A., and Poczai, P. (2022). Change in secondary metabolites and expression pattern of key rosmarinic acid related genes in Iranian lemon balm (*Melissa officinalis* L.) ecotypes using methyl jasmonate treatments. *Molecules* 27, 1715. doi: 10.3390/molecules27051715

Knapp, M. J., and Klinman, J. P. (2003). Kinetic studies of oxygen reactivity in soybean lipoxygenase-1. *Biochemistry* 42, 11466–11475. doi: 10.1021/bi0300884

Kolomiets, M. V., Hannapel, D. J., Chen, H., Tymeson, M., and Gladon, R. J. (2001). Lipoxygenase is involved in the control of potato tuber development. *Plant Cell* 13, 613–626. doi: 10.1105/tpc.13.3.613

Kovalchuk, I., Pellino, M., Rigault, P., Van Velzen, R., Ebersbach, J., Ashnest, J. R., et al. (2020). The genomics of Cannabis and its close relatives. *Annu. Rev. Plant Biol.* 71, 713–739. doi: 10.1146/annurev-arplant-081519-040203

Kroumova, A. B., Xie, Z., and Wagner, G. J. (1994). A pathway for the biosynthesis of straight and branched, odd-and even-length, medium-chain fatty acids in plants. *Proc. Natl. Acad. Sci. U.S.A.* 91, 11437–11441. doi: 10.1073/pnas.91.24.11437

Letunic, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127,128. doi: 10.1093/bioinformatics/btl529

Li, L., Chang, Z., Pan, Z., Fu, Z. Q., and Wang, X. (2008). Modes of heme binding and substrate access for cytochrome P450 CYP74A revealed by crystal structures of allene oxide synthase. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13883–13888. doi: 10.1073/pnas.0804099105

Li, P., Soudackov, A. V., and Hammes-Schiffer, S. (2018). Impact of mutations on the binding pocket of soybean lipoxygenase: implications for proton-coupled electron transfer. *J. Phys. Chem. Lett.* 9, 6444–6449. doi: 10.1021/acs.jpclett.8b02945

Li, J., Wang, X., Jiang, R., Dong, B., Fang, S., Li, Q., et al. (2021). Phytohormone-based regulation of trichome development. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.734776

Li, Y., Qiu, L., Zhang, Q., Zhuansun, X., Li, H., Chen, X., et al. (2020). Exogenous sodium diethyldithiocarbamate, a Jasmonic acid biosynthesis inhibitor, induced resistance to powdery mildew in wheat. *Plant Direct.* 4 (4), e00212. doi: 10.1002/pld3.212

Liavonchanka, A., and Feussner, I. (2006). Lipoxygenases: occurrence, functions and catalysis. *J. Plant Physiol.* 163, 348–357. doi: 10.1016/j.jplph.2005.11.006

Lichtman, A. H., Lux, E. A., McQuade, R., Rossetti, S., Sanchez, R., Sun, W., et al. (2018). Results of a double-blind, randomized, placebo-controlled study of nabiximols oromucosal spray as an adjunctive therapy in advanced cancer patients with chronic uncontrolled pain. *J. Pain Symptom Manage.* 55, 179–188.e171. doi: 10.1016/j.jpainsymman.2017.09.001

Liu, S., and Han, B. (2010). Differential expression pattern of an acidic 9/13-lipoxygenase in flower opening and senescence and in leaf response to phloem feeders in the tea plant. *BMC Plant Biol.* 10, 1–15. doi: 10.1186/1471-2229-10-228

Liu, J., Zhou, Y., Li, J., Wang, F., and Yang, Y. (2020). Comprehensive genomic characterization and expression analysis of the lipoxygenase gene family in watermelon under hormonal treatments. *Agriculture* 10, 429. doi: 10.3390/agriculture10100429

Livingston, S. J., Bae, E. J., Unda, F., Hahn, M. G., Mansfield, S. D., Page, J. E., et al. (2021). Cannabis glandular trichome cell walls undergo remodeling to store specialized metabolites. *Plant Cell Physiol.* 62, 1944–1962. doi: 10.1093/pcp/pcab127

Livingston, S. J., Quilichini, T. D., Booth, J. K., Wong, D. C., Rensing, K. H., Laflamme-Yonkman, J., et al. (2020). Cannabis glandular trichomes alter morphology and metabolite content during flower maturation. *Plant J.* 101, 37–56. doi: 10.1111/tpj.14516

Livingston, S. J., Rensing, K. H., Page, J. E., and Samuels, A. L. (2022). A polarized supercell produces specialized metabolites in cannabis trichomes. *Curr. Biol.* 32, 4040–4047.e4044. doi: 10.1016/j.cub.2022.07.014

Luo, X., Reiter, M. A., d'Espaux, L., Wong, J., Denby, C. M., Lechner, A., et al. (2019). Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* 567, 123–126. doi: 10.1038/s41586-019-0978-9

Marks, M. D., Tian, L., Wenger, J. P., Omburo, S. N., Soto-Fuentes, W., He, J., et al. (2009). Identification of candidate genes affecting Δ9-tetrahydrocannabinol biosynthesis in *Cannabis sativa*. *J. Exp. Bot.* 60, 3715–3726. doi: 10.1093/jxb/erp210

Matsumoto, S., Utsumi, Y., Kozuka, T., Iwamura, M., Nakai, T., Yamauchi, D., et al. (2024). CT image-based 3D inflorescence estimation of *Chrysanthemum seticuspe*. *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1374937

McGarry, R. C., Lin, Y.-T., Kaur, H., Higgs, H., Arias-Gaguancela, O., and Ayre, B. G. (2024). Disrupted oxylipin biosynthesis mitigates pathogen infections and pest infestations in cotton (*Gossypium hirsutum*). *J. Exp. Bot.* 75, 7365–7380. doi: 10.1093/jxb/erae394

Mechoulam, R., and Parker, L. A. (2013). The endocannabinoid system and the brain. *Annu. Rev. Psychol.* 64, 21–47. doi: 10.1146/annurev-psych-113011-143739

Meyer, M. P., Tomchick, D. R., and Klinman, J. P. (2008). Enzyme structure and dynamics affect hydrogen tunneling: The impact of a remote side chain (I553) in soybean lipoxygenase-1. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1146–1151. doi: 10.1073/pnas.0710643105

Minor, W., Steczko, J., Stec, B., Otwinowski, Z., Bolin, J. T., Walter, R., et al. (1996). Crystal structure of soybean lipoxygenase L-1 at 1.4 Å resolution. *Biochemistry* 35, 10687–10701. doi: 10.1021/bi960576u

Mochizuki, S., Sugimoto, K., Koeduka, T., and Matsui, K. (2016). Arabidopsis lipoxygenase 2 is essential for formation of green leaf volatiles and five-carbon volatiles. *FEBS Lett.* 590, 1017–1027. doi: 10.1002/1873-3468.12133

Naimuzzaman, M., Hasan, M. M., Kumer, A., Hossin, A. Y., Harun-Ur-Rashid, M., Roy, S. K., et al. (2025). Computational and *In silico* study of novel fungicides against combating root rot, gray mold, fusarium wilt, and cereal rust. *PloS One* 20, e0316606. doi: 10.1371/journal.pone.0316606

Nakashima, A., von Reuss, S. H., Tasaka, H., Nomura, M., Mochizuki, S., Iijima, Y., et al. (2013). Traumatin-and dinortraumatin-containing galactolipids in *Arabidopsis*: their formation in tissue-disrupted leaves as counterparts of green leaf volatiles. *J. Biol. Chem.* 288, 26078–26088. doi: 10.1074/jbc.M113.487959

Newie, J., Andreou, A., Neumann, P., Einsle, O., Feussner, I., and Ficner, R. (2016). Crystal structure of a lipoxygenase from Cyanothece sp. may reveal novel features for substrate acquisition. *J. Lipid Res.* 57, 276–287. Available online at: https://www.jlr.org/article/S0022-2275(20)35463-8/fulltext.

Offenbacher, A. R., Hu, S., Poss, E. M., Carr, C. A., Scouras, A. D., Prigozhin, D. M., et al. (2017). Hydrogen–deuterium exchange of lipoxygenase uncovers a relationship between distal, solvent exposed protein motions and the thermal activation barrier for catalytic proton-coupled electron tunneling. *ACS Cent. Sci.* 3, 570–579. doi: 10.1021/acscentsci.7b00142

Omare, M. O., Kibet, J. K., Cherutoi, J. K., and Kengara, F. O. (2021). Current trends in the use of *Cannabis sativa*: beyond recreational and medicinal applications. *Open Access Libr. J.* 8, 1–15. doi: 10.4236/oalib.1107132

Oultram, J. M. J., Pegler, J. L., Eamens, A. L., Gordon, R., Korbie, D. J., and Grof, C. P. L. (2024). Exogenously applied gibberellic acid alters cannabinoid profile in *Cannabis sativa* L. *Agronomy* 14, 2417. doi: 10.3390/agronomy14102417

Pattanaik, S., Patra, B., Singh, S. K., and Yuan, L. (2014). An overview of the gene regulatory network controlling trichome development in the model plant, *Arabidopsis*. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00259

Piersimoni, L., Kastritis, P. L., Arlt, C., and Sinz, A. (2021). Cross-linking mass spectrometry for investigating protein conformations and protein–protein interactions─ a method for all seasons. *Chem. Rev.* 122, 7500–7531. doi: 10.1021/acs.chemrev.1c00786

Rashid, A., Khajuria, M., Ali, V., Faiz, S., Jamwal, S., and Vyas, D. (2023). Genome-wide identification and expression profiling of WRKY family suggest their potential role in cannabinoid regulation in *Cannabis sativa* L. *Ind. Crops Prod.* 206, 117706. doi: 10.1016/j.indcrop.2023.117706

Raza, A., Charagh, S., Zahid, Z., Mubarik, M. S., Javed, R., Siddiqui, M. H., et al. (2021). Jasmonic acid: a key frontier in conferring abiotic stress tolerance in plants. *Plant Cell Rep.* 40, 1513–1541. doi: 10.1007/s00299-020-02614-z

Rodrigues Magalhães, H. C., Alves Filho, E. G., Rivero Meza, S. L., Oliveira, A., Garruti, D. S., and Purgatto, E. (2023). Effect of methyl jasmonate on the biosynthesis of volatile compounds associated with the ripening of grape tomato fruits. *J. Agric. Food Chem.* 71, 4696–4705. doi: 10.1021/acs.jafc.2c06215

Ruddat, V. C., Mogul, R., Chorny, I., Chen, C., Perrin, N., Whitman, S., et al. (2004). Tryptophan 500 and arginine 707 define product and substrate active site binding in soybean lipoxygenase-1. *Biochemistry* 43, 13063–13071. doi: 10.1021/bi0489098

Salamone, S., Waltl, L., Pompignan, A., Grassi, G., Chianese, G., Koeberle, A., et al. (2022). Phytochemical characterization of *Cannabis sativa* L. chemotype V reveals three new dihydrophenanthrenoids that favorably reprogram lipid mediator biosynthesis in macrophages. *Plants* 11, 2130. doi: 10.3390/plants11162130

Sanadhya, P., Kumar, A., Bucki, P., Fitoussi, N., Carmeli-Weissberg, M., Borenstein, M., et al. (2021). Tomato divinyl ether-biosynthesis pathway is implicated in modulating of root-knot nematode *Meloidogyne javanica*'s parasitic ability. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.670772

Sands, L. B., Haiden, S. R., Ma, Y., and Berkowitz, G. A. (2023). Hormonal control of promoter activities of *Cannabis sativa* prenyltransferase 1 and 4 and salicylic acid mediated regulation of cannabinoid biosynthesis. *Sci. Rep.* 13, 8620. doi: 10.1038/s41598-023-35303-4

Scala, A., Allmann, S., Mirabella, R., Haring, M. A., and Schuurink, R. C. (2013). Green leaf volatiles: a plant's multifunctional weapon against herbivores and pathogens. *Int. J. Mol. Sci.* 14, 17781–17811. doi: 10.3390/ijms140917781

Schachtsiek, J., Hussain, T., Azzouhri, K., Kayser, O., and Stehle, F. (2019). Virus-induced gene silencing (VIGS) in *Cannabis sativa* L. *Plant Methods* 15, 157. doi: 10.1186/s13007-019-0542-5

Schenk, G., Neidig, M. L., Zhou, J., Holman, T. R., and Solomon, E. I. (2003). Spectroscopic characterization of soybean lipoxygenase-1 mutants: the role of second coordination sphere residues in the regulation of enzyme activity. *Biochemistry* 42, 7294–7302. doi: 10.1021/bi027380g

Schlotzhauer, W. S., Pair, S. D., and Horvat, R. J. (1996). Volatile constituents from the flowers of Japanese honeysuckle (*Lonicera japonica*). *J. Agric. Food Chem.* 44, 206–209. doi: 10.1021/jf950275b

Schuurink, R., and Tissier, A. (2020). Glandular trichomes: micro-organs with model status? *New Phytol.* 225, 2251–2266. doi: 10.1111/nph.16283

Segraves, E. N., Chruszcz, M., Neidig, M. L., Ruddat, V., Zhou, J., Wecksler, A. T., et al. (2006). Kinetic, spectroscopic, and structural investigations of the soybean lipoxygenase-1 first-coordination sphere mutant, Asn694Gly. *Biochemistry* 45, 10233–10242. doi: 10.1021/bi060577e

Shakil, S. S. M., Gowan, M., Hughes, K., Azam, M. N. K., and Ahmed, M. N. (2021). A narrative review of the ethnomedicinal usage of *Cannabis sativa* Linnaeus as traditional phytomedicine by folk medicine practitioners of Bangladesh. *J. Cannabis Res.* 3, 1–12. doi: 10.1186/s42238-021-00063-3

Shen, J., Tieman, D., Jones, J. B., Taylor, M. G., Schmelz, E., Huffaker, A., et al. (2014). A 13-lipoxygenase, TomloxC, is essential for synthesis of C5 flavour volatiles in tomato. *J. Exp. Bot.* 65, 419–428. doi: 10.1093/jxb/ert382

Sievers, F., and Higgins, D. G. (2021). "The Clustal Omega Multiple Alignment Package." in *Multiple Sequence Alignment*, ed. K Katoh (New York, NY: Humana). *Methods Mol. Biol.* 2231. doi: 10.1007/978-1-0716-1036-7_1

Sirikantaramas, S., Morimoto, S., Shoyama, Y., Ishikawa, Y., Wada, Y., Shoyama, Y., et al. (2004). The gene controlling marijuana psychoactivity: molecular cloning and heterologous expression of Δ^1-tetrahydrocannabinolic acid synthase from *Cannabis sativa* L. *J. Biol. Chem.* 279, 39767–39774. doi: 10.1074/jbc.M403693200

Sirikantaramas, S., and Taura, F. (2017). "Cannabinoids: biosynthesis and biotechnological applications," in Cannabis sativa *L. -* Botany and Biotechnology. Eds. S. Chandra, H. Lata and M. A. ElSohly (Springer Nature, Cham), 183–206. doi: 10.1007/978-3-319-54564-6_8

Sirikantaramas, S., Taura, F., Tanaka, Y., Ishikawa, Y., Morimoto, S., and Shoyama, Y. (2005). Tetrahydrocannabinolic acid synthase, the enzyme controlling marijuana psychoactivity, is secreted into the storage cavity of the glandular trichomes. *Plant Cell Physiol.* 46, 1578–1582. doi: 10.1093/pcp/pci166

Skrzypczak-Jankun, E., Bross, R. A., Carroll, R. T., Dunham, W. R., and Funk, M. O. (2001). Three-dimensional structure of a purple lipoxygenase. *J. Am. Chem. Soc* 123, 10814–10820. doi: 10.1021/ja011759t

Small, E. (2015). Evolution and classification of *Cannabis sativa* (marijuana, hemp) in relation to human utilization. *Bot. Rev.* 81, 189–294. doi: 10.1007/s12229-015-9157-3

Song, W. C., and Brash, A. R. (1991). Purification of an allene oxide synthase and identification of the enzyme as a cytochrome P-450. *Science* 253, 781–784. doi: 10.1126/science.1876834

Song, C., Cao, Y., Dai, J., Li, G., Manzoor, M. A., Chen, C., et al. (2022). The multifaceted roles of MYC2 in plants: toward transcriptional reprogramming and stress tolerance by jasmonate signaling. *Front.Plant Sci.* 13. doi: 10.3389/fpls.2022.868874

Song, H., Wang, P., Li, C., Han, S., Lopez-Baltazar, J., Zhang, X., et al. (2016). Identification of lipoxygenase (*LOX*) genes from legumes and their responses in wild type and cultivated peanut upon *Aspergillus flavus* infection. *Sci. Rep.* 6, 35245. doi: 10.1038/srep35245

Spitzer-Rimon, B., Duchin, S., Bernstein, N., and Kamenetsky, R. (2019). Architecture and florogenesis in female *Cannabis sativa* plants. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00350

Steczko, J., Donoho, G. P., Clemens, J. C., Dixon, J. E., and Axelrod, B. (1992). Conserved histidine residues in soybean lipoxygenase: functional consequences of their replacement. *Biochemistry* 31, 4053–4057. doi: 10.1021/bi00131a022

Steel, L., Welling, M., Ristevski, N., Johnson, K., and Gendall, A. (2023). Comparative genomics of flowering behavior in *Cannabis sativa*. *Front. Plant Sci.* 14, 1227898. doi: 10.3389/fpls.2023.1227898

Stout, J. M., Boubakir, Z., Ambrose, S. J., Purves, R. W., and Page, J. E. (2012). The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in *Cannabis sativa* trichomes. *Plant J.* 71, 353–365. doi: 10.1111/j.1365-313X.2012.04949.x

Tagen, M., and Klumpers, L. E. (2022). Review of delta-8-tetrahydrocannabinol (Δ^8-THC): Comparative pharmacology with Δ^9-THC. *Br. J. Pharmacol.* 179, 3915–3933. doi: 10.1111/bph.15865

Tanney, C. A., Backer, R., Geitmann, A., and Smith, D. L. (2021). Cannabis glandular trichomes: A cellular metabolite factory. *Front. Plant Sci.* 12, 1923. doi: 10.3389/fpls.2021.721986

Taura, F., Morimoto, S., and Shoyama, Y. (1996). Purification and characterization of cannabidiolic-acid synthase from *Cannabis sativa* L.: Biochemical analysis of a novel enzyme that catalyzes the oxidocyclization of cannabigerolic acid to cannabidiolic acid. *J. Biol. Chem.* 271, 17411–17416. doi: 10.1074/jbc.271.29.17411

Thiele, E. A., Marsh, E. D., French, J. A., Mazurkiewicz-Beldzinska, M., Benbadis, S. R., Joshi, C., et al. (2018). Cannabidiol in patients with seizures associated with Lennox-Gastaut syndrome (GWPCARE4): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* 391, 1085–1096. doi: 10.1016/S0140-6736(18)30136-3

Tomchick, D. R., Phan, P., Cymborowski, M., Minor, W., and Holman, T. R. (2001). Structural and functional characterization of second-coordination sphere mutants of soybean lipoxygenase-1. *Biochemistry* 40, 7509–7517. doi: 10.1021/bi002893d

Tong, X., Qi, J., Zhu, X., Mao, B., Zeng, L., Wang, B., et al. (2012). The rice hydroperoxide lyase OsHPL3 functions in defense responses by modulating the oxylipin pathway. *Plant J.* 71, 763–775. doi: 10.1111/j.1365-313x.2012.05027.x

Toporkova, Y. Y., Smirnova, E. O., and Gorina, S. S. (2024). Epoxyalcohol synthase branch of lipoxygenase cascade. *Curr. Issues Mol. Biol.* 46, 821–841. doi: 10.3390/cimb46010053

Toth, J. A., Stack, G. M., Carlson, C. H., and Smart, L. B. (2022). Identification and mapping of major-effect flowering time loci *Autoflower1* and *Early1* in *Cannabis sativa* L. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.991680

Urits, I., Borchart, M., Hasegawa, M., Kochanski, J., Orhurhu, V., and Viswanath, O. (2019). An update of current cannabis-based pharmaceuticals in pain medicine. *Pain Ther.* 8, 41–51. doi: 10.1007/s40122-019-0114-4

Vancanneyt, G., Sanz, C., Farmaki, T., Paneque, M., Ortego, F., Castañera, P., et al. (2001). Hydroperoxide lyase depletion in transgenic potato plants leads to an increase in aphid performance. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8139–8144. doi: 10.1073/pnas.141079498

Vogt, J., Schiller, D., Ulrich, D., Schwab, W., and Dunemann, F. (2013). Identification of lipoxygenase (LOX) genes putatively involved in fruit flavour formation in apple (*Malus domestica*). *Tree Genet. Genomes* 9, 1493–1511. doi: 10.1007/s11295-013-0653-5

Wang, Y., Liu, M., Ge, D., Akhter-Bhat, J., Li, Y., Kong, J., et al. (2020). Hydroperoxide lyase modulates defense response and confers lesion-mimic leaf phenotype in soybean (*Glycine max* (L.) Merr.). *Plant J.* 104, 1315–1333. doi: 10.1111/tpj.15002

Wang, J., Song, L., Gong, X., Xu, J., and Li, M. (2020). Functions of jasmonic acid in plant regulation and response to abiotic stress. *Int. J. Mol. Sci.* 21, 1446. doi: 10.3390/ijms21041446

Wang, Y., Wang, D., Li, S., Zhai, Y., Zhao, Y., Zhong, F., et al. (2024). Methyl jasmonate and salicylic acid enhance the total flavonoid, phenolics, and cannabidiol contents of *Cannabis sativa* L. adventitious roots. *Plant Cell Tissue Organ Cult.* 159, 1–14. doi: 10.1007/s11240-024-02909-z

Wang, X., Zeng, Y., Li, L., Fan, R., Yan, Y., Guo, L., et al. (2024). Unraveling jasmonate-mediated plant immunity against green peach aphid (*Myzus persicae*) in peach: an integrated transcriptomic and RNA sequencing approach. *Plant Growth Regul.* 104, 1245–1259. doi: 10.1007/s10725-024-01198-y

Wasternack, C., and Feussner, I. (2018). The oxylipin pathways: biochemistry and function. *Annu. Rev. Plant Biol.* 69, 363–386. doi: 10.1146/annurev-arplant-042817-040440

Wasternack, C., and Hause, B. (2013). Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. *Ann. Bot.* 111, 1021–1058. doi: 10.1093/aob/mct067

Weber, H., Chételat, A., Caldelari, D., and Farmer, E. E. (1999). Divinyl ether fatty acid synthesis in late blight–diseased potato leaves. *Plant Cell* 11, 485–493. doi: 10.1105/tpc.11.3.485

Welling, M. T., Deseo, M. A., O'Brien, M., Clifton, J., Bacic, A., and Doblin, M. S. (2023). Metabolomic analysis of methyl jasmonate treatment on phytocannabinoid production in *Cannabis sativa*. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1110144

Welling, M. T., Liu, L., Kretzschmar, T., Mauleon, R., Ansari, O., and King, G. J. (2020). An extreme-phenotype genome-wide association study identifies candidate cannabinoid pathway genes in *Cannabis*. *Sci. Rep.* 10, 18643. doi: 10.1038/s41598-020-75271-7

Welling, M. T., Shapter, T., Rose, T. J., Liu, L., Stanger, R., and King, G. J. (2016). A belated green revolution for *Cannabis*: virtual genetic resources to fast-track cultivar development. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01113

Wu, Y., Trejo, H. X., Chen, G., and Li, S. (2021). Phytoremediation of contaminants of emerging concern from soil with industrial hemp (*Cannabis sativa* L.): a review. *Environ. Dev. Sustain.* 23, 114405–14435. doi: 10.1007/s10668-021-01289-0

Xie, Z., Mi, Y., Kong, L., Gao, M., Chen, S., Chen, W., et al. (2023). *Cannabis sativa*: origin and history, glandular trichome development, and cannabinoid biosynthesis. *Horticult. Res.* 10, uhad150. doi: 10.1093/hr/uhad150

Xie, L., Yan, T., Li, L., Chen, M., Ma, Y., Hao, X., et al. (2021). The WRKY transcription factor AaGSW2 promotes glandular trichome initiation in *Artemisia annua*. *J. Exp. Bot.* 72, 1691–1701. doi: 10.1093/jxb/eraa523

Xu, J., van Herwijnen, Z. O., Dräger, D. B., Sui, C., Haring, M. A., and Schuurink, R. C. (2018). SlMYC1 regulates type VI glandular trichome formation and terpene biosynthesis in tomato glandular cells. *Plant Cell* 30, 2988–3005. doi: 10.1105/tpc.18.00571

Yamamoto, F., Iwanaga, F., Al-Busaidi, A., and Yamanaka, N. (2020). Roles of ethylene, jasmonic acid, and salicylic acid and their interactions in frankincense resin production in *Boswellia sacra* Flueck. trees. *Sci. Rep.* 10, 16760. doi: 10.1038/s41598-020-73993-2

Yan, X., Cui, L., Liu, X., Cui, Y., Wang, Z., Zhang, H., et al. (2022). *NbJAZ3* is required for jasmonate-meditated glandular trichome development in *Nicotiana benthamiana*. *Physiol. Plant*. 174, e13666. doi: 10.1111/ppl.13666

Yokoyama, M., Kurusu, T., Ohno, H., Ifuku, O., Harada, R., and Tada, Y. (2025). Oxylipin KODA enhances the early growth of rice (*Oryza sativa* L.) under low-temperature stress at night to simulate a natural temperature condition. *Plant Biotechnol.* 42, 73–77. doi: 10.5511/plantbiotechnology.24.1218a

Youn, B., Sellhorn, G. E., Mirchel, R. J., Gaffney, B. J., Grimes, H. D., and Kang, C. (2006). Crystal structures of vegetative soybean lipoxygenase VLX-B and VLX-D, and comparisons with seed isoforms LOX-1 and LOX-3. *Proteins* 65, 1008–1020. doi: 10.1002/prot.21182

Zhao, Y., Wu, Y., Zhang, X., Zhu, X., Hou, Y., Chen, J., et al. (2025). Methyl jasmonate attenuates chilling injury of prune fruit by maintaining ROS homeostasis and regulating GABA metabolism and energy status. *Postharvest Biol. Technol.* 220, 113303. doi: 10.1016/j.postharvbio.2024.113303

Zoeller, M., Stingl, N., Krischke, M., Fekete, A., Waller, F., Berger, S., et al. (2012). Lipid profiling of the Arabidopsis hypersensitive response reveals specific lipid peroxidation and fragmentation processes: biogenesis of pimelic and azelaic acid. *Plant Physiol.* 160, 365–378. doi: 10.1104/pp.112.202846

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 16 May 2025
DOI 10.3389/fpls.2025.1558321

OPEN ACCESS

EDITED BY
Surendra Pratap Singh,
Chhatrapati Shahu Ji Maharaj University, India

REVIEWED BY
Miguel Alfonso,
Spanish National Research Council (CSIC), Spain
Shouchuang Wang,
Hainan University, China

*CORRESPONDENCE
Guoying Zhou
✉ zhougy@nwipb.cas.cn

†These authors have contributed equally to this work

RECEIVED 10 January 2025
ACCEPTED 14 April 2025
PUBLISHED 16 May 2025

CITATION
Wang T, Zhao S, Wang B, Li J, Ye Z, Zhang F, Ma H and Zhou G (2025) Genome-wide analysis of MYB transcription factors in four *Rheum* L. plants provides new insights into the synthesis of Anthraquinones. *Front. Plant Sci.* 16:1558321.
doi: 10.3389/fpls.2025.1558321

Genome-wide analysis of MYB transcription factors in four *Rheum* L. plants provides new insights into the synthesis of Anthraquinones

Tao Wang[1,2†], Shuo Zhao[1,2†], Bo Wang[1,2], Jianan Li[1,2], Zengrong Ye[1,3], Famei Zhang[1,2], Huiyuan Ma[1,4] and Guoying Zhou[1,2*]

[1]Key Laboratory of Tibetan Medicine Research, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, Qinghai, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]College of Agriculture and Animal Husbandry, Qinghai University, Xining, China, [4]College of Life Sciences, Qinghai Normal University, Xining, China

R. tanguticum (*Rheum tanguticum* Maxim. ex Regel) is a herbaceous plant belonging to Polygonaceae family and *Rheum* L. genus. It holds considerable value in culinary and medicinal realms, primarily due to their rich Anthraquinones (AQs) content. Understanding the molecular mechanisms that regulate AQs biosynthesis is a prerequisite for increasing their yield. MYB transcription factors (TFs) can regulate the synthesis of a variety of plant secondary metabolites. However, only a few research have explored the role of MYB TFs in *Rheum* L. species. In this study, 1054 MYB genes from four *Rheum* L. species were identified. The number of MYB genes in each species was similar, distributed across 11 chromosomes. To investigate the phylogeny of identified MYB TFs, they were classified into four subfamilies. Sequence characteristics, phylogenetic relationships, evolutionary trends, and tissue expression of MYB genes in *Rheum* L. species were further studied. Subsequently, 12 MYB genes were selected, which shown differential expression in different tissues. Further research on these genes indicated a significant correlation with genes in shikimate pathway and polyketide pathway of AQs biosynthesis. Protein-protein interaction simulations in *Arabidopsis thaliana* and qRT-PCR experiments further confirmed this situation. This research lays the foundation for studying molecular mechanisms by which MYB TFs regulates AQs biosynthesis in four *Rheum* L. species.

KEYWORDS

MYBs, anthraquinones, genome-wide analysis, RNA-Seq, qPCR, *Rheum* L. plants

Highlights

- A total of 1054 MYB genes were identified in four *Rheum* L. species.
- MYBs reflect the genetic diversity of the three genuine rhubarbs (*R. palmatum*, *R. officinale* and *R. tanguticum*).
- Tandem duplication makes a significant contribution to the expansion of the MYBs in *R. tanguticum*.
- Some MYB genes play a key role in AQs biosynthesis.

1 Introduction

Rheum L. plants is a tall perennial herb in Polygonaceous family, with about 60 species worldwide, mainly in temperate and subtropical alpine regions of Asia (Wang et al., 2005). Certain species among them possess exceptionally high medicinal and nutritional values, thereby securing a significant position in global applications. *R. palmatum*, *R. officinale* and *R. tanguticum* are *Rheum* L. species included in the Chinese Pharmacopoeia and of significant medicinal value, commonly referred to as "Rhubarbs" (C.P.C., 2020). *R. nobile*, which is similar in habitat and closely related phylogenetically, is utilized only in a restricted geographical area (Tibet). Research on *R. nobile* is currently relatively limited (Table 1). Rhubarbs contain a variety of important secondary metabolites, including over 200 types of secondary metabolites (Xiang et al., 2020). Among these active components, AQs such as aloe-emodin, rhein, chrysophanol, etc., exhibit significant pharmacological activity and possess effects such as antibacterial, anti-inflammatory, and anticancer properties (Wang et al., 2021).

AQs and their derivatives are aromatic polyketone compounds that play important roles in plants, such as photoprotection and enhancing plant disease resistance. The complete biosynthetic pathway of anthraquinones in plants remains unclear. However, two incompletely resolved anthraquinone biosynthetic pathways have been tentatively proposed in plants: the polyketide pathway and the shikimate/o-succinylbenzoate pathway. Due to shared intermediates, the MEP (methylerythritol 4-phosphate) and MVA (mevalonic acid) pathways are also involved in the biosynthesis of AQs (Leistner, 1985; Leistner and Zenk, 1968). Therefore, identifying key functional genes involved in the synthesis pathway of AQs compounds is a crucial step in current research to increase the production of AQs in rhubarbs. Recently, some key structural genes involved in AQs biosynthesis have been discovered in some plants, such as polyketide synthase-like enzyme (PKSIII) in *Cassia tora*, and UDP-glycosyltransferase (UGT) in *Fagopyrum esculentum*, etc (Kang et al., 2020; Zhao H. et al., 2023). However, the research identification of transcription factors related to AQs biosynthesis is still almost a blank slate.

Almost all eukaryotic organisms contain the MYB transcription factors (TFs), which are renowned for its ability to regulate a wide range of physiological processes in plants (Martin and Paz-Ares, 1997). These processes include environmental adaptation, hormone signal transduction, development, metabolic regulation, and so on (Allan et al., 2008). The first MYB TF cloned in plants was *ZmMYBC1* in maize (Paz-Ares et al., 1987). MYB TFs is one of the most prominent families of transcription factors in plants, consisting of three conserved functional domains: DNA-binding domain (DBD), transcription activation domain (TAD), and the not yet fully defined negative regulatory domain (NRD). Among them, DBD is the most conserved and commonly referred to as MYB domain (Frampton, 2004). MYB domain is composed of 1-4 repeating units of about 52 amino acids each, forming three alpha-helices, among which the second and third helices form a helix-turn-helix (HTH) structure (Ogata et al., 1996).

Based on adjacent repetitive units, MYB TFs can be divided into 4 subfamilies: 1R-MYB, R2R3-MYB, 3R-MYB, and 4R-MYB (Ogata et al., 1996). In studies identifying and analyzing MYB TFs present in plants, R2R3-MYB TFs are usually the most abundant, followed by 1R-MYB and 3R-MYB TFs. 4R-MYB TFs are less common, and 5R-MYB as well as newly discovered MYB TFs are even rarer (Jin and Martin, 1999; Yanhui et al., 2006; Thiedig et al., 2021; Si et al., 2023). There is a significant variation in the number of MYB genes across different species, ranging from less than 100 to over 500. For instance, 54 MYB genes have been identified and analyzed in mango (Zhang et al., 2022). In contrast, there are 524 in *Gossypium hirsutum* L (Salih et al., 2016). Most studies focus on a single plant species. To date, there is still limited information available on comparative analysis of MYB genes among closely related species.

The available information is scarce in comparison to the significance of MYB TFs and AQs in *Rheum* L. species, particularly for the three types of rhubarbs and their important related species, *R. nobile*. In several influential databases, no research related to MYB TFs of these four *Rheum* L. plants has been found. This study conducted a comprehensive genomic comparative analysis of MYB TFs in four species of *Rheum* L. plants. 263, 267, 259 and 265 MYB genes in *R. tanguticum*, *R. officinale*, *R. palmatum* and *R. nobile* have been achieved respectively. Subsequently, a systematic study was conducted through comparative genomic, HPLC and RNA sequencing (RNA-seq) analyses to explore the phylogenetic differences of MYB TFs among four rhubarbs, focusing on their tissue-specific expression patterns and protein-protein interaction relationships in *R. tanguticum*. The response of MYB genes to multiple confirmed structural genes involved in AQs synthesis provides a basis for further understanding the role of these genes in AQs biosynthesis. These results reflect a comprehensive molecular evolutionary analysis of MYB TFs and a detailed understanding of specific MYB proteins involved in AQs biosynthesis.

2 Materials and methods

2.1 Data resources

The whole genome sequence and annotation of *R. tanguticum* were obtained from self-sequencing data, while the data for *R.*

Abbreviations: *R. tanguticum, Rheum tanguticum* Maxim.ex Balf; *R. palmatum, Rheum palmatum; R. offcinale, Rheum offcinale* Baill; *R. nobile, Rheum nobile* Hook. f. & Thomson.

TABLE 1 AQs content in four *Rheum* L. species: A comparative review.

Ethnopharmacological classification	Species	Evaluation of *AQs* Content	References
Authentic rhubarbs	*R. tanguticum* *R. palmatum* *R. offcinale*	As most important three species in *Rheum* genus, *R. palmatum* l., *R. tanguticum* Maxim., and *R. officinale* Baill. have been widely used for medicinal (roots).	Su et al., 2020 Tan et al., 2023 Zhuang et al., 2020
Regional rhubarb	*R. nobile*	Besides Authentic rhubarbs, some plants in this genus produced in Hebei provinces et al. also were used as a rhubarb`s substitute by local people.	Yang et al., 2024

palmatum, *R. officinale*, and *R. nobile* were sourced from published genomic articles (Feng et al., 2023; Zhang H. et al., 2024; Zhang T. et al., 2024). MYB protein sequences of *Arabidopsis thaliana* were obtained from the TAIR database (https://www.arabidopsis.org/). MYBs Hidden Markov Model (HMM) profile files were downloaded from Pfam database (http://pfam.xfam.org/), including: Myb_DNA-binding (PF00249), Myb_DNA-bind_2 (PF08914), Myb_DNA-bind_3 (PF12776), Myb_DNA-bind_4 (PF13837), Myb_DNA-bind_5 (PF13873), Myb_DNA-bind_6 (PF13921), and Myb_DNA-bind_7 (PF15963).

2.2 Identification of MYB genes in four rhubarbs

To identify MYB genes in four *Rheum* L. plants (*R. tanguticum*, *R. palmatum*, *R. officinale*, and *R. nobile*), these strategies were employed. HMM profile for the MYB domain (Pfam accession numbers: PF00249, PF08914, PF12776, PF13837, PF13873, PF13921, and PF15963) was used to identify potential MYB genes in the genomes of four *Rheum* L. species using HMMER 3.4 software (http://hmmer.janelia.org/), with an E-value threshold of 1e-2 (Cui et al., 2024). Only proteins predicted to contain at least one MYB domain by online tool SMART (https://smart.embl.de/) were identified as candidate MYB members in rhubarbs. Subsequently, BLAST analyses (E-value of 1e-10) were conducted using the MYB candidate genes from the four *Rheum* L. plants as queries against the rhubarbs genome to verify the integrity of the identified MYBs from the database of rhubarbs (Altschul et al., 1990; Katiyar et al., 2012; Yao et al., 2024). Sequences obtained from these strategies were merged and redundant sequences were removed. Finally, CDD database (version 3.21) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) were used to further analyze and validate these sequences. Proteins confirmed by this database were considered candidate MYB TFs.

2.3 Physicochemical property analysis of four *Rheum* L. plants MYB proteins

Physicochemical properties of MYB TFs, such as molecular weight and isoelectric point, were analyzed using ExPASy (http://www.expasy.ch/tools/pi_tool.html). Protein Parameter Calc in Tbtools (version 2.119) was used to validate and complement these data (Chen et al., 2023).

2.4 Construction of phylogenetic tree of rhubarbs MYB TFs

Using the identified 263, 259, 267, 265, 197 MYB TFs from *R. tanguticum*, *R. palmatum*, *R. offcinale*, *R. nobile* and *Arabidopsis thaliana*, two phylogenetic trees were constructed with *R. tanguticum* as the core. Multiple sequence alignments of the conserved MYB domains were conducted using the ClustalW software implemented in MEGA 11 (version 11.0.13) (Kumar et al., 2016). Subsequently, a neighbor-joining (NJ) phylogenetic tree based on the MYB domain alignment was constructed by MEGA 11 according to JTT model (Kumar et al., 2016). Relative branch support was evaluated by 1000 bootstrap replicates, branch lengths were calculated by pairwise comparisons of genetic distances, and missing data were treated by pairwise deletions of gaps. Additionally, the ChiPlot (https://www.chiplot.online/) and iTOL tree (https://itol.embl.de/) websites were used for tree file modification and visualization.

2.5 Sequence analysis of MYB TFs in four *Rheum* L. species

Online MEME Suite (https://meme-suite.org/meme/) was used to perform motif analysis on MYB TFs (Bailey et al., 2009). Detected maximum number of motifs was set to 10, with the site distribution set to any, and all other parameters default. The motif identification images were saved. Batch CD-Search tool on the NCBI website (https://www.ncbi.nlm.nih.gov), was utilized to detect conserved domains in each plant species, with the E-value threshold set to 1e-10 and all other parameters default. Additionally, we prepared phylogenetic analysis files and genomic gff3 files for the four *Rheum* L. species.

Based on results from MEME, CD-search, phylogenetic analysis, and the genomic gff3 files, the identified MYB genes from four *Rheum* L. species were submitted to TBtools for protein motif composition and gene structure analysis, with the results visualized in graphical form (Chen et al., 2023).

2.6 Analysis of *cis*-acting elements of four *Rheum* L. MYB TFs

2000 bp *cis*-acting elements in the MYB TFs were obtained from the online website Plant CARE (http://bioinformatics.psb.ugent.be/

webtools/plantcare/), and redundancies were removed after screening. Chiplot (https://www.chiplot.online/) was used for heatmap visualization.

2.7 Analyses of chromosome distribution, duplication events of MYB TFs in four *Rheum* L. plants

Chromosomal distribution information of MYB TFs in four *Rheum* L. plants was derived from the annotated genomic data in their genome database. Gene Location Visualize (Advanced) tool of TBtools was used to visualize the chromosomal localization of MYBs in rhubarbs, and MapChart 2.3.2 was used to validate the results (https://www.wur.nl/). For the syntenic analysis of MYB genes in the four *Rheum* L. species, MCScanX with default settings was used to identify gene pairs of segmental duplications within their genomes (Wang et al., 2012). Tandem duplications were identified as two MYB genes separated by no more than one intervening gene. Ka and Ks of evolution were calculated using the Simple Ka/Ks Calculator (Chen et al., 2023). The divergence times (T) were calculated as T = Ks/(2λ) × 10^-6 Mya, with the approximate value for the clock-like rate λ = 2.44×10^-9 for rhubarbs (Li et al., 2023).

2.8 Expression profiling of *R. tanguticum* MYB TFs in various tissues

To determine expression profiles of MYB TFs, we obtained the transcriptomic data of four types of tissues—root, stem, leaf, and seed—from previous studies conducted in our laboratory. Quality control of the raw reads was performed using FastQC (v0.12.0). After trimming Illumina adapter sequences and removing low-quality bases with the FASTX-Toolkit, clean reads were aligned against the reference genome using HISAT2 (v2.2.1) with default parameter settings (Kim et al., 2015; Liu C. et al., 2020). The aligned reads were then converted into Bam files and sorted using SAMtools (v1.9.1). Subsequently, resulting aligned reads were processed by StringTie to assemble reads into genes and measure the expression levels of each gene (Pertea et al., 2015; Liu Y. et al., 2020). The obtained RNA-seq data were standardized using the “limma” package in R software. Subsequently, differentially expressed genes (DEGs) were assessed using the “DESeq2” package. Genes with an adjusted p-value <0.05 and an absolute log2 fold change (FC) >2 across four different tissues were considered as DEGs, with gene expression levels calculated as the average of three biological replicates. After escaping, number of fragments per kilobase of transcript per million mapped reads (FPKM) was used to visualize the distribution of gene expression levels (https://www.bioinformatics.com.cn) (Tang et al., 2023).

Previous studies have identified several genes encoding enzymes involved in the biosynthesis of AQs from the polyketide pathway, including 24 *PKSIII* genes. Candidate genes generated from the initial analysis and structural genes confirmed to be involved in AQs synthesis were characterised by Pearson’s index using OriginPro (2025) to screen for *RtanMYB* TFs that may regulate AQs biosynthesis.

2.9 Analysis of the AQs content

The determination of AQs was carried out using a high-performance liquid chromatography (HPLC) system (Agilent 1260 Infinity II). Reference standards for five AQs were obtained from Sigma (USA). Reagents used included methanol, acetonitrile, phosphoric acid, and formic acid (HPLC grade; Shandong Yuwang Group, CHN). A total of 5 AQs were measured, namely, aloe emodin (110795–201007), rhein (110757–200206), emodin (110756–200110), chrysophanol (110758–201013), and physcion (110796–201118).

A conical flask containing 25 mL of methanol was used to weigh in 0.5 grams of powder, and the total weight was meticulously recorded. After 1 hour of heating and refluxing, the sample was cooled to room temperature and reweighed. Any loss of weight during this process was compensated for by the addition of an equivalent volume of methanol. The solution was then filtered, and the resulting filtrate was collected to serve as the final sample solution for the determination of AQs. Chromatographic separation was achieved using a Unitary C18 column (4.6 × 250 mm, 5 μm, 100 Å). The mobile phase consisted of two components: solvent A (0.1% phosphoric acid in water) and solvent B (methanol), with elution performed in ultrapure water. The gradient elution program was as follows: from 0 to 40 min, 85% solvent B and 15% solvent A. Quantification of AQs was performed using the external standard method, with detection at a wavelength of 254 nm and an injection volume of 10 μL. The methodological approach was based on a previous study on rhubarb (Zhao S. et al., 2024) to ensure the accuracy and reliability of the experimental methods.

2.10 Subcellular localization

WoLF PSORT (https://wolfpsort.hgc.jp/) was employed to predict the subcellular localization of MYB TFs in four *Rheum* L. plants. DEGs of *R. tangutian* were comprehensively analyzed using WoLF PSORT, TBtool, and Cell-PLoc to determine their locations in subcellular structures.

2.11 Protein-protein interaction network of MYB TFs involved in the biosynthesis of AQs in *R. tanguticum*

STRING database with default parameters (https://string-db.org/) was used to predict, execute and visualise the potential protein-protein interaction networks of *RtanMYB* TFs based on known homologs.

2.12 Expression verification of candidate MYB TFs in *R. tanguticum*

DEGs related to AQs biosynthesis pathway were verificated by qRT-PCR. Primers were designed using PRIMER 5.0, and primers were validated for specificity using NCBI Primer BLAST, with *R. tanguticum* 18S rRNA as an internal reference (Supplementary Table S9). Sample RNA extraction of the four fractions (roots, stems, leaves and seeds) was first performed followed by reverse transcription operations.

Using TB Green Premix Ex Taq II (with TliRNase H Plus) (Takara Biomedical Technology Co., Ltd., Beijing, China) and an ABI 7500 real-time PCR system (Life Technologies, Foster City, CA, USA), transcript levels of these genes were quantified by quantitative real-time PCR (qRT-PCR). qRT-PCR procedures were as follows: denaturation at 95°C for 30 s, followed by 40 cycles of 95°C for 5 s and 60°C for 34 s; and unwinding at 95°C for 5 s, 60°C for 1 min and 95°C for 5 s. The results were obtained by using the ABI 7500 Real-Time PCR System (Life Technologies, Foster City, CA, USA). Data were analysed using the 2^ (-ΔΔCT) method (Schmittgen and Livak, 2008). In this study, each gene was tested using three identical samples, with each set representing an independent biological replicate. Additionally, to ensure the accuracy of the data, we performed three technical replicates for each sample.

3 Result

3.1 Identification of the MYB genes in the four *Rheum* L. species

1054 MYB TFs were identified from four *Rheum* L. species, including 263, 259, 267, 265 MYB TFs from *R. tanguticum*, *R. palmatum*, *R. offcinale*, *R. nobile* (Supplementary Table S1). Based on their relative linear order on chromosomes, they were renamed. Detailed information about these genes could be found in Supplementary Tables S1–S3. Identified MYB TFs were classified into 4 types: 1R-MYB, R2R3-MYB, 3R-MYB, 4R and Atypical-MYB (Supplementary Table S2). Among these types, R2R3-MYB was most abundant in four *Rheum* L. species. In *R. tanguticum*, *R. palmatum*, and *R. nobile*, the proportions of 3R-MYB and 4R-MYB were relatively small; while in *R. offcinale*, the proportion of 1R-MYB was relatively small. It was worth noting that, unlike general plant species, the number of MYB TFs in the 3R-MYB, 4R & Atypical-MYB groups had expanded to a greater extent in four rhubarbs, especially in *R. tanguticum*; the number of 3R-MYB and 4R & Atypical-MYB TFs (39.55%) was very close to that of R2R3-MYB (45.25%) (Supplementary Table S2).

In addition, since four *Rheum* L. species were widespread in the wild and cultivated as economically important crops. Moreover, MYBs played an important role in stress resistance, nutrient acquisition, and metabolic synthesis. Therefore, this study also conducted an internal comparative analysis of 1054 MYB TFs across the four rhubarbs. Results indicated that there was no linear correlation between the number of MYB genes and the size of the genome in each species (Supplementary Table S2). Furthermore, the number of MYB TFs varied among them (Supplementary Table S2), with significant interspecific divergence among subfamily members, except for the 3R-MYB. The proportion of 1R-MYB in *R. officinale* was relatively low (11.61%) compared to that in *R. nobile* (19.25%). In *R. tanguticum*, the proportion of R2R3-MYB was relatively low (45.25%), but the proportion of atypical MYBs was relatively high (22.43%). Studies had shown that a smaller number of atypical genes (ATGs) was associated with smaller differences between species, suggesting that large number of ATGs tended to undergo extreme forms of birth and death evolution (Nei and Rooney, 2005). MYB family was an important TF family, and its rapid expansion might have enhanced the adaptability of plants to environmental stresses and promoted establishment of survival signal transduction networks, physiology, and metabolic pathways under stress. This pattern was common in the evolution of gene families and was mainly caused by species-specific gene duplication (Panchy et al., 2016).

3.2 *Cis*-acting target sequence analysis of genes that encode enzymes involved in AQs biosynthesis reveal the potential role of MYB TFs

Many studies had confirmed the regulatory role of MYB TFs in mediating the biosynthesis of various secondary metabolites. However, there were no reports on the involvement of MYB in AQs biosynthesis. Gene expression regulation was largely mediated by *cis*-regulatory elements (CREs), which played a critical role in modulating gene functions across various biological processes (Lyu et al., 2024). To ascertain whether MYB TFs were involved in the regulation of AQs biosynthesis in rhubarbs, we analyzed the 2-kb upstream promoter regions of 24 type III polyketide synthases (PKSIII) that have been confirmed to participate in AQs biosynthesis in *R. tanguticum*.

45 CREs related to various biological activities were detected, including plant growth and development (Supplementary Figure S1). Notably, binding sites for MYB TFs are widely present among these 24 *RtPKSIII* genes. This finding confirmed the potential of MYB TFs to play a regulatory role in AQs biosynthesis.

3.3 Phylogenetic analysis of MYB TFs in four *Rheum* L. species

In plants, PKSIII genes such as chalcone synthases (CHSs) were involved in the biosynthesis of plant specialized metabolites, especially in the biosynthesis of acetate-pathway-derived flavonoids, stilbenes, and aromatic polyphenols (Lyu et al., 2024; Millard et al., 2019; Kang et al., 2020). CHSs in the flavonoid

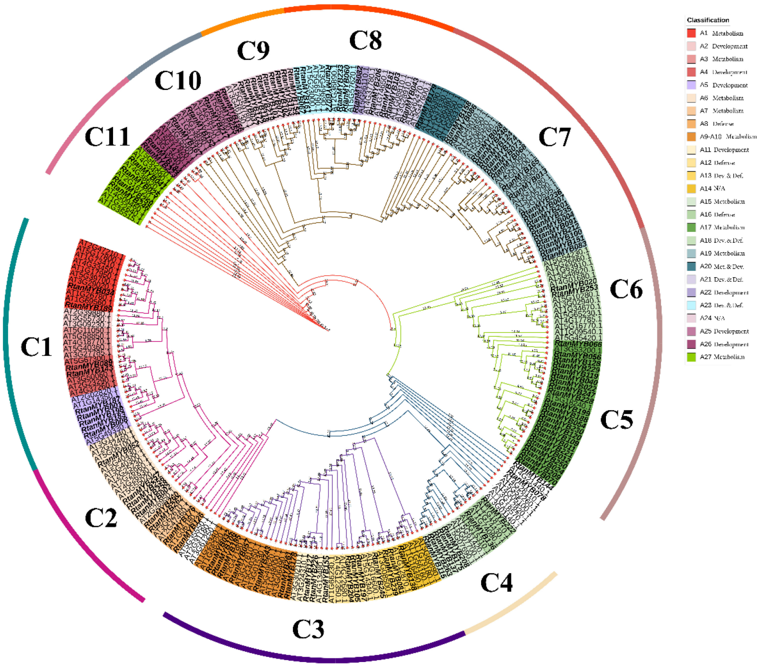


FIGURE 1
The phylogenetic tree of R2R3-MYB TFs from *R. tanguticum* and *Arabidopsis thaliana*. These proteins were divided into 11 clusters based on their topological structure. Additionally, based on the functional classification of *Arabidopsis thaliana* R2R3-MYBs, 27 subgroups were distinguished, each assigned a number (A1-A27).

pathway had been reported to be regulated by R2R3-MYB TFs in many studies (Golovko, 2023; Tuan et al., 2015; Zhu et al., 2023), which provided us with a reliable idea for screening MYB TFs that regulated AQs biosynthesis. Constructing phylogenetic trees based on model plants could offer a promising approach for uncovering gene functions. Therefore, to gain a clearer understanding of the functions of MYB TFs in AQs biosynthesis, phylogenetic trees were constructed using R2R3-MYB TFs from *R. tanguticum* (Figure 1).

The phylogenetic relationship of R2R3-MYB TFs in *R. tanguticum* and *Arabidopsis thaliana* was analyzed using MEGA12 (Figure 2). These genes were divided into 11 clades (C1-C11) according to the topology of the phylogenetic tree. Based on the TAIR (https://www.arabidopsis.org/) and prior research on *Arabidopsis thaliana* R2R3-MYB TFs, the A1 to A27 subfamilies and potential functions of their genes were indicated in Figure 2 (Dubos et al., 2010). For instance, the *RtanMYB* TFs in C7 were categorized into groups A19 and A20. Genes in these subfamilies might regulate anthocyanin metabolism and have certain defensive functions. The clustering on the phylogenetic tree of *Arabidopsis thaliana* - based subgroups into larger topological structures might imply convergent evolution of these genes in *R. tanguticum* (C1, C2). Similarly, some genes that originally belonged to the same subgroup lost their original classification trend in topological differentiation, likely due to gene functional divergence (C3 and unshaded branches on the tree).

CHSs shared between the AQs polyketide pathway and acetate-pathway-derived flavonoids is the breakthrough for mining MYB TFs regulating AQs biosynthesis (Kang et al., 2020; Tang et al., 2024). We screened some subgroups (A1, A3, A6, A7, A9, A10, A14, A17, A18, A19, A20, A27) involved in the regulation of secondary metabolite synthesis (especially the flavonoid pathway) and properly expanded them based on their topological cluster categories (C1, C2, C3, C5, C6, C7, C11), identifying 67 MYB TFs with potential functions. Analysis of the sequences of 67 MYBs shown they generally have typical R2 and R3 domains (Figure 3). Moreover, some sequences contain specific conserved motif signatures (bHLH, ANDV/ENVI et al.) for flavonoid pathway metabolites (Table 2). Previous studies had demonstrated that certain MYB TFs, which modulated flavonoid biosynthesis, could either activate or inhibit CHSs. These genes not only possessed the typical R2-R3 domain but also contained signature conserved domains such as ENVI, TLLLFR, or SG7. Further screening and analysis based on these conserved domains would lay an important groundwork for identifying MYB TFs that regulated CHSs in the AQs biosynthesis (polyketide pathway).

Subsequently, we performed a phylogenetic analysis of MYB TFs across all four species based on the subfamily classification of MYB TFs (Supplementary Figures S3–S6). The same method was used to divide different clusters on the phylogenetic tree. Using Arabidopsis thaliana as an internal reference, the contracting and expansion of gene numbers in different subgroups revealed that MYB TFs have undergone different evolutions after the differentiation of species in the *Rheum* genus. Notably, the number of MYB TFs in stress - related subgroups is similar in *R. tanguticum* and *R. nobile*, which may be related to their similar habitats (Chen et al., 2024).

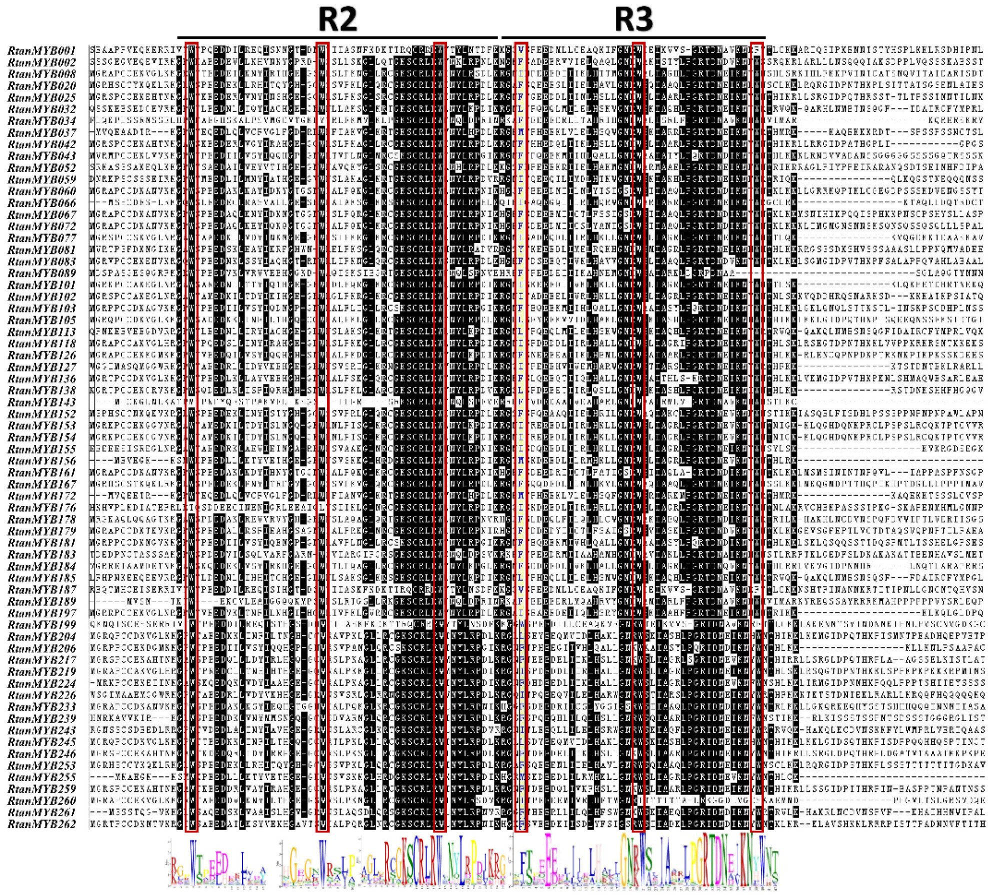


FIGURE 2
Amino acid sequence analysis of R2R3-MYB TFs in *R. tanguticum* that may be involved in AQs biosynthesis. The conserved sequences of key features are highlighted with a black background, and the conserved "W" in the R2 and R3 domains is marked with a red box.

3.4 Motif, domain and structure analysis of MYBs in four *Rheum* L. species

More systematic evolutionary relationships among four *Rheum* L. species are revealed through comprehensive sequence analysis. To understand the conserved structural domains of MYB proteins in the four *Rheum*. L species, we conducted Motif Elicitation (MEME) analysis on each identified MYB TFs. Ten different conserved motifs were discovered in the MYB TFs (Supplementary Figures S7–S10). Among these motifs, we detected many conserved amino acids, particularly tryptophan residue (W), which was a notable feature of the MYB domain (Wu et al., 2022). It could be observed that closely related genes have similar types and numbers of motifs, but each gene has a different domain structure (Figure 3; Supplementary Figures S11–S13). Different motifs contained different types of conserved amino acids. This was of great significance for differentiation of protein functions.

RtanMYB TFs containing motif 1, motif 4 and motif 7 were closely related (Figure 3), indicating that they have evolved distinct characteristic sequences from other MYB proteins during the process of evolution. Similar distribution of motifs among MYB proteins within the same subgroup suggested that motif distribution might be related to function. A similar situation also occured with motif 6 and motif 7 (Figure 3). *RobMYB* TFs containing motif 1 and motif 3, *RhpMYB* TFs containing motif 3 and motif 5, and *RnoMYB* TFs containing motif 1 and motif 2 were also the same (Supplementary Figures S11–S13). In addition, we observed that motif 3 was present in almost all members of *RtanMYB* TFs, indicating that motif 3 played an essential role in influencing the function of MYB TFs. In *RobMYB* TFs, motif 2 also played the same role, and the same situation occurs in the *R. palmatum* and *R. nobile* (Supplementary Figures S11–S13). These motifs might exist as the most conserved core MYB domain-forming motifs.

To elucidate the gene structure of the MYB family, we analyzed the introns and exons of the MYB genes. Within *RtanMYB* TFs (Figure 3), MYB genes clustered in the same subgroup have roughly similar exon/intron structures, numbers of exons, and gene lengths. In contrast, other subgroups exhibited greater variation in the number of exons and introns, and this structural diversity of this genes might reflect the evolutionary divergence between homologous MYBs. Furthermore, this could indicate a tendency for genes within a subfamily to evolve conserved functions, while members between subfamilies might undergo more functional differentiation during the evolutionary process. CDS structures of

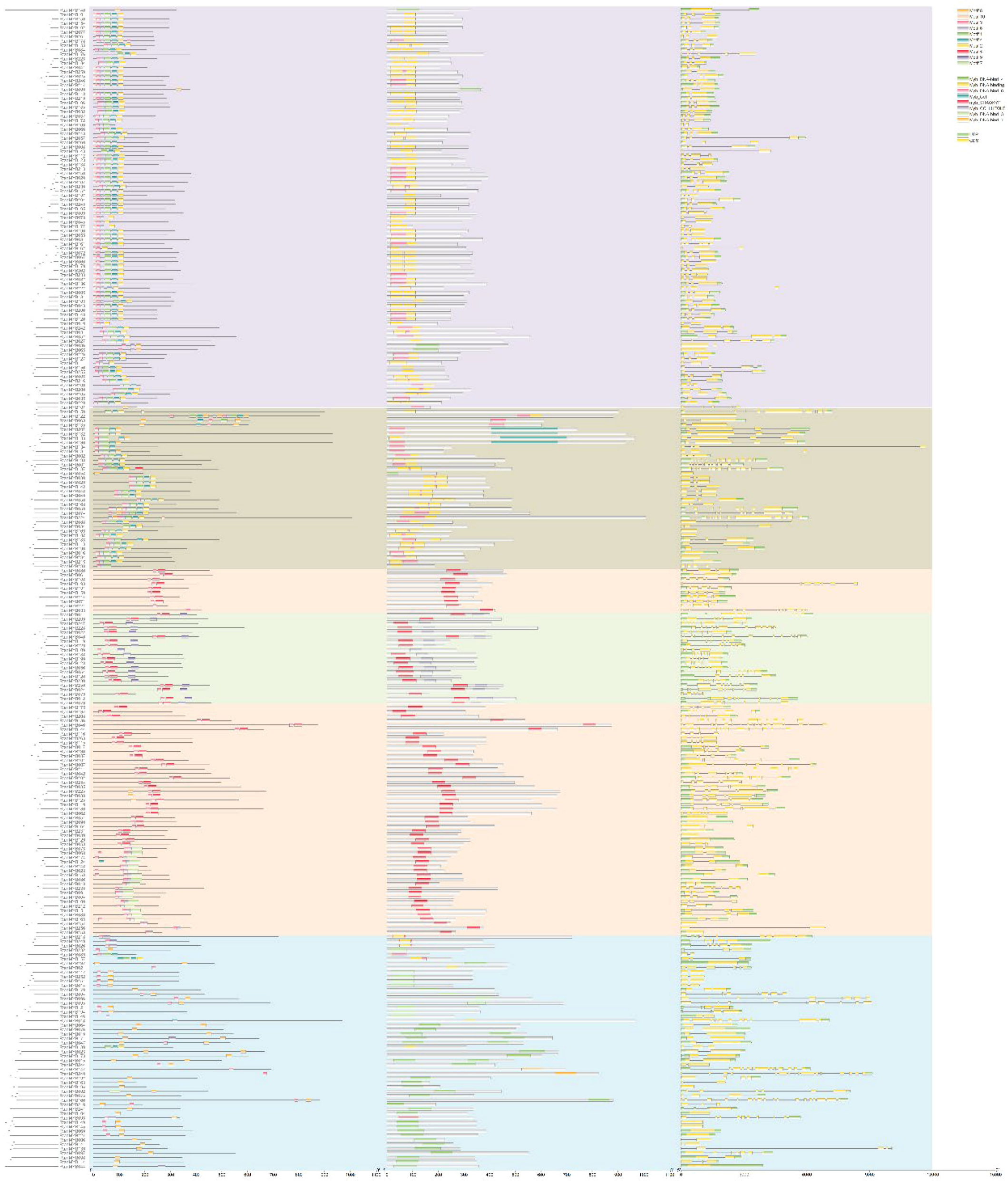


FIGURE 3
The evolutionary tree, conserved motifs, conserved domains, and gene structure of the 263 MYB TFs in *R. tanguticum*.

other three *Rheum* L. species also shown similar patterns (Supplementary Figures S11–S13).

3.5 Analysis of *cis*-acting elements in the promoter region and protein function of MYB TFs

AQs provide photoprotection and enhance disease resistance in plants, which is crucial for high altitude plants. Many studies have reported that environmental conditions affect the activation of MYB TFs regulating plant secondary metabolite synthesis (Ji et al., 2022). Analysis of promoter *cis*-acting elements could provide insights into stress response patterns of genes. 2kb upstream promoter region of MYB TFs was analyzed for CREs. 47 CREs related to plant growth and development and stress response were detected (Supplementary Table S4).

From a variety of CREs, five highly significant ones were screened using keywords related to environmental factors and stresses, such as “motifs” and “boxes”. Most of these response elements are MeJA (CGTCA - motif), ABRE, MBS, and LTR. Notably, MeJA has been reported to significantly induce the expression of genes related to AQs biosynthesis and increase AQs content in plants (Tang et al., 2024). Therefore, screening for

TABLE 2 Characteristics of motifs in MYB TFs related to AQs based on CHSs (Karppinen et al., 2021).

Motifs	Sequence features	Function/ presence
bHLH interaction	[D/E]Lx2[R/K]x3Lx6Lx3R	all MYBs (Except flavonol-regulating MYBs)
ANDV \NDEI	[A/S]NDV\NDE[I/V]	C2 repressors, PA-type MYBs
C1 \C2 \C3 \C4	[K/R]Px3[K/T][F/Y]	C2 repressors, MYB5, PA1-type
TLLLFR	Lx3GIDPx[T/N]H[R/K], pdLNL[D/E]L, dFLGL	subclade D2 of C2 repressors
TT2-box	TLxLF[R/G]	TT2/PA2-type MYBs
PA1	V[V/I]R[T/P][K/R]Ax[R/K]C	PA1-type MYBs
G-28	K[I/V]x2PKPxRx2S[I/L]	PA1-type MYBs
SG7	EKLYEEYLQLLx7-10QLDSFAESLLI	Flavonol-regulating MYBs
SG7-2	[K/R][R/x][R/K]xGRT[S/x][R/G]xx[M/x]K	Flavonol-regulating MYBs

potential MYB TFs involved in AQs biosynthesis regulation based on the number of CGTCA-motifs in specific MYB TFs is a reliable approach. In this regard, *RtanMYB042*, *RtanMYB131*, *RnoGMYB061*, *RhpMYB063* et al. show great potential.

3.6 Chromosomal location, duplication events of MYB genes

Chromosome localization shown that these genes were distributed across 11-12 chromosomes in each species, number and density of MYB genes on each chromosome were unevenly distributed. For example, in *R. tanguticum*, there was a relatively low number of genes found on chromosome 10 (15), while a higher number of genes are found on chromosomes 1 and 3 (>30). Additionally, MYB genes were primarily distributed at the ends of the chromosomes, with fewer in the centromeric regions.

Gene duplication contributed to gene amplification during the process of molecular evolution. They could occur through the following mechanisms: segmental duplication, tandem duplication or retroposition (Qiao et al., 2019). 7 tandemly duplicated genes in *R. tanguticum* were distributed across 3 chromosomes (Supplementary Table S5), with the highest number of genes found on chromosome 2 (Figure 4). Some of these tandemly duplicated genes had been identified as atypical MYBs, and there was also a transition between R2R3-MYB and 3R-MYB. Additionally, 51 pairs of genes related to segmental duplication events were identified, and these genes are similarly dispersed, with members from every subfamily being found. Similar results were also found in other three species (Supplementary Table S5; Figures 4, 5). If a gene lacks introns, is polyadenylated at the 3' end, has short direct repeat sequences at both ends, and is transcribed via mRNA to a chromosomal location different from that of the donor gene, then the gene is considered a retrogene (Liu A. et al., 2020). However, no MYBs were found met these criteria; therefore, it was hypothesized that the MYB TFs in rhubarbs had a low susceptibility to retroposition. Based on aforementioned results, both tandem and segmental duplications might play a key role in the diversity of MYB genes across the four *Rheum* L. species, with segmental duplication being predominant.

Some studies hypothesized that the size of a gene family in many species is a infeasible indicator of genetic diversity, which could be applied broadly to many gene families (Simionato et al., 2007). This study illustrated the extent of commonality in MYB gene repertoires between each pair of the four *Rheum* L. species, with the main indicators being the number of MYB orthologs and ATGs in the concerned species (Figure 6). For example, there were 223 and 205 orthologs between *R. tanguticum* and *R. palmatum*, with 97 one-to-one orthologs, and between *R. tanguticum* and *R. officinale*, this number was 98 (Supplementary Table S6). Therefore, only 36.5% (on average) of MYB genes presented in the most recent common ancestor (MRCA) of *R. tanguticum*, *R. palmatum* and *R. officinale* are conserved in all three species. In contrast, this value is only 33% between *R. tanguticum* and *R. nobile*. However, due to gene duplication events, the ancestral MYB genes in the MRCA have generated 126 and 108 MYB genes in *R. tanguticum* and *R. palmatum* respectively. On average, 73% of the MYB gene repertoires in the four *Rheum* L. plants are shared. Among the six comparisons between the four species, the MYB gene repertoires of *R. tanguticum* and *R. palmatum* are the most similar (86.1%, Supplementary Table S6), revealing their closest evolutionary relationship. However, the proportion of shared genes between the other three species pairs is not high (< 80%). Thus, MYB members in the four *Rheum* L. species revealed by this study are genetically diverse.

The ratio of nonsynonymous to synonymous substitutions (Ka/Ks) was an effective indicator for determining positive selection pressure after duplication and was commonly used to understand the direction of evolution of coding sequences and the intensity of selection (Lynch and Conery, 2000). A Ka/Ks ratio of 1 indicated neutral selection, a Ka/Ks ratio less than 1 indicated purifying selection, and a Ka/Ks ratio greater than 1 indicated positive selection (Lynch and Conery, 2000).

To investigate the role of these duplication events in the evolution and functional divergence of *RtanMYB* genes, we calculated Ka, Ks, and Ka/Ks for each gene pair. Among them, the average Ka and Ks values for tandem duplications of *RtanMYB* genes were lower than those for segmental duplications (Supplementary Table S5), while the average Ka/Ks of segmental duplication genes was lower than that of tandem duplication, and might be candidate genes that have undergone positive selection (>0.5), but the overall Ka/Ks value of *RtanMYB* genes was less than 0.5 (Supplementary Table S5). Similar results were also found in the orthologous gene pairs of *R. palmatum* and *R. offcinale*. The duplication events of segmental and tandem duplication genes in

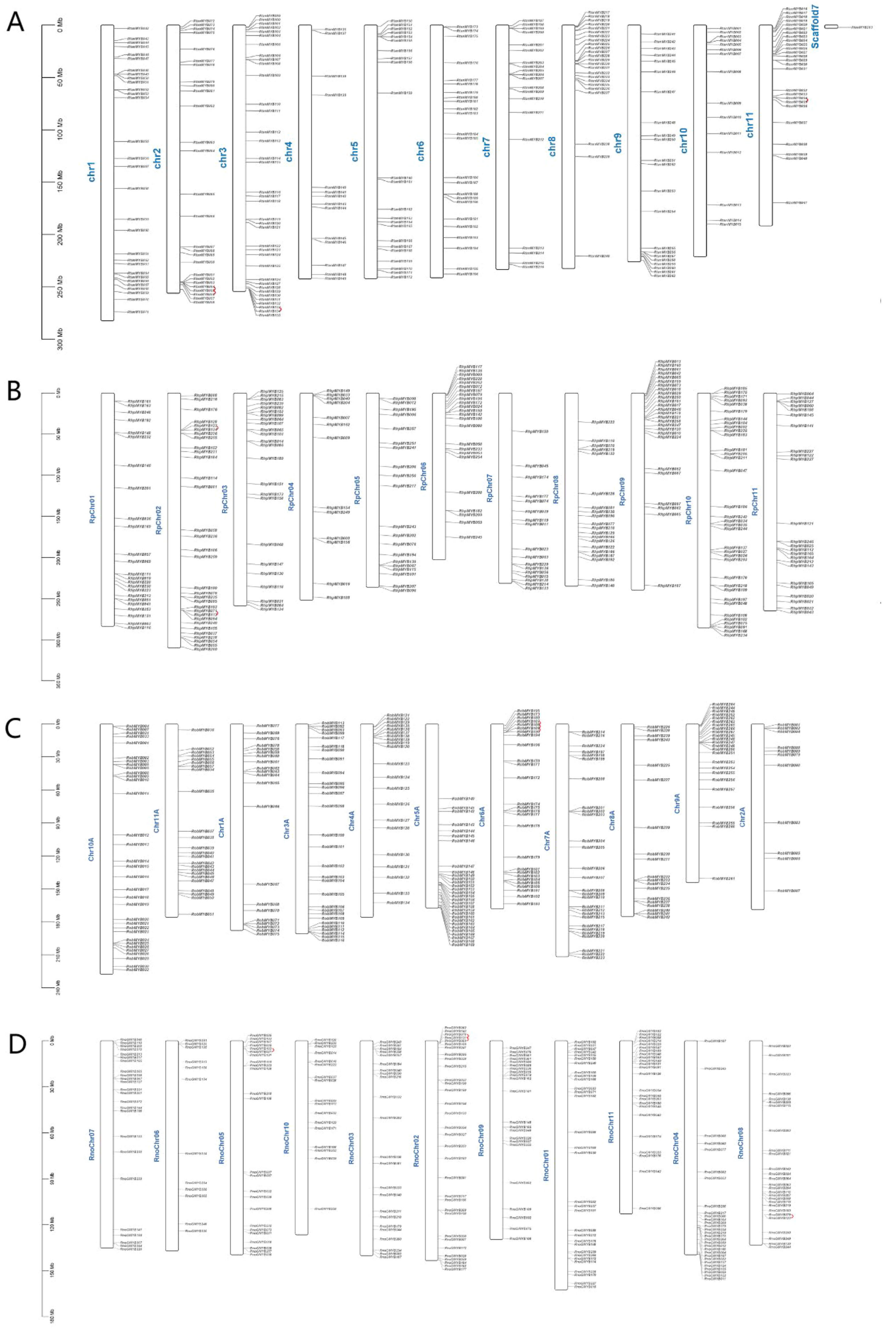


FIGURE 4
The tandem duplication events and distribution of MYB TFs on the 11 chromosomes of the Rheum genus plants. The scale represents the length of the Rheum chromosomes. **(A)** The tandem duplication events and distribution of *RtanMYB* TFs on the chromosomes; **(B)** The tandem duplication events and distribution of *RhpMYB* TFs on the chromosomes; **(C)** The tandem duplication events and distribution of *RobMYB* TFs on the chromosomes; **(D)** The tandem duplication events and distribution of *RnoMYB* TFs on the chromosomes.

R. tanguticum were estimated to have occurred approximately 194 million years ago (Mya) and 47 Mya. These results suggested that these duplication events played an important role in the evolution and functional divergence of *RtanMYB* genes. MYB TFs in *R. palmatum* and *R. offcinale* exhibited the same phenomenon. The divergence times of *R. palmatum* and *R. offcinale* were 210 and 164 Mya, respectively.

3.7 Expression profiling of MYB TFs of *R. tanguticum* in different tissues

PCA of gene expression in different parts of *R. tanguticum* shown great repeatability within each group and could clearly distinguish between different parts. On PC1 (Principal Component 1), roots were significantly separated from stems,

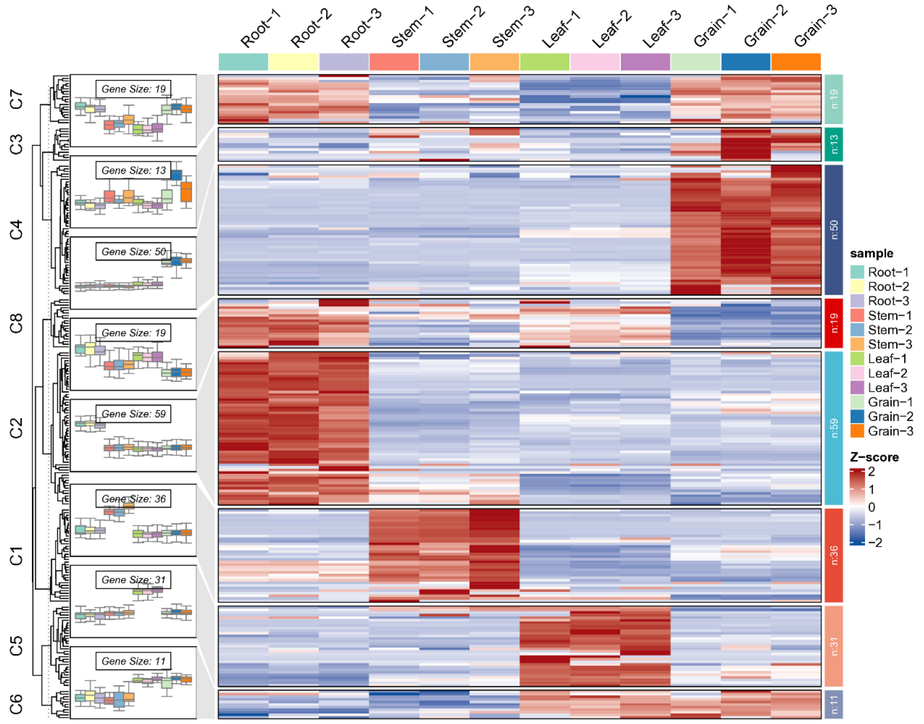


FIGURE 5
Expression heatmap of candidate genes in different tissues of *R. tanguticum* (A clustering heatmap is constructed using Log2 FPKM values, with further annotation in Supplementary Table S7).

leaves, and seeds, indicating that genes expressed in the roots have significant differences from those in other parts. On PC2, roots and seeds were similar, and both were significantly separated from stems and leaves (Supplementary Figure S2B).

Comparative transcriptome analysis revealed that among these 263 *R. tanguticum* MYB TFs, 21 genes were not detected in expression across the four selected tissues (excluding genes with poor reproducibility). Remaining *RtanMYB* TFs were expressed in at least one tissue, with no genes found to be highly expressed in all four components simultaneously. There were 97 up-regulated and 78 down-regulated differential expression MYB genes (DEMGs) in the root compared to stem and leaf samples; In comparison with the seed samples, there were fewer up-regulated DEGs, totaling 78, and a higher number of down-regulated DEMGs, amounting to 93 (Figure 5; Supplementary Table S7). Clustering diagram shown different expression levels among these four groups (Figure 5). Clustering analysis results divided these genes into 8 groups, whose expression profiles in each component did not have universal

FIGURE 6
Duplication and synthesis of MYB TFs in *Rheum* L. plants. Comparative physical mapping indicates that *RtanMYB* TFs are homologous to *R. palmatum*, *R. offcinale*, and *R. nobile*.

similarity, except for MYB TFs in the root and seed, which have a more similar expression profile in C7 and are clustered in a branch. In other clusters, there are significant differences in the expression levels of DEGs, with 42, 97, 82 and 36 highly expressed genes in leaves, roots, seeds, and stems respectively (Figure 5; Supplementary Table S7).

The expression divergence of duplicated genes typically occured after functional differentiation (Duarte et al., 2006). This study analyzed the expression differences between MYB genes that underwent duplication events. The divergence in their clustering pattern diagram suggested a high degree of variability in their expression patterns, indicating that the expression patterns of most gene pairs tended to be uncorrelated in the subject tissues (Supplementary Tables S5, S7; Supplementary Figure S8). Additionally, it was found that there was a phenomenon where the expression level of one gene in a pair of genes was higher than that of the other (Supplementary Tables S5, S7; Supplementary Figure S8). Many studies had demonstrated that tissue-specific expression divergence was one of the most critical indicators of functional differentiation of duplicated genes, and this divergence usually increased expression diversity and led to the evolution of subfunctionalization or neofunctionalization (Li et al., 2005). Although many duplicated genes shown very low expression values, most of them tend to exhibit tissue-specific expression (Supplementary Tables S5, S7; Supplementary Figure S8). Therefore, the expression of MYB genes produced by tandem duplication and segmental duplication exhibited divergence, which might herald the emergence of new biological functions after gene duplication events, facilitating the regulation of various physiological processes.

3.8 Analysis of AQs accumulation pattern and gene co-expression in different tissues of *R. tanguticum*

To investigate the degree of correlation between tissue-specific gene expression and tissue-specific accumulation of AQs, we employed HPLC for quantitative analysis of AQs distribution across various plant components. The results of HPLC analysis showed that, except for Emodin, the other four AQs had higher contents in the roots, followed by leaves and seeds. The content of Emodin was the highest in seeds. Moreover, signals for the five AQs were not detected in the stems (Figure 7B). Considering that AQs compounds mainly accumulate in roots, seeds and leaves of rhubarbs, MYB TFs that highly expressing in roots, seeds, and leaves in medicinal plants were most likely involved in regulating the biosynthesis of AQs. WGCNA analysis revealed four typical tissue-specific gene expression modules, with genes in these modules highly expressed in roots, leaves, and seeds (Figure 7A). The Pearson correlation characterization of the tissue-specific expression of genes within the four modules and the accumulation patterns of the five AQs further confirmed the strong correlation between these highly expressed genes and AQs accumulation. Among them, 49 MYB TFs (Size > 750 bp) were selected (Figure 5; Supplementary Table S7). To further investigate and confirm *RtanMYB* TFs related to the AQs biosynthetic pathway, we visualized the correlation between *RtanMYB* TFs and genes encoding key enzymes involved in AQs biosynthesis (Figure 8A; Supplementary Figure S2A). The results shown that *RtanMYB059*, *RtanMYB083*, *RtanMYB113*, *RtanMYB118*, *RtanMYB183*, *RtanMYB224*, *RtanMYB245*, *RtanMYB042*,

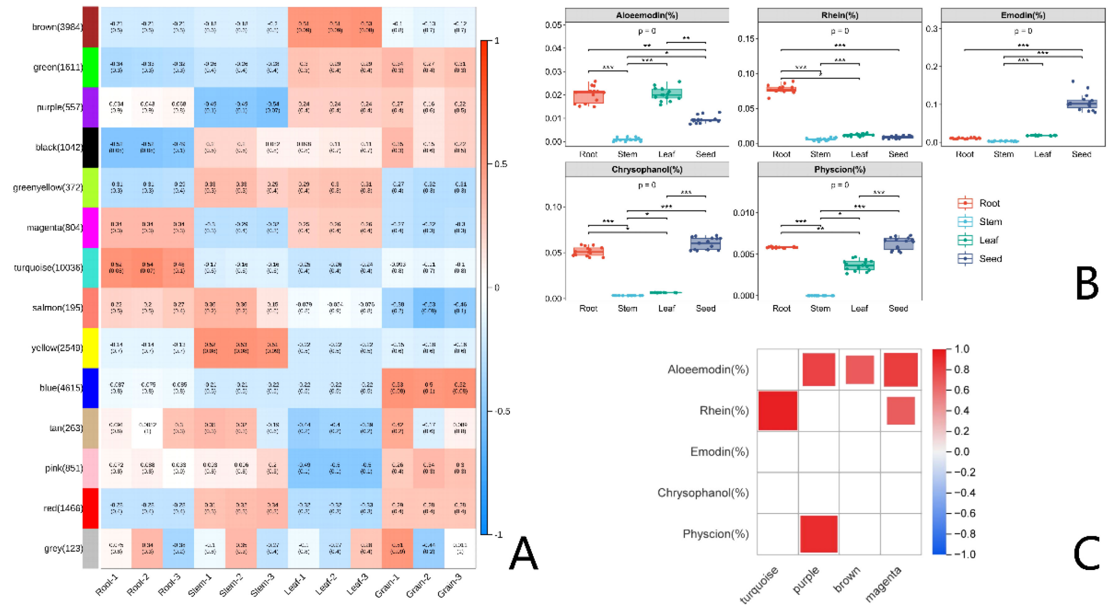


FIGURE 7

The correlation analysis of *RtanMYB* TFs in *R. tanguticum*. **(A)** Weighted Gene Co-Expression Network Analysis. **(B)** Quantitative analysis (HPLC) of AQs in different components of *R. tanguticum*. **(C)** The Pearson's correlation coefficients of AQs with MYBs possibly involved in AQs biosynthesis.

RtanMYB085, *RtanMYB155*, *RtanMYB199* and *RtanMYB217* exhibited strong potential to be involved in AQs synthesis.

3.9 Protein-protein interaction network of MYB TFs involved in the biosynthesis of AQs in *R. tanguticum*

Employing the STRING database, we further predicted the regulatory networks of the candidate MYB TFs and the relevant genes in their protein-protein interaction networks, as illustrated in Figure 8B. Studies had shown that CHS and UGT were involved in the biosynthesis of AQs in plants (Mund and Čellárová, 2023). PPI (Protein-Protein Interaction) simulations for these selected genes revealed that CHS and UGT (UDP-Glycosyltransferase), as central hub proteins, had widespread interaction phenomena with the screened *RtanMYB* TFs. This result indicated that there might be multiple pairs of interacting proteins in the AQs biosynthesis regulatory network of *R. tanguticum*. Centered on CHS, 8 MYB TFs were found to interact with it. Notably, *RtanMYB118* not only interacted with CHS but also with 4 UGT genes, suggesting its general role in this regulatory network. *RtanMYB042*, *RtanMYB059*, and *RtanMYB119* also demonstrated similar potential. We speculated that *RtanMYB* TFs might regulate the accumulation of AQs through these two enzyme genes.

3.10 Expression verification of candidate MYB TFs in *R. tanguticum*

Transcription factors regulated growth, development, and secondary metabolism through their own transcriptional activity. Previous work had provided ample evidence for the involvement of these 12 *RtanMYB* TFs in the biosynthesis of AQs, and qRT-PCR experiments further validated the expression of these 12 transcription factors under actual conditions (Figure 9). The results shown that *RtanMYB118* and *RtanMYB199* were highly expressed in all components except the stem, while the other 10 *RtanMYB* TFs were highly expressed in one or two components, consistent with the transcriptome data. Previous studies had indicated that AQs mainly accumulate in the roots of *R. tanguticum*, considering this situation, there might be a "biosynthesis-transport" scenario in the accumulation of AQs in *R. tanguticum*. Subcellular localization analysis of *RtanMYB118* and *RtanMYB199* shown that they were nuclear-localized (Supplementary Table S1). Therefore, *RtanMYB118/199* was confirmed to play an important role in the biosynthesis of AQs in *R. tanguticum*.

4 Discussion

Rhubarbs contains three important *Rheum* L. species that rich in AQs compounds with significant pharmacological activities (Xiang et al., 2020; Zhao S. et al., 2023). In recent years, research on them mainly focused on the extraction of chemical components (Lu et al., 2024). Due to limitation of genomic information, structural genes and transcription factors that play key roles in biosynthesis of AQs are still unclear (Wang et al., 2023). MYB TFs play a crucial role in regulating plant physiological activities (Ma et al., 2023). They had increasingly attracted the attention of researchers in various plant species and has been extensively studied in model plants such as *Arabidopsis thaliana* (Yanhui et al., 2006), *Oryza sativa* (Katiyar et al., 2012), *Solanum lycopersicum* (Guo et al., 2022), as well as in many other species, including

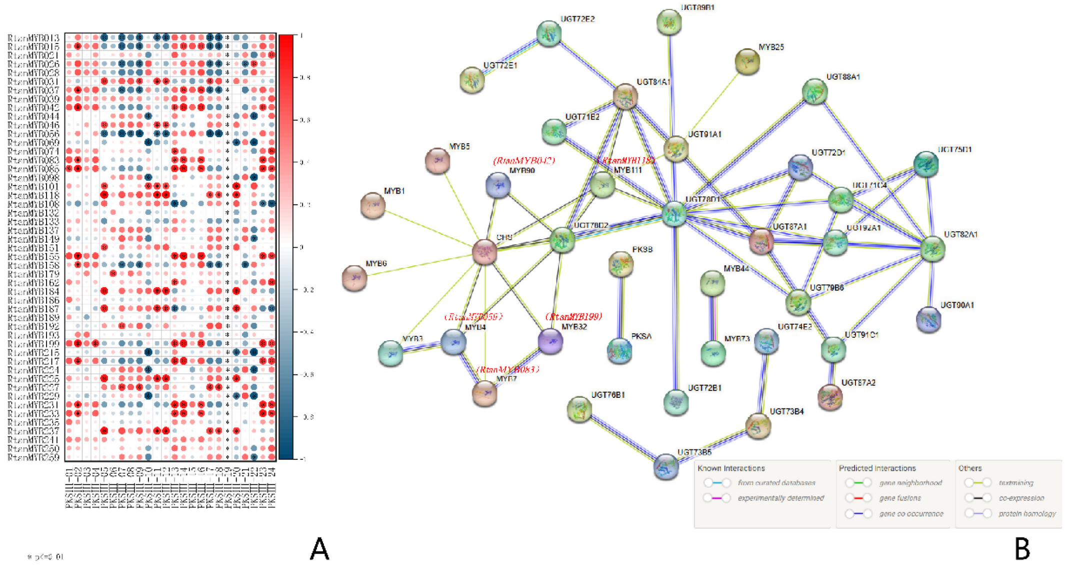


FIGURE 8
The expression analysis of *RtanMYBs*. **(A)** The characterization of Pearson correlation coefficients between PKS genes involved in AQs biosynthesis and MYB TFs (*p<=0.01). **(B)** The protein-protein interaction simulation of 12 *RtanMYB* TFs with the PKS and UGT gene families.

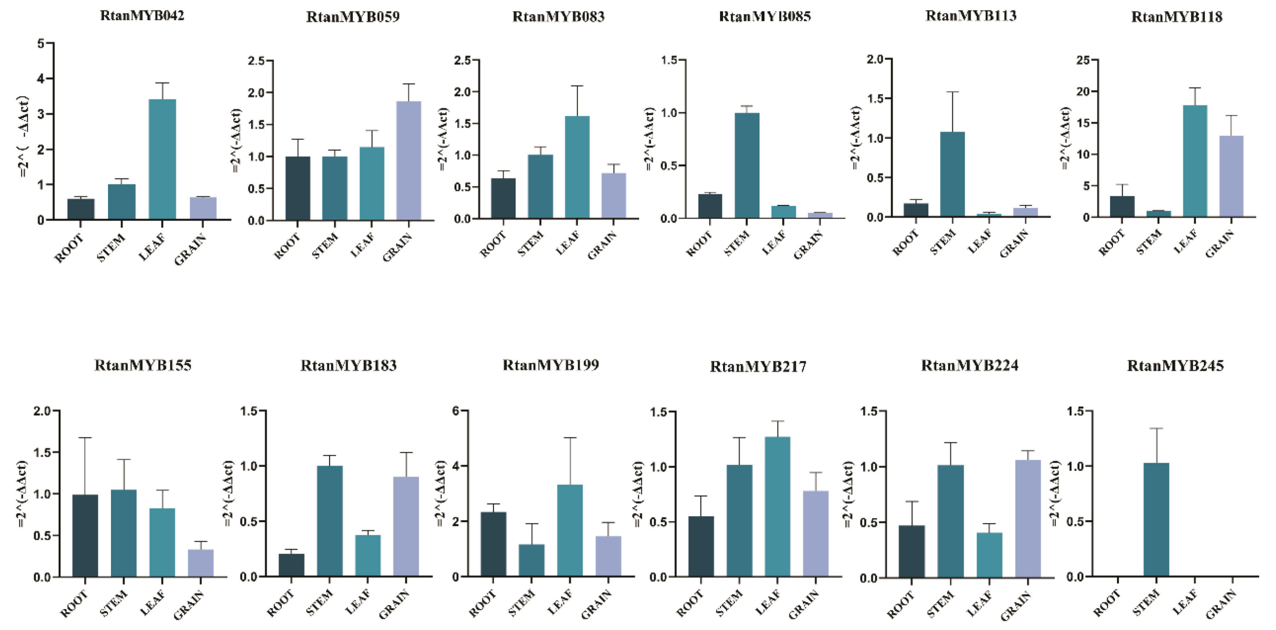


FIGURE 9
The qRT-PCR of 12 candidate *RtanMYB* TFs (n=3).

Morella rubra (Cao et al., 2021). Given the important regulatory role of MYB genes in plant physiology, understanding these genes is crucial for plant research, including breeding and crop improvement. Given the critical regulatory roles of MYBs in plant physiology, understanding these genes is essential for plant research, including breeding and crop improvement. At present, research on whole-genome MYB TFs of rhubarbs and their involvement in AQs biosynthesis is still quite limited. However, studies on the evolutionary relationships and functions of transcription factors using multi-omics approaches in other species have become relatively mature (Hao et al., 2023; Tao et al., 2018). This provides valuable experience for further research into the molecular mechanisms of AQ biosynthesis mediated by MYB TFs in rhubarbs.

4.1 Analysis of evolution and expansion of MYB TFs

Studies had shown that plant genomes typically evolve more rapidly than those of other eukaryotes, leading to greater genomic diversity even among closely related plant species (Kejnovsky et al., 2009). In this study, analysis of 1054 MYB genes identified from genomic sequences, as well as 197 MYB TFs from *Arabidopsis thaliana*, revealed that the difference in the number of MYB genes among the four *Rheum* L. species was not very large, and the number of genes did not have a linear correlation with genome size, but it did support the notion of genomic diversity mentioned above (Supplementary Table S2). Phylogenetic analysis can provide a clear picture of the evolutionary history of MYB TFs (Li et al., 2021). Evolutionary analysis indicated that MYB TFs from four species can be divided into four groups, with 1R-MYB and R2R3-MYB groups further divided into different subclasses (Supplementary Figures S3, S4). Within the (sub)groups, there was a significant variation in gene numbers between different species, a phenomenon particularly pronounced in the R2R3-MYB group. The differences in group division and the number of genes within each group reflect the evolutionary patterns of different species. Additionally, a decrease in gene pairs between *R. tanguticum* and *R. palmatum*, *R. tanguticum* and *R.nobile*, was observed. Considering the geographical distribution and cultivation status of these three species, this might be due to differences in domestication. Such differences could further shape the functional divergence of MYB TFs within different groups.

Introns offer insights into the evolutionary trajectories of genes and their corresponding proteins (Li et al., 2024). Their relative positioning can further drive the diversification of gene families (Li et al., 2024). MYB TFs in *R. tanguticum* exhibit variation in intron numbers, suggesting that MYB genes have undergone intron gain or loss during evolution, potentially leading to functional divergence (Figure 3). Research indicated that genes with a lower intron count may correlate with higher expression levels (He et al., 2021). qPCR validation of *RtanMYB* genes, which were identified as having higher expression levels in the transcriptome profile, supports this hypothesis (Supplementary Table S7; Figure 9).

4.2 Gene duplication events play important roles in MYBs family expansion

Gene duplication events are considered an important mechanism leading to the expansion of gene families and are prevalent in plant genomes (Panchy et al., 2016). Synteny analysis helps to identify the functional and evolutionary relationships of genes within the same species or between different species. This study reveals that multiple paralogous gene clusters (52 pairs) and tandem duplicated gene clusters (4 pairs) in *R. tanguticum* originated from segmental and tandem duplications (Figures 4, 10). Among them, segmental duplication plays a larger role in the expansion of MYB TFs, a phenomenon widely recognized in the 1R-MYB and R2R3-MYB subfamilies. Similar situations also exist in *R. palmatum*, *R. officinale*, and *R. nobile*. Since duplicated genes always tend to undergo subfunctionalization or neofunctionalization to reduce functional redundancy, there may be expression differences between duplicated genes (Duarte et al., 2006). Therefore, new and old duplicated genes often have significant differences in epigenetic modifications, expression patterns, and other aspects, which are related to the functional divergence in expression. For transcription factors, this also leads to the rewiring of transcriptional regulatory networks and changes in transcription factor complexes. Therefore, studying the protein interactions of MYB TFs that have undergone duplication events is of great significance for exploring their functions. Most of the duplicated *RtanMYB* genes from tandem or segmental duplications tend to be unrelated in the four tested tissues, and the expression value of one gene in the tissues is much greater than that of the other gene in the pair (Supplementary Tables S5, S7; Supplementary Figure S8; Supplementary Table S7). This expression divergence can, to some extent, increase the diversity of expression patterns and promote the subfunctionalization or neofunctionalization of duplicated genes.

By constructing collinearity maps among the four *Rheum* L. species, the evolutionary relationships of MYB genes within these species were analyzed. The collinearity maps shown that the number of gene pairs between *R. tanguticum* and the other three species (*R. palmatum*, *R. officinale*, and *R. nobile*) are 103, 91, and 101, respectively. The variation in the number of collinear gene pairs, including the phenomenon of certain gene pairs specifically appearing between certain *Rheum* L. species (for example, *RtanMYB079/098*), reflects the conservation status between the two species and indicates that new duplication events have occurred within the MYB gene family during the divergence of these species. Notably, some MYB TFs were found to be associated with at least two collinear gene pairs, suggesting that these genes may be crucial for the evolution of the MYB gene family (Supplementary Table S6).

Ka/Ks analysis indicates that within or between four species of this four *Rheum* L. genus, MYB TFs have mostly undergone purifying selection inside the duplicated genomic elements during the process of speciation. Few of these genes have been subject to positive selection (Supplementary Table S5). These results suggest that MYB genes in four *Rheum* L. genus have experienced strong purifying selection during their evolutionary process, undergoing minor changes after duplication. In previous studies of the MYB gene family, such as *G. hirsutum* (Salih et al., 2016) and *H. coronarium* (Abbas et al., 2021), a high proportion of purifying selection and a low proportion of positive selection have also been reported.

4.3 Some *RtanMYB* TFs play critical roles in the synthesis of AQs

Structural genes encoding compound synthesis enzymes play an important role in the biosynthesis of plant secondary metabolites (Golovko, 2023). Although AQs biosynthesis has been studied in some medicinal plants, their biosynthetic pathways and regulatory mechanisms are not yet fully elucidated. Our previous studies comprehensively confirms that DAHPS, DHQS, DHQD/SDH, SMK, EPSP, CS, ICS, PHYLLO, MenE, and MenB are involved in the shikimate pathway, producing the precursor 1,4-dihydroxy-2-naphthoyl-CoA, and in the polyketide pathway, PKSIII and PKC are involved in the synthesis of AQs precursors. In *R. tanguticum*, PKSIII and UGT genes have been confirmed to participate in AQs biosynthesis. We have screened these enzyme genes from the whole-genome data of *R. tanguticum*, including those with completed functional validation and pending validation, and analyzed their transcriptome profiles (Supplementary Figure S2A).

CREs are crucial for controlling transcriptional regulation of various biological processes, one of which is the synthesis of plant secondary metabolites (Li et al., 2024). Existing research has reported that the promoter regions of 24 PKSIII genes and 97

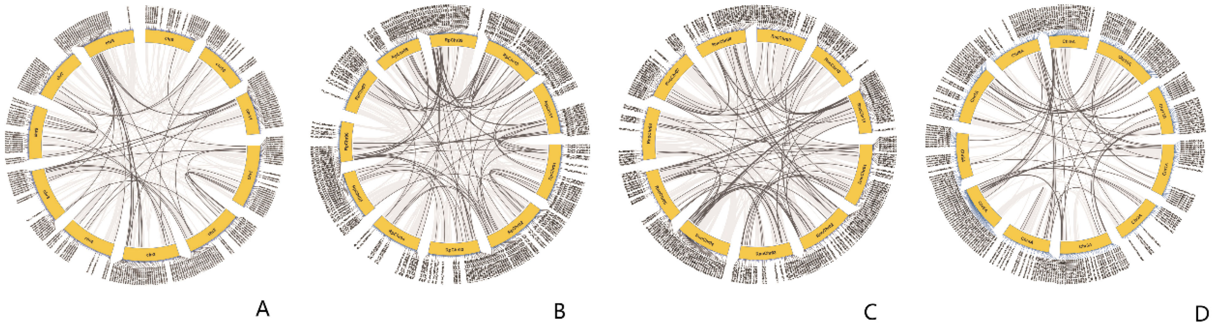


FIGURE 10
Analysis of duplication events in MYB TFs of *Rheum* L. plants. **(A)** Segmental duplication events of MYB TFs in *R. tanguticum*; **(B)** Segmental duplication events of MYB TFs in *R. palmatum*; **(C)** Segmental duplication events of MYB TFs in *R. nobile*; **(D)** Segmental duplication events of MYB TFs in *R. offcinale*.

UGT genes involved in AQs biosynthesis in *R. tanguticum* universally contain MYB binding sites (Supplementary Figure S1). Correlation analysis based on the Pearson index indirectly confirms this conclusion, indicating that genes for enzymes involved in the upstream and downstream of AQs biosynthesis generally exhibit strong positive or negative correlations with *RtanMYB* TFs. (Supplementary Figure S2A).

There have been no reports on the involvement of MYB TFs from closely related species of *R. tanguticum* in AQs biosynthesis. Therefore, this study successively carried out phylogenetic analysis, sequence analysis, and gene family evolution analysis. Based on the quantitative study of tissue-specific accumulation of AQs, tissue-specifically expressed *RtanMYB* TFs were screened. These were then correlated with the 24 PKSIII genes that have been confirmed to participate in AQs biosynthesis (Figure 8B). The results shown a strong response from 12 *RtanMYB* TFs, including *RtanMYB118* and *RtanMYB199*.

The protein interaction results of the 12 *RtanMYB* TFs simulated in *Arabidopsis thaliana* shown that homologs of these *RtanMYB* TFs interacted with the central UGT and CHS proteins. CHS and UGT are both key enzyme genes involved in plant AQs biosynthesis, and we speculate that they may regulate the accumulation of AQs through CHS and UGT.

The RNA-seq and qRT-PCR results are consistent at plant tissue level. Based on the RNA-seq results, the expression levels of some genes are relatively high in roots, leaves, and seeds, especially *RtanMYB118* and *RtanMYB199*. RT-qPCR shows that *RtanMYB118* and *RtanMYB199* exhibit the same expression pattern (Figure 9). Subcellular localization results indicate that both *RtanMYB118* and *RtanMYB199* have nuclear localization signals, which further confirms their transcriptional regulatory functions in AQs biosynthesis. Tissue-specific expression may confer different biological functions on family members. Functionally similar genes are closely related in the evolutionary process and often have high homology; the similarity in the expression patterns of *RtanMYB118* and *RtanMYB199* can explain the consistency of their functions.

Based on the evidence presented, *RtanMYB118/119* can serve as important targets for research on AQs biosynthesis and subsequent molecular breeding of *R. tanguticum*.

5 Conclusions

In summary, through comparative genomic analysis, 1054 MYB TFs were identified and characterized. Comparative analysis of molecular features and extensive phylogenetic analysis revealed the genetic diversity of MYB TFs in four *Rheum* L. plants. Traditional classification and optimization of the MYB TFs were completed based on sequence and evolutionary features, with four *Rheum* L. species of different medicinal value and domestication levels showing significant expansion or loss to varying degrees. Tandem and segmental duplications both contributed to the expansion of the MYBs, but segmental duplication events were markedly dominant. Expression profiles of four tissues revealed that duplicated genes generally exhibit divergent expression patterns, implying a tendency towards neofunctionalization or subfunctionalization after duplication. A detailed analysis of the expression of MYBs in different tissues allows us to identify DEGs in these tissues, and this data will make a significant contribution to future functional genomics research. In addition, this study also discovered and confirmed *RtanMYB118/199*, which affects the biosynthesis of AQs. In summary, phylogenetic analysis has laid the foundation for further research on the physiological functions of MYB TFs in rhubarbs; the MYB TFs in rhubarbs play an important regulatory role in the biosynthesis of AQs, providing an important research target for the molecular breeding studies of rhubarbs as a medicinal resource.

Data availability statement

Publicly available datasets were analyzed in this study This data can be found here: China National Center for Bioinformation (CNCB, https://www.cncb.ac.cn/) repository, with accession number PRJCA037598, PRJCA012373 and CNP0003451. National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/), with accession number PRJNA719574, PRJNA735904, PRJNA827652, and PRJNA1049137. The assembled genome and the genome annotation are available at figshare database (https://doi.org/10.6084/m9.figshare.25495309).

Author contributions

TW: Conceptualization, Investigation, Software, Validation, Visualization, Writing – original draft. SZ: Conceptualization, Methodology, Project administration, Validation, Writing – review & editing. BW: Investigation, Validation, Writing – review & editing. JL: Formal Analysis, Validation, Visualization, Writing – review & editing. ZY: Formal Analysis, Methodology, Validation, Visualization, Writing – review & editing. FZ: Formal Analysis, Software, Validation, Visualization, Writing – review & editing. HM: Investigation, Software, Visualization, Writing – review & editing. GZ: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Major Science and Technology Special Project of Qinghai Province (2023-SF-A5), Central Forestry Reform and Development Fund 2022 and Central Forestry Reform and Development Fund 2023.

Conflict of interest

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

Generative AI statement

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1558321/full#supplementary-material

Additional file 1: Supplementary Figure S1. The PCA analysis of unigenes based on the expression levels in four major components of R. tanguticum; each treatment was performed in three replicates. Supplementary Figure S2. The characterization of Pearson correlation coefficients between genes related to the shikimate pathway in AQs biosynthesis and MYB TFs (*p<=0.01). Supplementary Figure S3-S6. The phylogenetic tree of MYB TFs subfamily classification system for four Rheum species and Arabidopsis (1R-MYB, R2R3-MYB, 3R-MYB, 4R-MYB). Supplementary Figure S7-S10. The protein conserved motifs of MYB TFs in Rheum plants. Supplementary Figure S11-S13. Phylogenetic analysis, motif composition, conserved domains, and gene structure of MYB TFs in Rheum plants.

Additional file 2: Supplementary Table. The basic data on the screened MYB TFs from Rheum plants and their classification and analysis.

References

Abbas, F., Ke, Y., Zhou, Y., Yu, Y., Waseem, M., Ashraf, U., et al. (2021). Genome-wide analysis reveals the potential role of MYB transcription factors in floral scent formation in *Hedychium coronarium. Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.623742

Allan, A. C., Hellens, R. P., and Laing, W. A. (2008). MYB transcription factors that colour our fruit. *Trends Plant Sci.* 13, 99–102. doi: 10.1016/j.tplants.2007.11.012

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Cao, Y., Jia, H., Xing, M., Jin, R., Grierson, D., Gao, Z., et al. (2021). Genome-wide analysis of MYB gene family in Chinese bayberry (*Morella rubra*) and identification of members regulating flavonoid biosynthesis. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.691384

Chen, K., Wang, B., Chen, C., and Zhou, G. (2024). The relationship between niche breadth and phylogenetic characteristics of eight species of rhubarb on the Qinghai-Tibet Plateau, Asia. *Ecol. Evol.* 14, e11040. doi: 10.1002/ece3.11040

Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., et al. (2023). TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol. Plant* 16, 1733–1742. doi: 10.1016/j.molp.2023.09.010

Chinese Pharmacopoeia Commission (2020). *Pharmacopoeia of the People's Republic of China* (Beijing: China Medical Science Press).

Cui, D., Xiong, G., Ye, L., Gornall, R., Wang, Z., Heslop-Harrison, P., et al. (2024). Genome-wide analysis of flavonoid biosynthetic genes in Musaceae (Ensete, Musella, and Musa species) reveals amplification of flavonoid 3′,5′-hydroxylase. *AoB Plants* 16, plae049. doi: 10.1093/aobpla/plae049

Duarte, J. M., Cui, L., Wall, P. K., Zhang, Q., Zhang, X., Leebens-Mack, J., et al. (2006). Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of Arabidopsis. *Mol. Biol. Evol.* 23, 469–478. doi: 10.1093/molbev/msj051

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in arabidopsis. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Feng, T., Pucker, B., Kuang, T., Song, B., Yang, Y., Lin, N., et al. (2023). The genome of the glasshouse plant noble rhubarb (*Rheum nobile*) provides a window into alpine adaptation. *Commun. Biol.* 6, 706. doi: 10.1038/s42003-023-05044-1

Frampton, J. (2004). *Myb transcription factors: Their role in growth, differentiation and disease* (Netherlands: Springer). doi: 10.1007/978-1-4020-2869-4

Golovko, T. K. (2023). Plant anthocyanins: structure, biosynthesis regulation, functions, and ecology. *Russ J. Plant Physiol.* 70, 161. doi: 10.1134/S1021443723700292

Guo, H., Mao, M., Deng, Y., Sun, L., Chen, R., Cao, P., et al. (2022). Multi-omics analysis reveals that *SlERF.D6* synergistically regulates SGAs and fruit development. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.860577

Hao, Y., Xiang, L., Lai, J., Li, C., Zhong, Y., Ye, W., et al. (2023). SlERF.H6 mediates the orchestration of ethylene and gibberellin signaling that suppresses bitter-SGA biosynthesis in tomato. *New Phytol.* 239, 1353–1367. doi: 10.1111/nph.19048

He, F., Shi, Y. J., Zhao, Q., Zhao, K. J., Cui, X. L., Chen, L. H., et al. (2021). Genome-wide investigation and expression profiling of polyphenol oxidase (PPO) family genes uncover likely functions in organ development and stress responses in Populus trichocarpa. *BMC Genomics* 22, 731. doi: 10.1186/s12864-021-08028-9

Ji, X., Tang, J., and Zhang, J. (2022). Effects of salt stress on the morphology, growth and physiological parameters of *Juglansmicrocarpa* L. *Seedlings. Plants* 11, 2381. doi: 10.3390/plants11182381

Jin, H., and Martin, C. (1999). Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol Biol.* 41 (5), 577–585. doi: 10.1023/a:1006319732410

Kang, S. H., Pandey, R. P., Lee, C. M., Sim, J. S., Jeong, J. T., Choi, B. S., et al. (2020). Genome-enabled discovery of anthraquinone biosynthesis in *Senna tora*. *Nat. Commun.* 11, 5875. doi: 10.1038/s41467-020-19681-1

Karppinen, K., Lafferty, D.J., Albert, N.W., Mikkola, N., McGhie, T., Allan, A.C., et al. (2001). MYBA and MYBPA transcription factors co-regulate anthocyanin biosynthesis in blue-coloured berries. *New Phytol* 232, 1350–1367. doi: 10.1111/nph.17669

Katiyar, A., Smita, S., Lenka, S. K., Rajwanshi, R., Chinnusamy, V., and Bansal, K. C. (2012). Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. *BMC Genomics* 13, 544. doi: 10.1186/1471-2164-13-544

Kejnovsky, E., Leitch, I. J., and Leitch, A. R. (2009). Contrasting evolutionary dynamics between angiosperm and mammalian genomes. *Trends Ecol. Evol.* 24, 572–582. doi: 10.1016/j.tree.2009.04.010

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

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Leistner, E. (1985). "Biosynthesis of chorismate-derived quinones in plant cell cultures," in *Primary and Secondary Metabolism of Plant Cell Cultures.* Proceedings in Life Sciences. Eds. K. H. Neumann, W. Barz and E. Reinhard (Springer, Berlin, Heidelberg). doi: 10.1007/978-3-642-70717-9

Leistner, E., and Zenk, M. H. (1968). Mevalonic acid a precursor of the substituted benzenoid ring of anthraquinones. *Tetrahedron Lett.* 9, 1395–1396. Available at: https://www.academia.edu/25120436/ (Accessed 12, 2024).

Li, J. L., Li, H., Zhao, J. J., Yang, P., Xiang, X., Wei, S. Y., et al. (2024). Genome-wide identification and characterization of the RZFP gene family and analysis of its expression pattern under stress in *Populus trichocarpa*. *Int. J. Biol. Macromol.* 255, 128108. doi: 10.1016/j.ijbiomac.2023.128108

Li, H., Luo, Y., Gan, L., Ma, P. F., Gao, L. M., Yang, J. B., et al. (2021). Plastid phylogenomic insights into relationships of all flowering plant families. *BMC Biol.* 19, 232. doi: 10.1186/s12915-021-01166-2

Li, Y., Wang, Z., Zhu, M., Niu, Z., Li, M., Zheng, Z., et al. (2023). A chromosome-scale Rhubarb (Rheum tanguticum) genome assembly provides insights into the evolution of anthraquinones biosynthesis. *Commun. Biol.* 6, 867. doi: 10.1038/s42003-023-05248-5

Li, W. H., Yang, J., and Gu, X. (2005). Expression divergence between duplicate genes. *Trends Genet.* 21, 602–607. doi: 10.1016/j.tig.2005.08.006

Liu, C., Feng, C., Peng, W., Hao, J., Wang, J., Pan, J., et al. (2020). Genetic diversity analysis of Chinese plum (*Prunus salicina* L.) based on whole-genome resequencing. *Hortic. Res.* 7, 1–12. doi: 10.1007/s11295-021-01506-x

Liu, A., Liu, C., Haiying, L., Zhijun, W., Min, Z., Yan, X., et al. (2020). Phylogenetic analysis and transcriptional profiling of WRKY genes in sunflower (*Helianthus annuus* L.): Genetic diversity and their responses to different biotic and abiotic stresses. *Industrial Crops Prod.* 148, 112268. doi: 10.1016/j.indcrop.2020.112268

Liu, Y., Zeng, Y., Li, Y., Liu, Z., Lin-Wang, K., Espley, R. V., et al. (2020). Genomic survey and gene expression analysis of the MYB-related transcription factor superfamily in potato (*Solanum tuberosum* L.) *Int. J. Biol. Macromol.* 164, 2450–2464. doi: 10.1016/j.ijbiomac.2020.08.062

Lu, W., Chen, T., Shen, C., Feng, H., Lin, C., Song, Z., et al. (2024). Differential recognition and targeting isolation of HPLC co-efflux isomers using countercurrent chromatography based on LC–MS feature analysis. *Microchem. J.* 205, 111384. doi: 10.1016/j.microc.2024.111384

Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151–1155. doi: 10.1126/science.290.5494.1151

Lyu, R., Gao, Y., Wu, T., Ye, C., Wang, P., He, C., et al. (2024). Quantitative analysis of *cis*-regulatory elements in transcription with KAS-ATAC-seq. *Nat. Commun.* 15, 6852. doi: 10.1038/s41467-024-50680-8

Ma, R., Liu, B., Geng, X., Ding, X., Yan, N., Sun, X., et al. (2023). Biological function and stress response mechanism of MYB transcription factor family genes. *J. Plant Growth Regul.* 42, 83–95. doi: 10.1007/s00344-021-10557-2

Martin, C., and Paz-Ares, J. (1997). MYB transcription factors in plants. *Trends Genet.* 13, 67–73. doi: 10.1016/S0168-9525(96)10049-4

Millard, P. S., Kragelund, B. B., and Burow, M. (2019). R2R3 MYB transcription factors – functions outside the DNA-binding domain. *Trends Plant Sci.* 24, 934–946. doi: 10.1016/j.tplants.2019.07.003

Mund, N. K., and Čellárová, E. (2023). Recent advances in the identification of biosynthetic genes and gene clusters of the polyketide-derived pathways for anthraquinone biosynthesis and biotechnological applications. *Biotechnol. Adv.* 63, 108104. doi: 10.1016/j.bioteChadv.2023.108104

Nei, M., and Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet.* 39, 121–152. doi: 10.1146/annurev.genet.39.073003.112240

Ogata, K., Kanei-Ishii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., et al. (1996). The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation. *Nat. Struct. Biol.* 3, 178–187. doi: 10.1038/nsb0296-178

Panchy, N., Lehti-Shiu, M., and Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant Physiol.* 171, 2294–2316. doi: 10.1104/pp.16.00523

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A., and Saedler, H. (1987). The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J.* 6, 3553–3558. doi: 10.1002/j.1460-2075.1987.tb02684.x

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

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., et al. (2019). Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. *Genome Biol.* 20, 38. doi: 10.1186/s13059-019-1650-2

Salih, H., Gong, W., He, S., Sun, G., Sun, J., and Du, X. (2016). Genome-wide characterization and expression analysis of MYB transcription factors in *Gossypium hirsutum*. *BMC Genet.* 17, 129. doi: 10.1186/s12863-016-0436-8

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Si, Z., Wang, L., Ji, Z., Zhao, M., Zhang, K., and Qiao, Y. (2023). Comparative analysis of the MYB gene family in seven Ipomoea species. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1155018

Simionato, E., Ledent, V., Richards, G., Thomas-Chollier, M., Kerner, P., Coornaert, D., et al. (2007). Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* 7, 33. doi: 10.1186/1471-2148-7-33

Su, S., Wu, J., Gao, Y., Luo, Y., Yang, D., and Wang, P. (2020). The pharmacological properties of chrysophanol, the recent advances. *Biomed. Pharmacother.* 125, Article 110002. doi: 10.1016/j.biopha.2020.110002

Tan, S. M., Ho, L. N., Wong, Y. S., Abidin, C. Z. A., and Ong, S. A. (2023). Sustainable utilization of anthraquinone-rich Rheum officinale as electron shuttle in microbial fuel cell: strategy for stimulating monohydric phenols degradation and bioelectricity generation. *Chem. Eng. J.* 475, Article 146423. doi: 10.1016/j.cej.2023.146423

Tang, D., Chen, M., Huang, X., Zhang, G., Zeng, L., Zhang, G., et al. (2023). SRplot: A free online platform for data visualization and graphing. *PloS One* 18, e0294236. doi: 10.1371/journal.pone.0294236

Tang, J., Li, Y.-m., Wang, Y., Yan, F., Feng, Z., LV, R.-h., et al. (2024). Comparative transcriptome analysis and identification of candidate bZIP transcription factors involved in anthraquinone biosynthesis in *Rheum officinale* Baill. *Genomics* 116, 110948. doi: 10.1016/j.ygeno.2024.110948

Tao, C., Jin, X., Zhu, L., Xie, Q., Wang, X., and Li, H. (2018). Genome-wide investigation and expression profiling of APX gene family in *Gossypium hirsutum* provide new insights in redox homeostasis maintenance during different fiber development stages. *Mol. Genet. Genomics* 293, 685–697. doi: 10.1007/s00438-017-1413-2

Thiedig, K., Weisshaar, B., and Stracke, R. (2021). Functional and evolutionary analysis of the Arabidopsis 4R-MYB protein SNAPc4 as part of the SNAP complex. *Plant Physiol.* 185, 1002–1020. doi: 10.1093/plphys/kiaa067

Tuan, P. A., Bai, S., Yaegaki, H., and Peace, C. P. (2015). The crucial role of *PpMYB10.1* in anthocyanin accumulation in peach and relationships between its allelic type and skin color phenotype. *BMC Plant Biol.* 15, 280. doi: 10.1186/s12870-015-0664-5

Wang, D., Wang, X.-H., Yu, X., Cao, F., Cai, X., Chen, P., et al. (2021). Pharmacokinetics of Anthraquinones from Medicinal Plants. *Front. Pharmacol.* 12, 638993. doi: 10.3389/fphar.2021.638993

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49. doi: 10.1093/nar/gkr1293

Wang, P., Wei, J., Hua, X., Dong, G., Dziedzic, K., Wahab, A. T., et al. (2023). Plant AQs: Classification, distribution, biosynthesis, and regulation. *J. Cell Physiol.* 239, e31063. doi: 10.1002/jcp.31063

Wang, A., Yang, M., and Liu, J. (2005). Molecular Phylogeny, Recent Radiation and Evolution of Gross Morphology of the Rhubarb genus Rheum (Polygonaceae) Inferred from Chloroplast DNA trnL-F Sequences. *Ann. Bot.* 96, 489–498. doi: 10.1093/aob/mci201

Wu, Y., Wen, J., Xia, Y., Zhang, L., and Du, H. (2022). Evolution and functional diversification of R2R3-MYB transcription factors in plants. *Hortic. Res.* 9, uhac058. doi: 10.1093/hr/uhac058

Xiang, H., Zuo, J., Guo, F., and Dong, D. (2020). What we already know about rhubarb: a comprehensive review. *Chin. Med.* 15, 88. doi: 10.1186/s13020-020-00370-6

Yang, X., Dai, L., Yan, F., Ma, Y., Guo, X., Jenis, J., et al. (2024). The phytochemistry and pharmacology of three Rheum species: A comprehensive review with future perspectives. *Phytomedicine* 131, 155772. doi: 10.1016/j.phymed.2024.155772

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., et al. (2006). The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol. Biol.* 60, 107–124. doi: 10.1007/s11103-005-2910-y

Yao, X., Yang, S., Lai, D., Weng, W., Fan, Y., Wu, W., et al. (2024). Genome-wide identification, evolution, and expression level analysis of the TALE gene family in Sorghum bicolor. *BMC Plant Biol.* 24, 1152. doi: 10.1186/s12870-024-05735-9

Zhang, H., He, Q., Xing, L., Wang, R., Wang, Y., Liu, Y., et al. (2024). The haplotype-resolved genome assembly of autotetraploid rhubarb *Rheum officinale* provides insights into its genome evolution and massive accumulation of anthraquinones. *Plant Commun.* 5, 100677. doi: 10.1016/j.xplc.2023.100677

Zhang, H., Liu, Z., Luo, R., Sun, Y., Yang, C., Li, X., et al. (2022). Genome-Wide Characterization, Identification and Expression Profile of MYB Transcription Factor Gene Family during Abiotic and Biotic Stresses in Mango (Mangifera indica). *Plants (Basel)* 11, 3141. doi: 10.3390/plants11223141

Zhang, T., Zhou, L., Pu, Y., Tang, Y., Liu, J., Yang, L., et al. (2024). A chromosome-level genome reveals genome evolution and molecular basis of anthraquinone biosynthesis in *Rheum palmatum*. *BMC Plant Biol.* 24, 261. doi: 10.1186/s12870-024-04972-2

Zhao, H., He, Y., Zhang, K., Li, S., Chen, Y., He, M., et al. (2023). Rewiring of the seed metabolome during Tartary buckwheat domestication. *Plant Biotechnol. J.* 21, 150–164. doi: 10.1111/pbi.13932

Zhao, S., Xiong, F., Li, J., Ye, Z., Wang, L., Wang, T., et al. (2024a). Metabolomic characteristics and anthraquinones accumulation patterns of Rhubarb in different tissues and roots from different developmental stages. *Food Biosci.* 62, 105426. doi: 10.1016/j.fbio.2024.105426

Zhao, S., Xiong, F., Wang, L., Wang, B., Chen, K., Chen, C., et al. (2023). Study on the quality characteristics and geographical origin authentication of wild *Rheum tanguticum* in three authentic regions. *J. Food Composition Anal.* 123, 105463. doi: 10.1016/j.jfca.2023.105463

Zhu, J., Wang, Y., Wang, Q., Li, B., Wang, X., Zhou, X., et al. (2023). The combination of DNA methylation and positive regulation of anthocyanin biosynthesis by MYB and bHLH transcription factors contributes to the petal blotch formation in Xibei tree peony. *Hortic. Res.* 10, uhad100. doi: 10.1093/hr/uhad100

Zhuang, T., Gu, X., Zhou, N., Ding, L., Yang, L., and Zhou., M. (2020). Hepatoprotection and hepatotoxicity of Chinese herb Rhubarb (Dahuang): how to properly control the "General (Jiang Jun)" in Chinese medical herb. *Biomed. Pharmacother.* 127, Article 110224. doi: 10.1016/j.biopha.2020.110224

Frontiers in Plant Science

TYPE Original Research
PUBLISHED 16 July 2025
DOI 10.3389/fpls.2025.1586984

EDITED BY
Zongxia Yu,
Lushan Botanical Garden (CAS), China

REVIEWED BY
Chunmao Yuan,
Key Laboratory of Chemistry for Natural Products of Guizhou Province (CAS), China
Qiang Fei,
Guiyang University, China

*CORRESPONDENCE
Xiaoping Qin
qxp99@163.com
Xiao Ding
dingxiao@mail.kib.ac.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 03 March 2025
ACCEPTED 12 June 2025
PUBLISHED 16 July 2025

CITATION
Liu Y, Li J, Wu G, Wu X, Zhao Y, Ding X and Qin X (2025) Fungal inhibitory activity of sesquiterpenoids isolated from *Laggera pterodonta*.
Front. Plant Sci. 16:1586984.
doi: 10.3389/fpls.2025.1586984

Fungal inhibitory activity of sesquiterpenoids isolated from *Laggera pterodonta*

Yuxuan Liu[1†], Jinliang Li[2†], Guoxing Wu[1], Xiaoyun Wu[1], Yuhan Zhao[3], Xiao Ding[3*] and Xiaoping Qin[1*]

[1]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, College of Plant Protection, Yunnan Agricultural University, Kunming, China, [2]Testing Center, Yunnan Dehong Institute of Tropical Agricultural Science, Ruili, China, [3]State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

The phytochemical diversity of plants represents a valuable reservoir for novel agrochemical agents. Our preliminary investigations identified pronounced antifungal efficacy in *Laggera pterodonta* extracts, although this species remains critically understudied. Herein, six eudesmane-type sesquiterpenes (1–6) were isolated from *L. pterodonta* and structurally characterized through spectroscopic analysis. Additionally, the antifungal activity of these compounds against six plant-pathogenic fungal species was evaluated: *Phytophthora nicotianae*, *Fusarium oxysporum*, *Alternaria alternata*, *Gloeosporium fructigenum* Berk, *Colletotrichum fructicola*, and *Botrytis cinerea*. The results showed that the six compounds exhibited varying degrees of inhibitory effects on the six plant pathogens. Compound 1 showed the strongest antifungal effect in a dose-dependent way, with half-maximal effective concentration (EC_{50}) values of 12.56, 51.29, and 47.86 μg/mL against *P. nicotianae*, *F. oxysporum*, and *G. fructigenum* Berk, respectively. Compound 3 and compound 6 also showed notable inhibitory effects against *F. oxysporum* and *P. nicotianae* at 100 μg/mL, with inhibition rates of 58.82% and 73.92%, respectively. The minimum inhibitory concentrations (MICs) of compound 1 against *P. nicotianae* and *F. oxysporum* were 200 and 400 μg/mL, respectively. Mechanistic analysis revealed that compound 1 induced pronounced ultrastructural deformations in *P. nicotianae* and *F. oxysporum*, compromising membrane integrity and elevating permeability in both pathogens. Notably, the three bioactive compounds exhibited favorable ADMET (absorption, distribution, metabolism, excretion, and toxicity) profiles, demonstrating promising candidacy as novel herbicidal agents. These findings underscore their potential to advance phytogenic fungicide discovery.

KEYWORDS

Laggera pterodonta*, antibacterial activity, eudesmane-type sesquiterpenes, *Phytophthora nicotianae*, *Fusarium oxysporum

1 Introduction

The advent of chemical pesticides revolutionized agricultural pest, disease, and weed management, substantially improving crop productivity. However, their prolonged use has precipitated the "3R" (resistance, residue, and resurgence) dilemma, along with pervasive environmental contamination, posing significant challenges to sustainable societal development. Consequently, plant-derived pesticides that are characterized by low mammalian toxicity, rapid degradability, high efficacy, and environmental compatibility have garnered increasing attention. The exploration and application of such eco-friendly agrochemicals now represent a priority research area in modern plant protection science (Yang et al., 2022; Prastiwi et al., 2023; Saravana, 2022). Numerous plant-derived active ingredients, including pyrethrins, picloram, neem, anethole, tea saponin, and nicotine, have been researched for their pesticidal potential, have been proven to be potentially bioactive against several pests, and have been used in agricultural applications. Moreover, their safety has been widely recognized and has been registered and commercialized successfully in several countries.

A growing interest in plant-derived fungicides has catalyzed the identification of approximately 1,400 plant species exhibiting antimicrobial (bacteriostatic and bactericidal) properties. These species biosynthesize a diverse array of antimicrobial phytochemicals, including terpenoids, alkaloids, flavonoids, phenolic acids, and other structurally distinct classes of secondary metabolites (Bhandari et al., 2021). For example, four terpenoids isolated from plants of the genus *Aromatica* showed good inhibitory effects on a variety of plant pathogens, and the inhibitory effects were selective and complementary (Ding et al., 2017). The phytochemical constituents of *Sophora flavescens* Aiton, including kurarinol and trifolirhizin (in addition to the well-characterized matrine), exhibit broad-spectrum inhibitory activity against various phytopathogenic bacteria (Wang et al., 2020). Similarly, *Salix myrtillacea* seed essential oil demonstrates potent fumigant efficacy against multiple plant pathogens (Liu et al., 2022). However, compared to the extensive exploration of plant-derived insecticides, phytogenic fungicides remain structurally homogeneous and inadequately studied, representing a promising yet underexplored frontier for agrochemical innovation (Zhang et al., 2023).

Laggera pterodonta (DC.) Benth., a member of the Asteraceae family (genus Laggera), has been employed in traditional Chinese medicine for centuries owing to its well-documented pharmacological properties, including antioxidant, antitumor, antimicrobial, and analgesic activities (Wang, 2019; Li et al., 2023; Xie et al., 2021). The chemical composition of *L. pterodonta* is dominated by flavonoids, eucalyptane-type sesquiterpenoids, and their derivatives, in addition to coumarins, triterpenoids, fatty acids, and volatile oil components (Lu et al., 2014; Wei et al., 2017; Chen et al., 2012).

Currently, extensive research has elucidated the diverse pharmacological activities of *L. pterodonta*. For instance, sesquiterpenes and flavonoids isolated from this species exhibit significant anti-inflammatory properties (Xie et al., 2021). Similarly, its phytochemical constituents, including chrysosplenetin and absinthin, demonstrate notable antioxidant activity (Li et al., 2013). Notably, the flavonoid 5,7,3′,4′-tetramethoxy-3-hydroxyflavone displayed potent antitumor efficacy *in vitro*, as validated by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Cao et al., 2011). In contrast, the agricultural applications of *L. pterodonta* remain underexplored. Preliminary studies have indicated that its ethanol extract exhibits contact toxicity against aphids (*Aphis gossypii*) and East Asian locusts (*Locusta migratoria manilensis*) (Gao et al., 2019), while petroleum ether extracts showed inhibitory effects against *Colletotrichum gloeosporioides*, the causative agent of oil tea anthracnose (Luo et al., 2022). Building on prior findings that the ethyl acetate extracts of *L. pterodonta* possess robust antimicrobial activity, we initiated a phytochemical investigation targeting eudesmane-type sesquiterpenes to identify novel antimicrobial agents.

In this study, six sesquiterpenoids were isolated from *L. pterodonta* and structurally characterized (Figure 1), with their antifungal efficacy evaluated against a panel of phytopathogenic fungi. To elucidate the mechanistic basis of their inhibitory activity, the most bioactive candidates were subjected to preliminary mechanistic investigation via scanning electron microscopy (SEM) and propidium iodide (PI) fluorescence staining assays. Furthermore, ADMET (absorption, distribution, metabolism, excretion, and toxicity) predictions for these compounds were conducted via the admetSAR platform to prioritize those candidates with the most potential to be developed as novel, plant-derived fungicides or as lead compounds for structure–activity optimization. This approach aims to establish both a phytochemical scaffold and an empirical foundation for advancing the discovery of next-generation agrochemicals.

2 Materials and methods

2.1 General experimental procedures

Silica gel (60–80, 200–300, and 300–400 mesh) and thin-layer chromatography (TLC) plates were purchased from Qingdao Ocean Chemical Co., Ltd. (Qingdao, China). MCI gel (75–150 μm) was obtained from Chengdu Scientific Biochemical Co. (Chengdu, China), and Sephadex LH-20 gel (40–70 μm) was sourced from Pharmacia (USA). Semi-preparative high-performance liquid chromatography (HPLC, Shimadzu, Kyoto Prefecture, Japan) was performed using ultrapure water (Milli-Q purification Systems, Millipore, Burlington, MA, USA) and HPLC-grade acetonitrile (J.T. Baker, Phillipsburg, NJ, USA), with a YMC-Pack ODS-A column (5 μm, 10 × 250 mm; Shenzhen Kemis Technology Co., Shenzhen, China). Reagent-grade solvents, including petroleum ether, ethanol, acetone, ethyl acetate, methanol (MeOH), and chloroform, were supplied by Qingdao Ocean Chemical Co., Ltd. (China).

NMR spectra were measured using Bruker 500 and 600 MHz spectrometers while using Tetramethylsilane (TMS) as the endogenous reference. Infrared spectra of KBr particles were

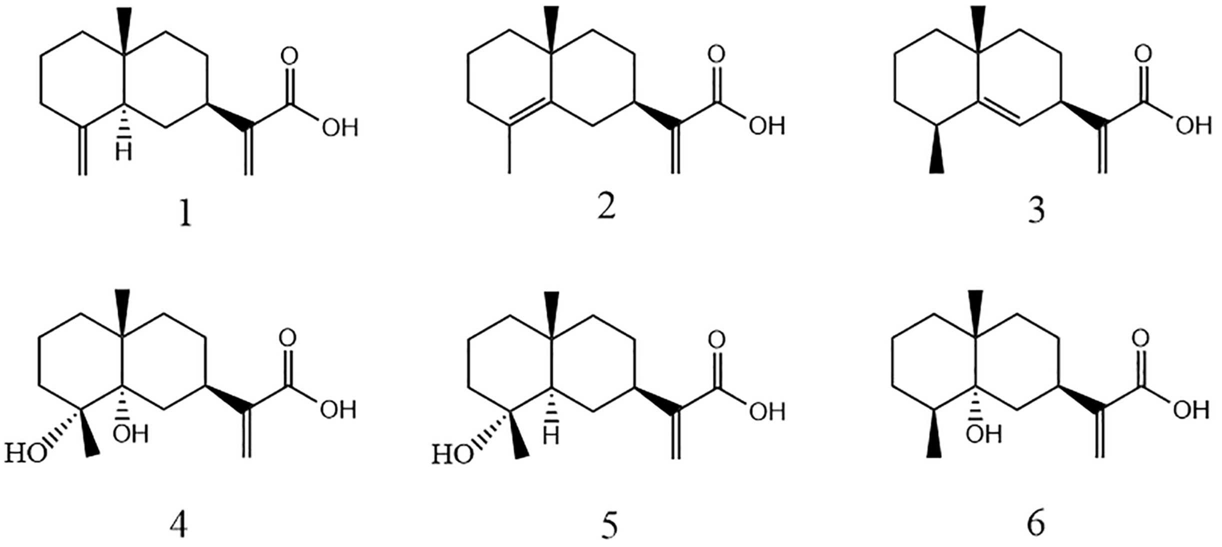


FIGURE 1
Chemical structures of the six compounds isolated from *Laggera pterodonta*.

measured using a BioRad FTS-135 spectrometer; rotameters were analyzed using a JASCO P-1020 digital polarimeter; and UV spectra were recorded using a Shimadzu UV-2401a. High-Resolution Electrospray Ionization Mass Spectrometry (HR-ESI-MS) was recorded on a triple quadrupole mass spectrometer (Agilent, USA).

2.2 Plant material

The above-ground parts of *L. pterodonta* were collected in July 2017 from the Baoshan countryside (25°C5′N, 99°C6′E), Yunnan Province, China. Prof. Hua Peng from the Kunming Institute of Botany, Chinese Academy of Sciences (CAS), identified the collected samples (via no. 1707016).

2.3 Extraction and isolation

Field-collected *L. pterodonta* (5 kg) was air-dried, mechanically pulverized, and exhaustively extracted with methanol (3 × 8 L) at ambient temperature. The combined filtrate was concentrated under reduced pressure to yield a crude extract (8 L), which was subsequently partitioned sequentially with petroleum ether and ethyl acetate (1:1, v/v; triplicate) to produce a petroleum ether-soluble fraction (68 g) and an ethyl acetate-soluble fraction (55 g). The ethyl acetate fraction was fractionated via silica gel column chromatography (10 × 100 cm) using a stepwise gradient elution of petroleum ether–acetone (100:1 to 1:1, v/v), yielding eight sub-fractions (Fr.1–Fr.8).

The Fr.4 (45.6 g) fraction was first purified using MCI column chromatography with a mobile phase of 10%–90% methanol–water in gradient elution to yield eight fractions (4A–4H). Fraction 4D was further eluted using forward silica gel column chromatography (mobile phase petroleum ether–ethyl acetate = 40:1–4:1) with gradient elution to yield five fractions (4D1–4D5). Fr.4D2 (3.9 g) was further purified using Sephadex LH-20 gel column chromatography (eluent was a mixture of methanol–dichloromethane = 6:4) and then again purified using semi-preparative HPLC [liquid phase conditions: the column was a YMC-pack ODS-A (5 mm, 10×250 mm), and the mobile phase was 58% acetonitrile in water at a flow rate of 3 mL/min, resulting in compounds 1 (100 mg, t_R = 59 min)), 2 (460 mg, t_R = 59 min), and 3 (135 mg, t_R = 55 min)]. Similarly, Fr.4F (2.1 g) was purified using Sephadex LH-20 gel column chromatography under methanol elution, followed by semi-preparative HPLC with a mobile phase of 47% acetonitrile in water, at a flow rate of 3 mL/min, to obtain compound 6 (1.9 g, t_R = 19 min).

Fr.5 (35.3 g) was purified using MCI column chromatography (gradient elution with 10%–90% methanol-water) to yield nine fractions (5A–5I). Fraction 5C was further purified using Sephadex LH-20 gel column chromatography (eluent was a mixture of methanol–dichloromethane = 6:4), followed by gradient elution using forward silica gel column chromatography (mobile phase petroleum ether–ethyl acetate = 40:1–1:1) to yield four fractions (5C1–5C4). After purifying Fr.5C2 (1.6 g) again using Sephadex LH-20 gel column chromatography (eluent was a mixture of methanol–dichloromethane = 6:4), it was purified by semi-preparative HPLC using 58% acetonitrile in water as the mobile phase at a flow rate of 3 mL/min to yield compounds 4 (67 mg, t_R = 27 min) and **5** (350 mg, t_R = 15 min).

2.4 Fungus material

Phytophthora nicotianae, *Fusarium oxysporum*, *Gibberella fructigenum* Berk, *Alternaria alternata*, *Colletotrichum fructicola*, and *Botrytis cinerea* strains were sourced from the Bio-Pesticide Development and Utilization Laboratory, Yunnan Agricultural University (Yunnan, China).

2.5 Screening of compounds for antifungal activity

The effect of the compounds on the mycelial growth of the pathogenic fungal strains was assessed using the mycelial growth rate method (Wang and Li, 2022). The compounds were dissolved in acetone to prepare 10 mg/mL stock solutions. Aseptically, 0.5 mL of each solution was added to 49.5 mL of sterile potato dextrose agar (PDA), yielding a drug-supplemented medium at a final concentration of 100 μg/mL. The medium was poured into Petri dishes, and 5-mm-diameter mycelial agar plugs (harvested from actively growing fungal colonies) were centrally inoculated onto each plate to standardize mycelial growth vigor. A solvent control containing an equivalent volume of acetone was included, and all treatments were performed in triplicate. The plates were then incubated at 27°C until the blank control filled the entire plate. Then, the diameter of the colonies was measured using the method of crossover, and the inhibition rate was calculated accordingly.

$$\text{Inhibition}\,(\%) = (\text{OD}_{\text{CK}} - \text{OD Compound}/\text{OD}_{\text{CK}} - \text{OD Fungus Cake}) \times 100\,\%$$

When the compounds showed good inhibitory activity against the tested fungal strains, the master compound solution was diluted into five concentration gradients of compound solutions (10, 5, 2.5, 1.25, and 0.625 mg/mL). Then, the drug-containing medium was prepared at concentrations of 100, 50, 25, 12.5, and 6.25 μg/mL. The rest of the operation was performed as described above. When the blank control grew all over the plate, the colony diameter was measured, and the inhibition rate was calculated using the crossover method. The experimental data were processed and statistically analyzed using Microsoft Excel 2011 and SPSS 20.0 software, with the inhibition rate converted to the corresponding chance value to obtain the virulence regression equation and to find the half-maximal effective concentration (EC_{50}) value.

2.6 Determination of MIC of compounds against *P. nicotianae* and *F. oxysporum*

Based on preliminary antifungal activity screening, compound 1—exhibiting broad-spectrum antifungal efficacy—was selected for further analysis. Its minimum inhibitory concentration (MIC) against *P. nicotianae* and *F. oxysporum* was assessed via the 96-well microdilution broth method (Isela et al., 2015). Prior to experimentation, *F. oxysporum* was maintained on PDA under standard conditions for a 7-day incubation period. Later, the spores were rinsed with sterile water and filtered through sterile gauze to remove the mycelium. The spores were then counted using a hemacytometer, and a spore suspension was prepared with a concentration of 1×10^6 CFU/mL. *P. nicotianae* was first treated with 0.1% potassium nitrate to induce spore production, and then spores were counted as described above to make a spore suspension with a concentration of 1×10^6 CFU/mL. Under aseptic conditions, 100 μL of a mixture of compounds and PDA and 100 μL of freshly prepared spore suspension (10^6 spores/mL) were taken from a sterile 96-well plate and shaken gently to mix well so that the effective concentrations of the compounds in each column of the treatment group were maintained at 12.5, 25, 50, 100, 200, 400, and 800 μg/mL. At the same time, the culture medium was set as a blank control, and three replicates were set for each treatment. The plates were capped, sealed, and incubated at a constant temperature of 28°C. The MIC was designated as the lowest drug concentration achieving complete inhibition of fungal growth, corresponding to an absence of visible fungal proliferation in test wells relative to the untreated control.

2.7 Effect of compounds on mycelial morphology

The effect of compound 1 on the mycelial morphology of *P. nicotianae* and *F. oxysporum* was first observed using the coating plate method. A spore suspension of 50 μL was aspirated and coated on the plates containing the drug medium (the concentration of the compound was the EC_{50} for inhibiting fungal growth), and a plate without the compound was used as the control group, which was incubated in an incubator with a constant temperature and humidity of 28°C, and then observed under a microscope, photographed, and recorded in a timely manner for data collection.

SEM was employed to assess ultrastructural alterations in fungal morphology. *P. nicotianae* and *F. oxysporum* were pre-cultured on PDA, and 5-mm mycelial plugs were excised from colony margins. For treatment groups, plugs were transferred to a drug-supplemented medium, while a drug-free medium served as the control. Upon full mycelial colonization in the control group (CK), samples from both groups were sectioned, mounted on stubs, and flash-frozen in super-cooled liquid nitrogen (−196°C) for 2 min. Specimens were subsequently sublimated in a preparation chamber at −140°C for 15 min, followed by dual gold sputter coating (60 s each). Morphological analysis was conducted using a cryo-scanning electron microscope (Sigma 300, Zeiss, Jena, Germany) under high-vacuum conditions.

2.8 Cell membrane integrity assay

PI is unable to penetrate intact living cell membranes, but it can enter the cells through damaged cell membranes and bind to DNA to produce fluorescence. Therefore, the extent of damage to cell membranes by compounds was further evaluated in *P. nicotianae* and *F. oxysporum* after treatment with PI-stained compounds. Following the method of Bo et al. (2014), the drug-treated fungus was washed twice with Phosphate Buffered Saline (PBS) solution and gently transferred to a slide. Then, 50 μL of PI dye (1 g/L) was added dropwise, and stained in the dark for 15 min. Excess dye was washed away with PBS, covered with a coverslip, and observed under an inverted fluorescence microscope with 535 nm as the excitation wavelength and 617 nm as the emission wavelength.

2.9 ADMET prediction of active compounds

The Simplified Molecular Input Line Entry System (SMILES) notations of the bioactive compounds were submitted to the ADMET prediction module of the admetSAR platform (http://lmmd.ecust.edu.cn/admetsar1), with pharmacokinetic profiles generated using an automated algorithm execution.

3 Results and discussion

3.1 Structural identification of compounds

3.1.1 Compound 1

Costic acid, colorless oil; the ^{13}C-NMR and DEPT of the compound showed 15 carbon signals as one methyl, eight methylene, two hypomethyl, and four quaternary carbons; the molecular formula of the compound was determined to be $C_{15}H_{22}O_2$ by combining with ESI-MS; the molecular weight of the compound was 234 (m/z: 233 $[M-H]^-$). The NMR data were as follows: ^{1}H-NMR (500 MHz, $CDCl_3$) d: 2.3 (1H, m, H-1), 2.01 (1H, m, H-1), 1.58 (1H, m, H-2), 1.52 (1H, m, H-2), 1.59 (1H, m, H-3), 1.34 (1H, m, H-3), 1.89 (1H, d, J = 11. 3 Hz, H-5), 1.66 (1H, m, H-6), 1.22 (1H, m, H-6), 2.53 (1H, m, H-7), 1.61 (1H, m, H-8), 1.46 (1H, m, H-8), 1.59 (1H, m, H-9), 1.28 (1H, m, H-9), 6.31 (1H, s, H-13), 5. 68 (1H, s, H-13), 0.55 (3H, s, H-14), 4.40 (1H, d, J = 1.5 Hz, H-15), and 4.70 (1H, d, J = 1.5 Hz, H-15). ^{13}C-NMR (150 MHz, $CDCl_3$) d: 36.8 (C-1), 23.4 (C-2), 41. 8 (C-3), 150.7 (C-4), 49.8 (C-5), 29.9 (C-6), 39.3 (C-7), 27.3 (C-8), 41.0 (C-9), 36.8 (C-10), 145.3 (C-11), 172.5 (C-12), 124.8 (C-13), 16.4 (C-14), and 105.5 (C-15). The data were in general agreement with those of the reference (Kalliopi et al., 2017), which determined the structure of the compound.

3.1.2 Compound 2

Isocostic acid, colorless oil, with the molecular formula $C_{15}H_{22}O_2$ and a molecular weight of 234 (m/z: 233 $[M-H]^-$) as determined by the ^{13}C-NMR and ESI-MS of the compound. The NMR data were as follows: ^{1}H-NMR (500 MHz, CD_3OD) d: 6.15 (1H, d, J = 1.1 Hz, H-13), 5.59 (1H, s, H-13), 1.1 (3H, s, H-15), and 1.62 (3H, s, H-14). ^{13}C-NMR (500 MHz, CD_3OD) δ: 41.4 (C-1), 20.1 (C-2), 34.1 (C-3), 126.0 (C-4), 135.6 (C-5), 32.5 (C-6), 41.9 (C-7), 29.0 (C-8), 43.4 (C-9), 35.5 (C-10), 147.6 (C-11), 170.64 (C-12), 122.8 (C-13), 19.4 (C-14), and 25.0 (C-15). The NMR spectral data of the compound were in general agreement with literature comparisons (Cruz and Martinez, 1982), and the structure of the compound was determined.

3.1.3 Compound 3

Eudesma-5,12-dien-13-oic acid, a colorless oily substance with the molecular formula $C_{15}H_{22}O_2$ and a molecular weight of 234 (m/z: 233 $[M-H]^-$) as determined by the ^{13}C-NMR and ESI-MS of the compound. The NMR data were as follows: ^{1}H-NMR (500 MHz, CD_3OD) d: 1.21 (1H, m, H-1), 1.52 (1H, m, H-1), 1.85 (1H, m, H-2), 1.42 (1H, m, H-2), 1.55 (1H, m, H-3), 1.57 (1H, m, H-3), 2. 45 (1H, m, H-4), 5.19 (1H, s, H-6), 3.30 (1H, m, H-7), 1.92 (1H, m, H-8), 1.42 (1H, m, H-8), 1.44 (2H, m, H-9), 6.10 (1H, m, H-12), 5.52 (1H, m, H-12), 1.17 (3H, s, H-14), and 1.19 (3H, d, J = 7.6 Hz H-15). ^{13}C-NMR (125 MHz, CD_3OD) d: 43.15 (C-1), 18.54 (C-2), 34.4 (C-3), 39.67 (C-4), 149.63 (C-5), 124. 78 (C-6), 39.89 (C-7), 27.76 (C-8), 42.87 (C-9), 35.52 (C-10), 148.1 (C-11), 170.95 (C-12), 123.20 (C-13), 27.81 (C-14), and 23.57 (C-15). The NMR spectral data of the compound were in general agreement with literature comparisons (Ying et al., 2006), so the structure of the compound was determined.

3.1.4 Compound 4

4α,5α-Dihydroxyeudesma-11(13)-en-12-oic acid, white acicular crystals. The molecular formula of the compound was determined to be $C_{15}H_{24}O_4$, and its molecular weight was 268 (m/z: 267 $[M-H]^-$) by ^{13}C-NMR and ESI-MS. The NMR spectral data were as follows: ^{1}H-NMR (500 MHz, $CDCl_3$) d: 1.86 (1H, m, H-1), 1.43 (1H, m, H-1), 1.90 (1H, m, H-2), 1.84 (1H, m, H-2), 1.88 (1H, m, H-3), 1.40 (1H, m, H-3), 1.62 (1H, m H-6), 1.61 (1H, m, H-6), 2.95 (1H, m, H-7), 1.02 (1H, m, H-8), 1.70 (1H, m, H-8), 1.10 (1H, m, H-9), 0.1.02 (1H, m, H-9), 5.71 (1H, bs, H-13), 6.31 (1H, s, H-13), 1.20 (3H, s, H-14), and 1.28 (3H, s, H-15).^{13}C-NMR (125 MHz, $CDCl_3$) d: 17.52 (C-1), 29.70 (C-2), 36.32 (C-3), 75.16 (C-4), 76.06 (C-5), 26.20 (C-6), 34.41 (C-), 31.25 (C-7), 31.25 (C-5), 26.20 (C-6), 34.41 (C-), 31.25 (C-7), 31.25 (C-7), 31.25 (C-), 31.25 (C-) 7), 31.25 (C-8), 37.69 (C-9), 36.83 (C-10), 144.80 (C-11), 171.53 (C-12), 125.28 (C-13), 25.88 (C-14), and 22.20 (C-15). The NMR spectral data of the compounds were in general agreement with literature comparisons (Ying et al., 2006), and the structure of the compound was determined.

3.1.5 Compound 5

Ilicic acid, white massive crystals; the ^{13}C-NMR and DEPT combined with ESI-MS of the compound determined its molecular formula to be $C_{15}H_{24}O_3$ and its molecular weight to be 252 (m/z: 253 $[M+H]^+$). The NMR spectral data were as follows: ^{1}H NMR (500 MHz, $CDCl_3$) d: 0.89 (3H, s, H-14), 1.11 (3H, s, H-15), 1.25 (1H, m, H-6), 1.31~1.66 (2H, m, H-1), 1.31~1.66 (2H, m, H-8), 1.31~1.66 (2H, m, H-9), 1.31~1.66 (2H, m, H-8), 1.31~1.66 (2H m, H-9), 1.31~1.66 (2H, m, H-2), 1.81 (1H, m, H-3), 1.50 (1H, m, H-5), 1.89 (1H, d, J = 12.0 Hz, H-6), 2.50 (1H, m, H-7), 6.25 (1H, s, H-12), and 5.62 (1H, s, H-12). ^{13}C-NMR (125 MHz, $CDCl_3$) δ: 18.8 (C-15), 20.1 (C-2), 22.4 (C-14), 26.8 (C-8), 27.0 (C-6), 34.7 (C-10), 40.0 (C-7), 40.9 (C-9), 43.3 (C-1), 44.5 (C-3), 72.6 (C-4), 55.0 (C-5), 124.2 (C-12), 145.3 (C-11), and 171.6 (C-13). The NMR spectral data of the compound were in almost perfect agreement with literature comparisons (San et al., 1990), and hence, the structure was determined to be ilicic acid.

3.1.6 Compound 6

5α-Hydroxy-4,15-dihydrocostic acid, white acicular crystals; the ^{13}C-NMR and DEPT of the compound showed that the compound had 15 carbon signals of two methyls, two methylene groups, two hypomethyls, and four quaternary carbons, which was determined to have a molecular formula of $C_{15}H_{24}O_3$ by combining with ESI-MS and a molecular weight of 252 (m/z: 234 $[M-H20]^-$).

The NMR spectral data were as follows: ^{1}H-NMR (500 MHz, $CDCl_3$) d: 6.24 (1H, s, H-13), 5.63 (1H, s, H-13), 0.90 (3H, s, H-15), 1.08 (3H, s, H-14), 2.05 (1H, ddd, J = 12.0, 12.0, 5.0 Hz, H-3), 1.34 (1H, d, J = 12.0 Hz, H-3), 1.58 (1H, dq, J = 7.0, 5.0 Hz, H-4), 1.20 (1H, dd, J = 12.5, 4.0 Hz, H-6), 1.86 (1H, dd, J = 12.5, 12.5 Hz, H-6), 3.06 (1H, m, H-7), 1.63 (1H, m, H-8), 1.53 (1H, m, H-8), 1.57 (1H, m, H-9), and 1.50 (1H, m, H-9). ^{13}C-NMR (125 MHz, $CDCl_3$) δ: 38.2 (C-1), 26.2 (C-2), 34.7 (C-3), 41.1 (C-4), 75.9 (C-5), 28.0 (C-6), 34.7 (C-7), 17.0 (C-8), 37.9 (C-9), 36.7 (C-10), 145.8 (C-11), 172.5 (C-12), 124.3 (C-13), 16.7 (C-14), and 21.8 (C-15). The data were in general agreement with those of the reference (Xie et al., 2011), which determined the structure of the compound.

3.2 Screening of compounds for fungal inhibitory activity

The antifungal activity of the six compounds was evaluated using a fungal growth rate assay (Table 1). At a concentration of 100 μg/mL, all compounds exhibited varying inhibitory effects against the tested phytopathogenic fungi. Notably, the eudesmane-type sesquiterpene acid compound 1 demonstrated pronounced inhibitory efficacy against *P. nicotianae*, *F. oxysporum*, and *Gibberella fujikuroi* Berk, with inhibition rates of 94.64%, 72.55%, and 70.46%, respectively. In contrast, its activity against *A. alternata*, *B. cinerea*, and *C. fructicola* was marginal, yielding inhibition rates of 15%–36%.

Compound 3 suppressed *F. oxysporum* mycelial growth by 58.82% but showed limited efficacy against the other five strains (16%–34% inhibition). Similarly, compound 6 inhibited *P. nicotianae* growth by 73.92%, while its activity against the remaining fungi ranged from 10% to 33%. The remaining three compounds (2, 4, and 5) displayed negligible antifungal effects against all tested pathogens, with inhibition rates consistently below 40%.

Structurally, compounds 1, 2, and 3 are isomers of each other, but compound 1 exhibited a broader spectrum and more fungal inhibitory activity as compared to the others. The differences among the three are manifested in the different positions of the double bonds; therefore, we speculate that the changes in the positions of the double bonds have a greater influence on the fungal inhibitory activity of eucalyptus-type sesquiterpene acid compounds and that the terminal double bonds of the compounds should exhibit greater activity. Meanwhile, compounds 5 and 6 are isomers of each other, and the structural difference between them is only the hydroxyl linkage site, but the inhibitory effect on *P. nicotianae* showed a significant difference. Combined with the structure and fungal inhibitory activity of compound 4, it is evident that the attachment of hydroxyl groups to C-3 of eucalyptane-type sesquiterpenoids significantly reduces their inhibitory effect against *P. nicotianae*. However, further studies are warranted to elucidate the broad-spectrum efficacy of this inhibitory effect against various phytopathogens.

3.3 Determination of the EC_{50} of compound 1 against three plant pathogenic fungi

Based on the results of the preliminary screening of the antimicrobial activity of the eucalyptane-type sesquiterpenoid 1, the pathogenic fungi that achieved more than 70% inhibition were selected for the determination of virulence equations (Table 2; Figure 2). As shown in Table 2, the eudesmane-type sesquiterpenoid 1 exhibited pronounced antifungal efficacy against the three phytopathogenic fungi, *P. nicotianae*, *F. oxysporum*, and *G. fujikuroi* Berk, with EC_{50} values of 12.56, 51.29, and 47.86 μg/mL, respectively.

The results of the activity test showed that the eucalyptane-type sesquiterpenoid 1 had a good inhibitory effect against *P. nicotianae* (Figure 3), which shows its potential to be further developed into a novel plant-derived fungicide. Meanwhile, in order to explore the principle of the inhibition of phytopathogenic fungi by compound 1, we selected *P. nicotianae* and *F. oxysporum* as the target strains and further investigated the mechanism of action of eucalyptolide-type sesquiterpenoids in plant pathogenic fungi.

3.4 Determination of the MIC of compound 1 against *P. nicotianae* and *F. oxysporum*

In order to further understand the inhibitory effect of compound 1 on plant pathogenic fungi, the MIC of compound 1

TABLE 1 Inhibition rates (%) of the compounds against six fungal pathogens.

Compound	*Phytophthora nicotianae*	*Fusarium oxysporum*	*Alternaria. alternata*	*Gloeosporium fructigenum* Berk	*Colletotrichum fructicola*	*Botrytis cinerea*
1	94.64 ± 1.39	72.55 ± 1.96	34.77 ± 2.16	70.46 ± 0.69	15.69 ± 1.96	35.29 ± 1.96
2	32.58 ± 1.63	16.34 ± 1.13	12.09 ± 1.13	16.67 ± 0.98	15.69 ± 1.96	15.03 ± 1.13
4	33.53 ± 0.34	58.82 ± 1.96	16.60 ± 0.93	25.16 ± 2.47	17.65 ± 0.59	18.56 ± 0.82
4	38.56 ± 2.26	13.86 ± 0.30	16.34 ± 1.13	13.86 ± 0.63	18.56 ± 2.46	8.69 ± 0.74
5	34.90 ± 1.03	17.78 ± 1.94	16.41 ± 1.67	19.54 ± 0.96	31.90 ± 1.24	13.95 ± 1.53
6	73.92 ± 0.86	15.82 ± 1.25	17.65 ± 1.96	32.09 ± 1.28	11.76 ± 1.69	13.07 ± 1.13

TABLE 2 Inhibitory activity of compound 1 against the mycelial growth of three plant pathogens.

Plant pathogenic fungus	Toxicity regression equation	EC_{50} (μg/mL)	Correlation coefficient (R^2)
Phytophthora nicotianae	y = 1.713x + 3.1137	12.56	0.979
Fusarium oxysporum	y = 1.3822x + 2.6427	51.29	0.9277
Gloeosporium fructigenum Berk	y = 1.3297x + 2.7718	47.86	0.9581

The effects of the tested compounds on three plant pathogens were repeated in triplicate.

against *P. nicotianae* and *F. oxysporum* was determined using the broth dilution method in a 96-well plate; the results are presented in Table 3. The mycelial growth of *F. oxysporum* persisted at concentrations of 0–100 μg/mL, whereas the complete inhibition of *P. nicotianae* and *F. oxysporum* was observed at 200 and 400 μg/mL, respectively. Consequently, the MICs of compound 1 against *P. nicotianae* and *F. oxysporum* were determined as 200 and 400 μg/mL, respectively.

3.5 Effect of compound 1 on the mycelial morphology of *P. nicotianae* and *F. oxysporum*

The EC_{50} of compound 1 for mycelial growth inhibition was applied to treat *P. nicotianae* and *F. oxysporum*, with morphological alterations assessed via light microscopy (LM) and SEM. As depicted in Figure 4, pronounced morphological aberrations were evident in both fungi following 24-hour exposure to compound 1 under identical LM magnification. In the control group of *P. nicotianae*, the mycelium was observed to be uniform in thickness, the surface was smooth and full, and it had formed an obvious mycelial structure. The mycelium in the treatment group was noticeably twisted and non-uniform in thickness; the mycelium was shorter and more forked, the surface was rough, and the mycelium as a whole was distributed in a deformed manner. In the *F. oxysporum* control group, the hyphae exhibited uniform morphology with robust mycelial networks, whereas treated samples displayed attenuated and fragmented hyphae, failing to establish defined mycelial structures after 24 hours. Radial expansion in the treatment group was markedly reduced compared to the controls.

Figure 5 presents the SEM images of *P. nicotianae* and *F. oxysporum* mycelia treated with compound 1 at the EC_{50} concentration for mycelial growth inhibition. *P. nicotianae* control mycelium was found to be tubular, smooth, and homogeneous in morphology, and fuller, while after the action of compound 1, the surface of the fungus became rough, with white material attached, and there were deformities, such as collapse, obvious loss of water, crumpling, and bending and folding, which was in line with the phenomenon observed under the light microscope. The *F. oxysporum* control mycelium surface was also relatively smooth, and the mycelium and spores were very full, while in the compound **1** treatment, the most obvious changes were that the diameter of the mycelium became thinner and shorter, the spores became smaller, and the growth rate lagged. This phenomenon is also consistent with the phenomenon observed during light microscopy imaging.

3.6 Effect of compound 1 on the membrane integrity of fungal cells

PI, a membrane-impermeant nucleic acid stain, fluoresces upon binding to intracellular DNA/RNA, serving as a marker for

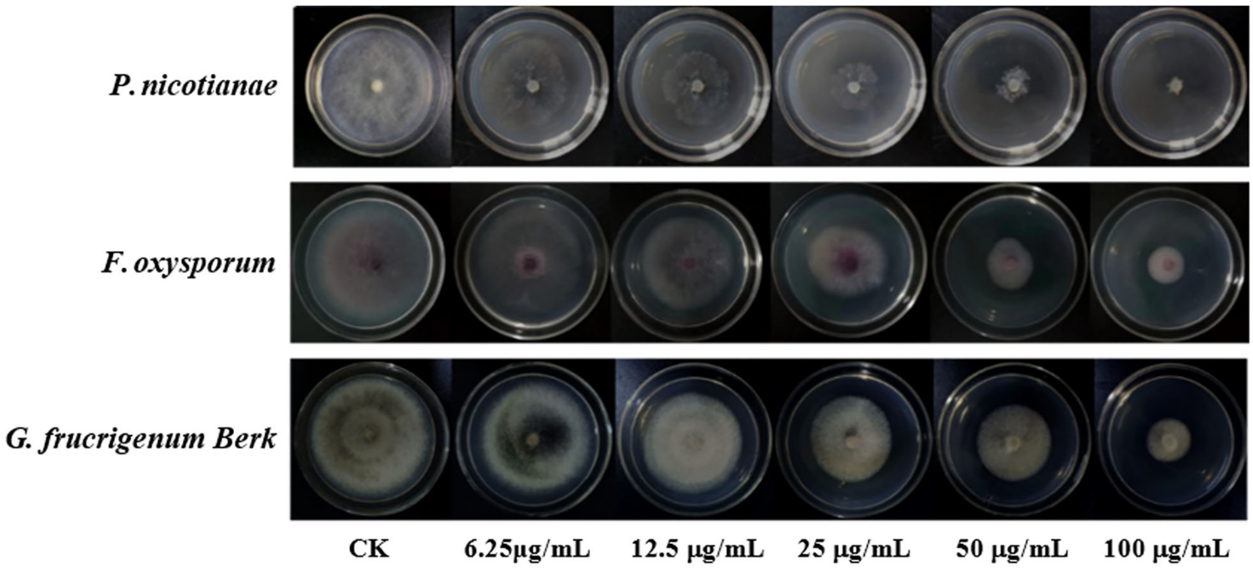


FIGURE 2
Inhibitory activity of compound 1 against mycelial growth of three plant pathogens.

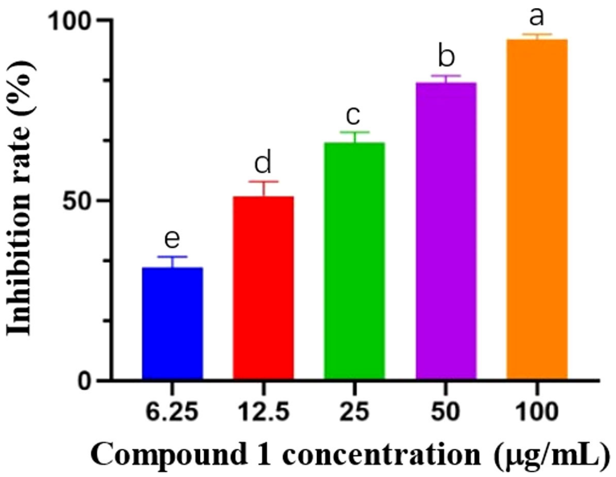


FIGURE 3
Inhibition of *Phytophthora nicotianae* by compound 1 isolated from *Laggera pterodonta*.

compromised membrane integrity. Late-stage apoptotic cells exhibit PI uptake, reflecting increased membrane permeability. As shown in Figure 6, no discernible PI fluorescence was observed in untreated controls under dark-field excitation. In contrast, compound 1-treated *P. nicotianae* and *F. oxysporum* exhibited pronounced red fluorescence, indicating membrane disruption and elevated permeability. These findings suggest that compound 1 induces structural damage to fungal cell membranes, likely mediating its antifungal activity. Collectively, the data support the hypothesis that the eudesmane-type sesquiterpenoid 1 exerts antifungal effects by compromising membrane integrity.

The current understanding of the antifungal mechanisms of plant-derived agents remains limited. Extant studies have suggested that these compounds primarily disrupt fungal physiology through three pathways: 1) compromising cell ultrastructure, 2) perturbing metabolic homeostasis, and 3) impairing bioenergetic processes (Zhao et al., 2022). For example, Li et al. (2017) found that tea tree oil could damage the cell wall and cell membrane of *B. cinerea* through increased membrane permeability, which ultimately affects the growth of *B. cinerea*. Furthermore, numerous studies have demonstrated that antifungal agents disrupt fungal cell walls and membrane integrity, inducing membrane hyperpermeability and thereby suppressing hyphal proliferation (Ou et al., 2019; He et al., 2018). Regarding the effects of fungal inhibitors on the cellular metabolism of fungal cells, Chen et al. (2018) found that curcumin could inhibit the activity of *Fusarium graminearum* Schw-related enzymes and affect tRNA synthesis and glucose metabolism using proteomics. Additionally, several reports have provided comprehensive information about the effect of bacteriostatic active substances on the cellular energy of fungi. For instance, emerging evidence demonstrates that curcumin perturbs both material metabolism and energy homeostasis in the phytopathogen *F. graminearum* Schw (Chen et al., 2018).

Based on the observed morphological aberrations and compromised membrane integrity in *P. nicotianae* and *F. oxysporum* mycelia treated with compound 1, we hypothesized that its antifungal activity involves dual mechanisms: 1) direct disruption of membrane architecture, increasing permeability and triggering leakage of essential metabolites, thereby altering hyphal morphology; and 2) interference with metabolic pathways critical for fungal proliferation, leading to growth arrest. While these findings provide preliminary insights, the inhibition mechanism of compound 1 remains incompletely resolved. Antifungal agents often exhibit multifactorial and polypharmacological modes of action, necessitating further investigation to delineate specific molecular targets and pathways. Nevertheless, the data collectively substantiate compound 1 as a promising candidate for the development of novel plant-derived fungicides, offering both a theoretical framework and methodological precedents for future studies.

3.7 ADMET prediction of active compounds

The ADMET property of a compound denotes the five key notes as absorption, distribution, metabolism, excretion, and toxicity of a compound in an organism (Deng et al., 2012). The main indicators are the blood–brain barrier (BBB), human intestinal absorption (HIA), Caco-2 permeability (CCP), Ames mutagenicity (ATT), carcinogenicity, and cytochrome (CYP2D6). ADMET prediction of compounds helps to evaluate the drug capability and safety of compounds in the early stages of drug development, which not only helps to increase the success rate of drug development but also reduces the cost of drug development.

Based on antifungal activity screening results, compounds 1, 3, and 6—exhibiting selective antifungal efficacy against key phytopathogens—were prioritized for ADMET profiling (Table 4). Computational predictions indicate favorable intestinal absorption and bioavailability for all three compounds, with

TABLE 3 The minimum inhibitory concentration of compound 1 against *Phytophthora nicotianae* and *Fusarium oxysporum*.

Concentration (μg/mL)	0	12.5	25	50	100	200	400	800
P. nicotianae	+++	+++	++	++	+	–	–	–
F. oxysporum	+++	+++	++	++	++	+	–	–

+++, a large amount of mycelium grows; +++, a medium amount of mycelium grows; +, a small amount of mycelium grows; –, no mycelium grows.

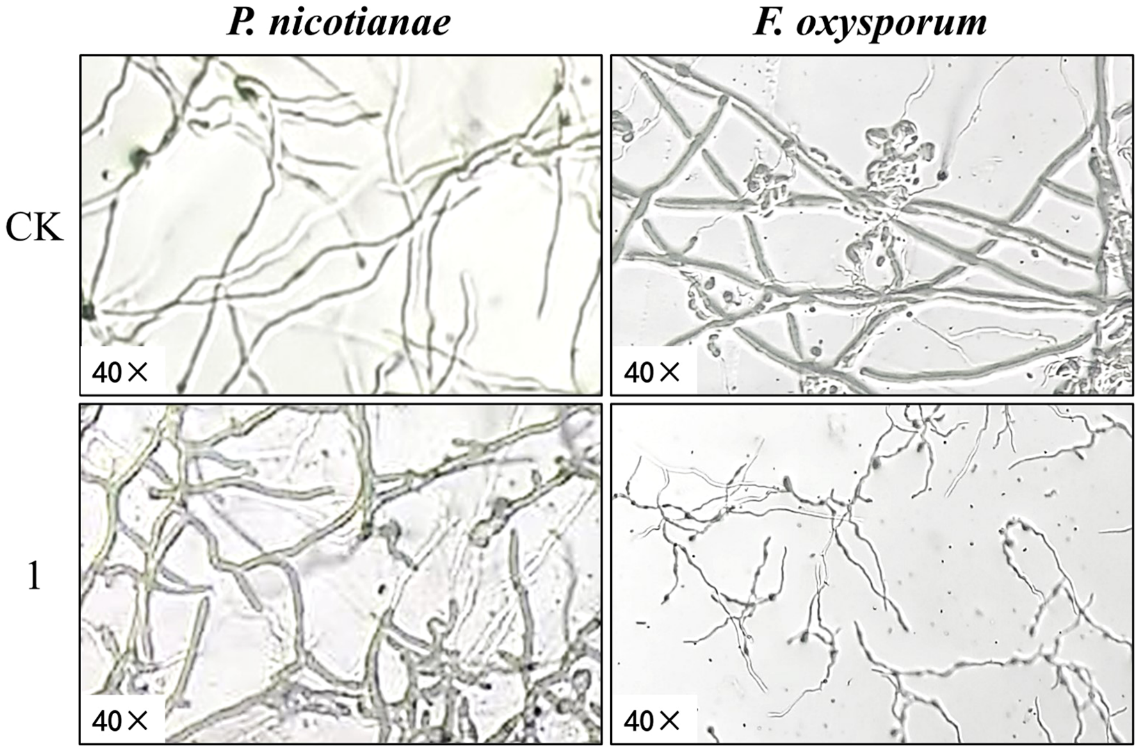


FIGURE 4
Effect of compound **1** on mycelium morphology of *Phytophthora nicotianae* and *Fusarium oxysporum* observed by light microscopy.

demonstrable permeability across intestinal epithelial barriers. Notably, none exhibited mutagenic or carcinogenic liabilities, suggesting low toxicological risks. In addition, as compared to compounds 1 and 3 (which were able to cross the blood–brain barrier), compound 6 did not easily cross the blood–brain barrier. However, the difference in their structures lies in the presence or absence of hydroxyl groups on the backbone, so we speculate that the presence or absence of backbone hydroxyl groups had a significant effect on the blood–brain barrier passage ability of eucalyptus-type sesquiterpenoids. These findings indicate that all three compounds have good ADMET properties and have characteristics that can be developed into novel drugs.

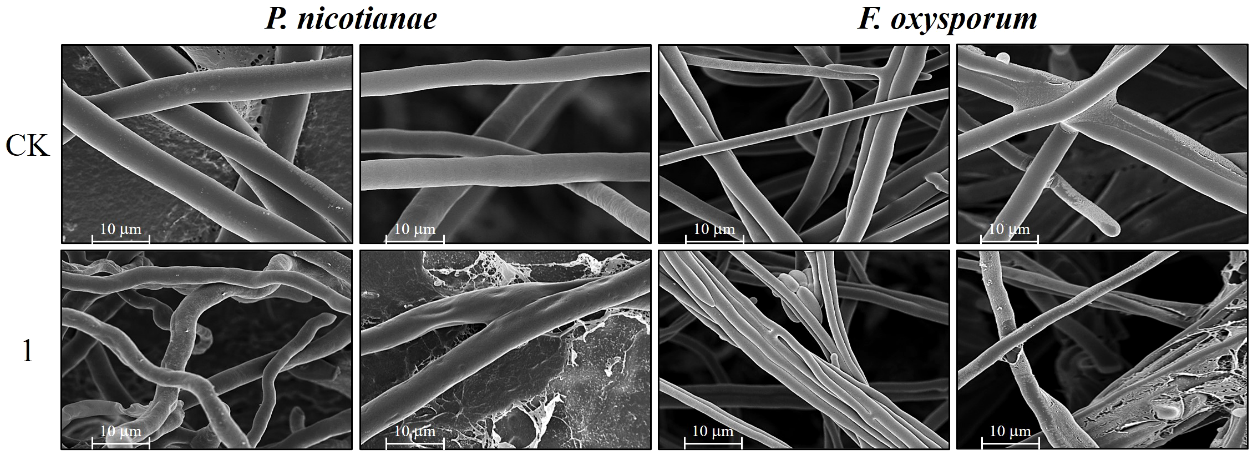


FIGURE 5
Scanning electron micrographs representing the mycelium of *Phytophthora nicotianae* and *Fusarium oxysporum* treated with compound 1.

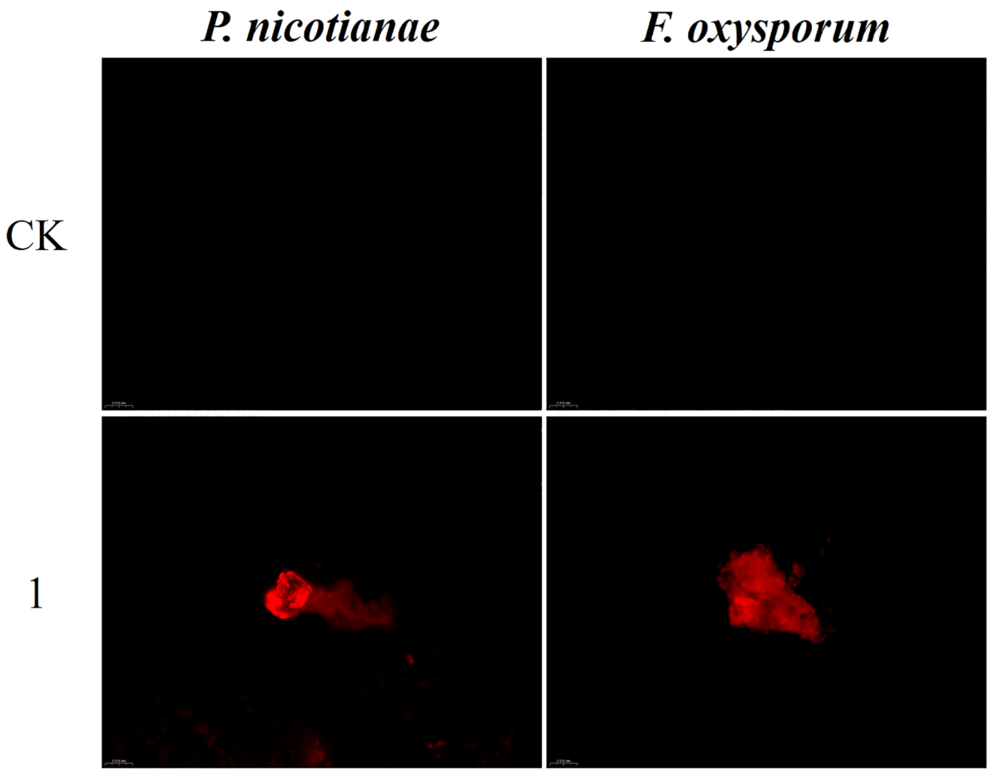


FIGURE 6
Effects of compound 1 on the cell membrane integrity of *Phytophthora nicotianae* and *Fusarium oxysporum*. CK represents the control.

TABLE 4 ADMET prediction results for compounds.

Compound	BBB	HIA	CCP	ATT	Carcinogenicity	CYP2D6 inhibitor
1	0.7250+	0.9950+	0.7249+	None	None	Non-inhibitor
3	0.7000+	0.9946+	0.8665+	None	None	Non-inhibitor
6	0.5000−	0.9956+	0.7609+	None	None	Non-inhibitor

BBB "+" in the table indicates that the drug molecules can easily pass through the blood–brain barrier. BBB "−" indicates that the drug molecules cannot easily pass through the blood–brain barrier; the closer the value is to 1, the better the permeability of the drug to the blood–brain barrier. HIA "+" indicates that the drug molecules can be absorbed or assimilated through human intestines; the closer the value is to 1, the better the absorption through the intestines. CCP "+" indicates that the drug molecules can easily penetrate the human intestinal cell line; the closer the value is to 1, the better the permeability of CCP. "None" indicates that the compound is not mutagenic or carcinogenic. ADMET, absorption, distribution, metabolism, excretion, and toxicity; BBB, blood–brain barrier; HIA, human intestinal absorption; CCP, Caco-2 permeability; ATT, Ames mutagenicity.

4 Conclusions

In this study, six eudesmane-type sesquiterpenoids isolated from *L. pterodonta* were evaluated for antifungal activity. The compounds exhibited varying inhibitory effects on six phytopathogens (*P. nicotianae*, *F. oxysporum*, *G. fujikuroi* Berk., *A. alternata*, *C. fructicola*, and *B. cinerea*). Compound 1 demonstrated potent inhibition against *P. nicotianae*, *F. oxysporum*, and *G. fujikuroi* (94.64%, 72.55%, and 70.46% inhibition, respectively), with EC_{50} values of 12.56, 51.29, and 47.86 μg/mL, respectively, and MICs of 200 and 400 μg/mL for *P. nicotianae* and *F. oxysporum*, respectively. Compound 3 inhibited *F. oxysporum* (58.82%), while compound 6 suppressed *P. nicotianae* (73.92%). Compound 1 induced severe morphological aberrations and compromised membrane integrity in *P. nicotianae* and *F. oxysporum*, elevating permeability and disrupting cellular homeostasis. All three bioactive compounds exhibited favorable ADMET profiles, underscoring their potential as eco-friendly, plant-derived fungicides, which may also serve as the basis for the development of new fungicides.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

YL: Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. JL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. GW: Funding acquisition, Project administration, Visualization, Writing – review & editing. XW: Methodology, Supervision, Writing – original draft. YZ: Methodology, Project administration, Resources, Validation, Writing – review & editing. XD: Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing. XQ: Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by the National Key R&D Program of China, 2024YFD1400600, and the Reserve Talents Project for Yunnan Young and Middle-aged Academic and Technical Leaders (202105AC160037), and the Yunnan Province Agricultural Joint Special Key Project (202301BD070001-141), and the National Natural Science Foundation of China (31660541), and the Yunnan Provincial Science, Technology Talents and Platform Program Project (202205AD160041).

Conflict of interest

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

Generative AI statement

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

Publisher’s note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1586984/full#supplementary-material

References

Bhandari, S., Yadav, P. K., and Sarhan, A. T. (2021). Botanical fungicides; current status, fungicidal properties and challenges for wide scale adoption: a review. *J. Rev. In Food And Agriculture.* 2, 63–68. doi: 10.26480/rfna.02.2021.63.68

Bo, T., Liu, M., Zhong, C., Zhang, Q., Su, Q. Z., Tan, Z. L., et al. (2014). Metabolomic analysis of antimicrobial mechanisms of ε-poly-L-lysine on Saccharomyces cerevisiae. *J. J. Agric. Food Chem.* 62, 4454–4465. doi: 10.1021/jf500505n

Cao, C. M., Shen, W. Z., Li, Y. L., Wang, H., and Gao, M. Y. (2011). Effects and mechanisms of 3,5-dihydroxy-6,7,3′,4′-tetramethoxyflavone from the traditional Chinese medicine Laggera pterodonta on apoptosis of human nasopharyngeal carcinoma CNE cells. *J. Adv. Biochem. Biophysics.* 38, 254–261. doi: 10.3724/SP.J.1206.2010.00539

Chen, G. Y., Chen, X. B., and Liu, G. M. (2012). Study on the chemical composition of Stinking Lingdan. *J. Anhui Agric. Science.* 40, 2022–2023. doi: 10.13989/j.cnki.0517-6611.2012.04.126

Chen, C., Long, L., Zhang, F., Chen, Q., Chen, C., Yu, X. R., et al. (2018). Antifungal activity, main active components and mechanism of Curcuma longa extract against Fusarium graminearum. *J. PloS One* 13, 194284. doi: 10.1371/journal.pone.0194284

Cruz, R., and Martinez, R. M. (1982). Stereoselective total synthesis of (±) – isocostic and (±)-3-Oxoisocostic acids. *Aust. J. Chem.* 35, 451–456. doi: 10.1071/CH9820451

Deng, Z. T., Qi, X., and Li, J. (2012). Research progress on high-throughput screening technology for *in vitro* ADMET of innovative drugs. *J. Adv. Modern Biomedicine.* 12, 1176–1178. doi: 10.13241/j.cnki.pmb.2012.06.051

Ding, L., Li, H. C., Wang, B. Q., and Liu, G. A. (2017). Bacteriostatic activities of the major secondary metabolites of four species of Isodon amethystoides from Gansu. *J. J. Northwest Normal University (Natural Sci. Edition).* 53, 82–87. doi: 10.16783/j.cnki.nwnuz.2017.02.015

Gao, T. T., Yang, Y. H., Yang, N., Li, B., Jin, H., Tao, K., et al. (2019). Extraction and isolation of locust-killing active components of Laggera pterodonta. *J. J. Sichuan University.* 56, 142–148. doi: 10.3969/j.issn.0490-6756.2019.01.025

He, J., Wu, D., Zhang, Q., Chen, H., Li, H. Y., Han, Q. H., et al. (2018). Efficacy and mechanism of cinnamon essential oil on inhibition of Colletotrichum acutatum isolated from ‘Hongyang’ Kiwifruit. *J. Front. Microbiol.* 9, 1288. doi: 10.3389/fmicb.2018.01288

Isela, S. F., Everardo, L. R., Georgina, R. L., Alejandrina, M. G. M., and Arturo, V. G. (2015). Influence of culture media on biofilm formation by Candida species and response of sessile cells to antifungals and oxidative stress. *J. BioMed. Res. Int.* 2015, 783639. doi: 10.1155/2015/783639

Kalliopi, S., Demosthenis, I., Apostolos, S., Anita, B., Athanassios, G., and Haralambos, E. (2017). Use of costic acid, a natural extract from Dittrichia viscosa, for the control of Varroa destructor, a parasite of the European honey bee. *J. Beilstein J. Organic Chem.* 13, 952–959. doi: 10.3762/bjoc.13.96

Liu, Q. J., Fu, L. X., Zhang, Y., Li, F. C., Yu, H., Li, J. L., et al. (2022). Fumigant activity and chemical composition analysis of essential oil of Dodonaea viscosa seeds against five plant pathogenic fungi. *J. South. J. Agriculture.* 53, 1935–1943. doi: 10.3969/j.issn.2095-1191.2022.07.016

Li, J. L., Li, F. C., Wu, G. X., Gui, F. R., Li, H. M., Xu, L. L., et al. (2023). Acetylcholinesterase inhibitory activity of sesquiterpenoids isolated from Laggera pterodonta. *J. Front. Plant Science.* 14, 1074184–1074184. doi: 10.3389/fpls.2023.1074184

Li, Y., Shao, X., Xu, J., Wei, Y. Y., and Xu, F. (2017). Tea tree oil exhibits antifungal activity against Botrytis cinerea by affecting mitochondria. *J. Food Chem.* 234, 62–67. doi: 10.1016/j.foodchem.2017.04.172

Li, S. H., Zhao, Q., Liu, F., and Lei, K. J. (2013). Identification and antioxidant activity of flavonoids in Laggera pterodonta. *J. Modern Food Sci. Technology.* 29, 1213–1216. doi: 10.13982/j.mfst.1673-9078.2013.06.013

Lu, P. W., Chen, L. J., and Li, W. (2014). Chemical constituents from Laggera pterodonta. *J. J. Chin. medicinal materials.* 37, 816–819. doi: 10.13863/j.issn1001-4454.2014.05.022

Luo, J. M., Zhan, X. Y., Wu, J. R., Liu, L., Hong, Y. T., Zhang, D. H., et al. (2022). Screening of plant extracts for fungus inhibitory activity against Colletotrichum gloeosporioides. *J. Chin. J. Biol. Control.* 38, 852–859. doi: 10.16409/j.cnki.2095-039x.2022.04.009

Ou, Y. Q. L., Duan, X. F., Li, L., and Tao, N. G. (2019). Cinnamaldehyde exerts its antifungal activity by disrupting the cell wall integrity of Geotrichumcitri-aurantii. *J. Front. Microbiol.* 10, 55. doi: 10.3389/fmicb.2019.00055

Prastiwi, S., Wagiyana, W. Y., and Alfarisy, F. K. (2023). Compatibility studies of entomopathogenic fungi and botanical pesticide for controlling *Spodoptera exigua.* J. *E3S Web Conferences.* 373, e100345. doi: 10.1051/e3sconf/202337307006

San, J. F., Castellano, G., and Marco, J. (1990). Sesquiterpene lactones from Artemisia herba-alba. *J. Phytochemistry.* 29, 541–545. doi: 10.1016/0031-9422(90)85114-U

Saravana, G. (2022). Plants and phytochemical activity as botanical pesticides for sustainable agricultural crop production in India-MiniReview. *J. J. Agric. Food Res.* 9, e100345. doi: 10.1016/J.JAFR.2022.100345

Wang, Y. X. (2019). Overview of the research on Laggera pterodonta, a commonly used folk medicine in Yunnan. *J. Yunnan J. Traditional Chin. Med.* 40, 74–75. doi: 10.16254/j.cnki.53-1120/r.2019.07.031

Wang, H., and Li, P. (2022). Determination of virulence of different fungicides against Venturia inaequalis. *J. Modern Pesticides.* 21, 55–57. doi: 10.3969/j.issn.1671-5284.2022.06.010

Wang, C. Y., Lu, C. S., and Chang, F. H. (2020). Research on the fungus inhibition of Alternaria panax Whetz by effective parts of Sophorae Flavescentis Radix. *J. South. Agriculture.* 14, 159–183. doi: 10.19415/j.cnki.1673-890x.2020.29.076

Wei, J. C., Chen, Y., Wei, Z. X., Ye, X., Que, Z. L., Yu, Y. H., et al. (2017). Progress of research on the pharmacology and quality analysis of hexandra chinensis. *J. Guangzhou Chem. Industry.* 45, 1–2. doi: 10.3969/j.issn.1001-9677.2017.02.001

Xie, Y. Q., Fan, W. T., Chen, X. Q., Li, R. T., and Zhang, Z. J. (2021). Chemical constituents from Laggera pterodonta. *J. Biochem. Systematics Ecology.* 94, 104222. doi: 10.1016/j.bse.2020.104222

Xie, W. D., Weng, C. W., Shen, T, and Gao, X. (2011). Sesquiterpenoids from aster himalaicus. *J. Chem. Nat. Compd+.* 47, 309–310. doi: 10.1007/s10600-011-9916-2

Yang, X. R., Liao, Y. F., Zhao, P. F., Qin, Y. M., and Ou, Y. X. H. (2022). Progress of research on plant-derived pesticides and their development and utilization. *J. South. Agriculture.* 16, 33–36. doi: 10.19415/j.cnki.1673-890x.2022.11.010

Ye, X., Qu, Z., Qi, Z., Handong, S., Françoise, G., and Yu, Z.. (2003). Eudesmane derivatives from Laggera pterodonta. *J. Fitoterapia.* 74, 459–463. doi: 10.1016/S0367-326X(03)00106-0

Ying, Q. X., Yue, D. L., and Yan, L. Q. (2006). Eudesma-5,12-dien-13-oic acid from Laggera pterodonta. *Acta Crystallogr E.* 62, o1844–o1845. doi: 10.1107/S1600536806012529

Zhang, T. Z., Li, J. H., Shen, A., and Li, S. Z. (2023). Bacteriostatic and preservative effects of multimethoxy flavonoid extracts from citrus peels on red grapes. *J. Food Industry Sci. Technol.* 44, 143–150. doi: 10.13386/j.issn1002-0306.2022090038

Zhao, Y. T., Wang, X. E., Zhao, Y., Pan, T., and Liu, L. M. (2022). Current status and prospect of research on fungus inhibiting active ingredients and inhibitory mechanism of plant sources. *J. Agric. Technology.* 42, 26–29. doi: 10.19754/j.nyyjs.20220830007

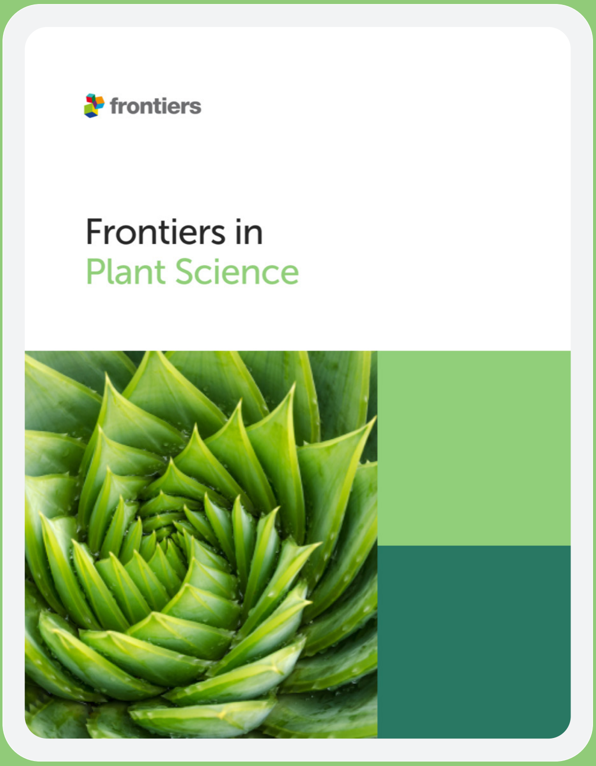
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →


frontiers

Frontiers in
Plant Science

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us
+41 (0)21 510 17 00
frontiersin.org/about/contact