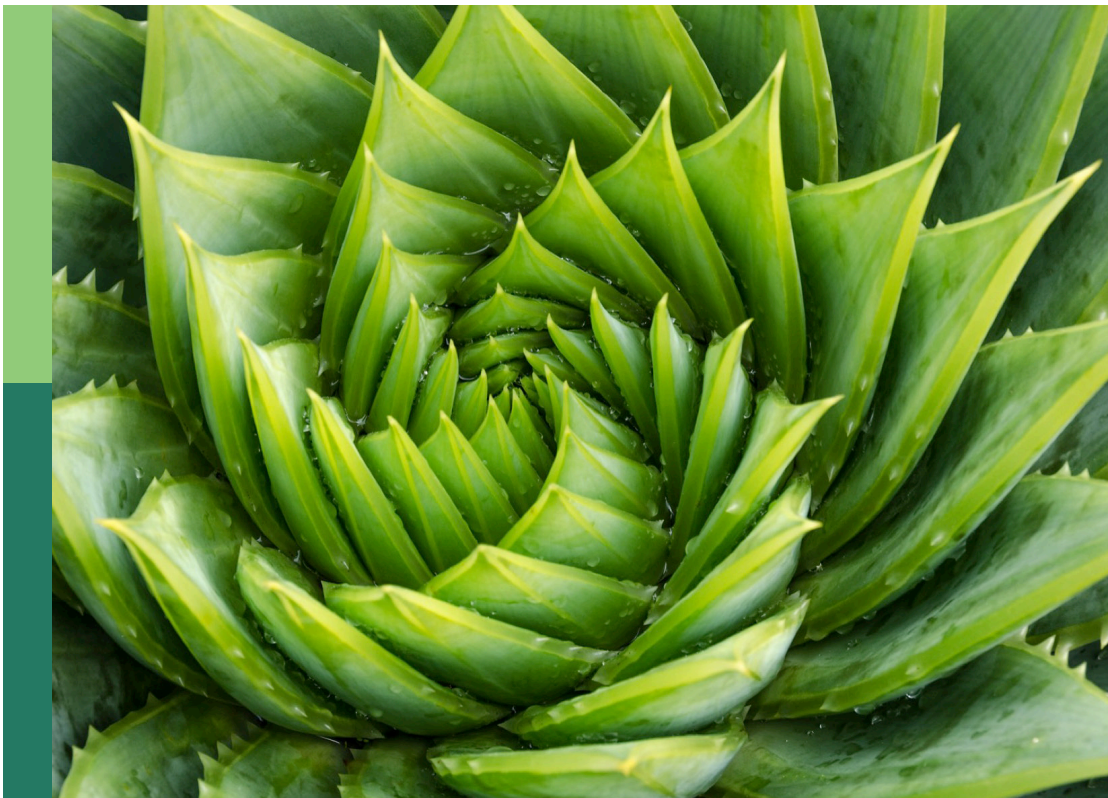
frontiers | Research Topics

Plant specialized metabolisms: physiological responses and molecular mechanisms

Edited by
Weiwei Zhang, Xiaomeng Liu, Feng Xu and Moonhyuk Kwon

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-6519-3
DOI 10.3389/978-2-8325-6519-3

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 specialized metabolisms: physiological responses and molecular mechanisms

Topic editors

Weiwei Zhang — Yangtze University, China

Xiaomeng Liu — Wuhan Polytechnic University, China

Feng Xu — Yangtze University, China

Moonhyuk Kwon — Gyeongsang National University, Republic of Korea

Citation

Zhang, W., Liu, X., Xu, F., Kwon, M., eds. (2025). *Plant specialized metabolisms: physiological responses and molecular mechanisms*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6519-3

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 17 June 2025
DOI 10.3389/fpls.2025.1629842

OPEN ACCESS

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

*CORRESPONDENCE
Weiwei Zhang
wwzhangchn@163.com
Moonhyuk Kwon
mkwon@gnu.ac.kr

RECEIVED 16 May 2025
ACCEPTED 29 May 2025
PUBLISHED 17 June 2025

CITATION
Cui X, Zhang W, Liu X, Xu F and Kwon M (2025) Editorial: Plant specialized metabolisms: physiological responses and molecular mechanisms.
Front. Plant Sci. 16:1629842.
doi: 10.3389/fpls.2025.1629842

Editorial: Plant specialized metabolisms: physiological responses and molecular mechanisms

Xinru Cui[1], Weiwei Zhang[1*], Xiaomeng Liu[2], Feng Xu[1] and Moonhyuk Kwon[3*]

[1]College of Horticulture and Gardening, Yangtze University, Jingzhou, China, [2]School of Modern Industry for Selenium Science and Engineering, National R&D Center for Se-rich Agricultural Products Processing Technology, Wuhan Polytechnic University, Wuhan, China, [3]Division of Applied Life Science (BK21 Four), Anti-Aging Bio Cell Factory Regional Leading Research Center, Research Institute of Molecular Alchemy, Gyeongsang National University, Jinju, Republic of Korea

KEYWORDS

specialized metabolites, biosynthetic mechanisms, regulatory networks, environmental adaptations, transcription factors

Editorial on the Research Topic:
Plant specialized metabolisms: physiological responses and molecular mechanisms

Plant specialized metabolites (PSMs), also referred to as "secondary metabolites", encompass a wide range of compounds such as flavonoids, quinones, terpenoids, alkaloids, and phenolic compounds. These metabolites have evolved to perform specific physiological and ecological roles, significantly contributing to plant adaptation under diverse environmental conditions, stress resistance, development processes, and metabolic regulation. Notably, certain PSMs possess physiological activities and pharmacological effects that hold promise in the prevention and treatment of various human diseases, including cancer, aging-related disorders, and cardiovascular conditions. Additionally, some PSMs have valuable industrial applications, such as natural rubber. The biosynthesis of PSMs exhibits remarkable diversity and complexity, varying across species, organs, tissues, and developmental stages. This biosynthetic complexity is governed by genetic factors and tightly regulated through interactions among transcription factors (TFs), *cis*-regulatory elements, and environmental cues (both biotic and abiotic stimuli).

Elucidating the biosynthetic pathways and gene functions associated with specific metabolites is essential for understanding their regulatory mechanisms. For instance, Long et al. employed comparative metabolomics across wolfberry and six other species (maize, rice, wheat, soybean, tomato, and grape), identifying 16 metabolites specific to wolfberry. By comparing the copy number of key enzymes in metabolite synthesis and degradation, they found that the phenyllactate degradation gene *UGT1* had the lowest copy

number among the six species, whereas the riboflavin and phenyllactate synthesis genes *RFK* and *HPPR* had higher copy numbers. This suggests that the copy numbers of *RFK, HPPR*, and *UGT1* may be the main reasons for the specific accumulation of riboflavin and phenyllactate in wolfberry. In *Rubus chingii*, Xu et al. identified 32 RcBBX transcriptional factors. By integrating BBX expression profiles across organs (roots, stems, leaves, flowers, fruits), developmental stages, and abscisic acid (ABA) treatments, they predicted RcBBX26 as a potential activator of anthocyanin biosynthesis and accumulation in red chestnut fruits. Moreover, the expression trends of seven anthocyanin biosynthetic genes (*Rc4CL4, Rc4CL5, Rc4CL6, Rc4CL12, RcUFGT8, RcUFGT9, and RcUFGT11*) were consistent with *RcBBX26* and anthocyanin accumulation during fruit ripening, indicating that *RcBBX26* positively regulates anthocyanin biosynthesis by activating target gene expression. In mung bean, Cho et al. identified six genes encoding four key enzymes (*CCoAOMT1*, *CYP81E1*, *DFR*, *HCT*), which commonly affect the levels of secondary metabolites (catechin, chlorogenic acid, formononetin, genistin). Regulatory network analysis revealed that NAC042 and MYB74 TFs orchestrate the expression of these enzymes, enhancing flavonoid content. These results could improve the nutritional value of mung beans and contribute to developing high-quality mung bean varieties.

Deciphering the biosynthetic mechanisms of various PSMs and increasing the production of valuable metabolites through gene editing are the key objectives in this field. Functional characterization plays a crucial role in unraveling complex molecular mechanisms. For example, Ma et al. utilized yeast two-hybrid and two-luciferase assays to show that CaMYBA, CaMYC, and CaTTG1 form an MYB-bHLH-WD40(MBW) complex, which directly bind promoters of anthocyanin synthesis structural genes such as *CaANS* to promote transcription and anthocyanin accumulation in pepper leaves. Covarrubias et al. found that overexpression of *SlLIP1* in tomato enhances fruit-specific *SlLIP1* transcripts, accompanied by increased bound and unbound linoleic acid (LA) content in fruit. Targeted metabolomics analysis of polar metabolites using LC-MS/MS showed that the LA content increase was associated with modifications at the level of transcripts of various genes involved in LA biosynthesis. Martinez-Sanchez et al. used RNA-seq and qPCR to analyze anthocyanin profiles and the expression of genes encoding anthocyanin biosynthetic enzymes revealed that SlATV represses anthocyanin biosynthesis by inhibiting key gene expression. Transient and stable transformation showed that pepino R2R3 MYB113 (AN1-like) is a key transcriptional activator for pepino anthocyanin accumulation. The study by Wolters et al. elucidated the interaction of dandelion TkSRPP with TkUGT80B1 providing a new connection between TkSRPPs and triterpenoid saponin synthesis in *T. koksaghyz* latex. This result will help to further elucidate the network of proteins linking TkSRPPs, stress responses, and NR biosynthesis within the cellular complexity of latex.

Plant adaptation to environmental changes is closely associated with the dynamic changes in the profiles of PSMs. Therefore, exposing plants to specific induced conditions is an effective way to identify changes in the corresponding metabolites and their regulatory genes. In tobacco, Wei et al. demonstrated the transformation of tobacco metabolites through different roasting processes. The analysis revealed a series of differentially expressed metabolites (DEMs) among fresh leaf, normal curing (NC), excessive curing (EC), and insufficient curing (IC) leaves at the end of 42°C, 54°C, and 68°C, respectively, suggesting that the roasting process regulates the transformation of tobacco metabolites and has a significant effect on the formation of tobacco quality. In addition, LC-MS/MS identified 845 metabolites, with flavonoids as the most abundant class. Petrovic et al. found that the specialized metabolism of *Nepeta nuda* L. leaves is more reprogrammable in response to differential growth conditions than that in inflorescences. Guo et al. combined proteomic and metabolomic analysis of the changes in tobacco leaves under under both topping and non-topping conditions. They found that the expression of proteins such as chalcone synthase (CHS), chalcone isomerase (CHI), naringenin 3-dioxygenase (F3H), and flavonoid 3'-monooxygenase (F3'H) was upregulated, and metabolites like pinocembrin, kaempferol, trifolin, rutin, and quercetin also increased thus enhancing the biosynthetic pathways of "flavonoid" and "flavone and flavonol".

Evolutionary divergence also shapes metabolite biosynthesis. Farzana et al. revealed and characterized a new ɣTMT-like enzyme, perivine Nb-methyltransferase (TePeNMT) from the plant *Tabernaemontana elegans*, distinguishing it from other ɣTMTs and ɣTMT-like NMTs. Their findings suggest that parallel evolution of ancestral gTMTs may be responsible for the occurrence of perivine N-methylation in *T. elegans* and *Catharanthus. roseus*.

Finally, this Research Topic includes a review by Wang et al. on different factors affecting the synthesis and accumulation of secondary metabolites in *Ficus carica*, including varieties, tissue type, environmental factors (e.g., light), stresses (e.g., high temperature, low temperature, drought, nutrient deficiencies, salinity), hormonal treatments, and developmental factors. Furthermore, they discussed the role of structural genes and TFs in the biosynthesis of secondary metabolites, specifically anthocyanins and furanocoumarins. The results of this research have important application prospects for further research and development of new *F. carica* varieties.

Author contributions

XC: Writing – original draft. WZ: Writing – original draft, Writing – review & editing. XL: Writing – review & editing. FX: Writing – review & editing. MK: Writing – review & editing.

Acknowledgments

We extend our appreciation to all the authors for contributing their scholarly work to our Research Topic. Additionally, we would like to convey our profound gratitude to all reviewers for their valuable assistance in the assessment of the manuscripts.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial associations that might be interpreted 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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 April 2024
DOI 10.3389/fpls.2024.1392175

OPEN ACCESS

EDITED BY
Moonhyuk Kwon,
Gyeongsang National University, Republic of Korea

REVIEWED BY
Hyung Won Ryu,
Korea Research Institute of Bioscience and Biotechnology (KRIBB), Republic of Korea
Jaeyoung Kwon,
Korea Institute of Science and Technology, Republic of Korea

*CORRESPONDENCE
Cheng Jin
jincheng@hainanu.edu.cn
Yuanyuan Zhang
Yuanyuan.zhang@hainanu.edu.cn

†These authors share first authorship

RECEIVED 27 February 2024
ACCEPTED 15 April 2024
PUBLISHED 26 April 2024

CITATION
Long Q, Zhang C, Zhu H, Zhou Y, Liu S, Liu Y, Ma X, An W, Zhou J, Zhao J, Zhang Y and Jin C (2024) Comparative metabolomics combined with genome sequencing provides insights into novel wolfberry-specific metabolites and their formation mechanisms. *Front. Plant Sci.* 15:1392175. doi: 10.3389/fpls.2024.1392175

Comparative metabolomics combined with genome sequencing provides insights into novel wolfberry-specific metabolites and their formation mechanisms

Qiyuan Long[1,2†], Changjian Zhang[1,2†], Hui Zhu[1,2], Yutong Zhou[2], Shuo Liu[1,2], Yanchen Liu[1,2], Xuemin Ma[3], Wei An[4], Jun Zhou[5], Jianhua Zhao[4], Yuanyuan Zhang[1,2*] and Cheng Jin[1,2*]

[1]School of Breeding and Multiplication (Sanya Institute of Breeding and Multiplication), Hainan University, Sanya, China, [2]School of Tropical Agriculture and Forestry, Hainan University, Haikou, Hainan, China, [3]Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, Umeå, Sweden, [4]National Wolfberry Engineering Research Center, Wolfberry Science Research Institute, Ningxia Academy of Agriculture and Forestry Sciences, Yinchuan, China, [5]College of Biological Science and Engineering, North Minzu University, Yinchuan, China

Wolfberry (*Lycium*, of the family Solanaceae) has special nutritional benefits due to its valuable metabolites. Here, 16 wolfberry-specific metabolites were identified by comparing the metabolome of wolfberry with those of six species, including maize, rice, wheat, soybean, tomato and grape. The copy numbers of the riboflavin and phenyllactate degradation genes *riboflavin kinase* (*RFK*) and *phenyllactate UDP-glycosyltransferase* (*UGT1*) were lower in wolfberry than in other species, while the copy number of the phenyllactate synthesis gene *hydroxyphenyl-pyruvate reductase* (*HPPR*) was higher in wolfberry, suggesting that the copy number variation of these genes among species may be the main reason for the specific accumulation of riboflavin and phenyllactate in wolfberry. Moreover, the metabolome-based neighbor-joining tree revealed distinct clustering of monocots and dicots, suggesting that metabolites could reflect the evolutionary relationship among those species. Taken together, we identified 16 specific metabolites in wolfberry and provided new insight into the accumulation mechanism of species-specific metabolites at the genomic level.

KEYWORDS

metabolome, nutrition, riboflavin, phenyllactate, copy number variation

1 Introduction

Wolfberry (*Lycium*, of the family Solanaceae) has excellent nutritional value, with a history dating back thousands of years (Jin et al., 2013). The wolfberry genus contains ~80 species, with a discrete geographic distribution from South and North America to Australia, Eurasia, the Pacific Islands, and South Africa. There are 7 species and 3 varieties in China, mainly distributed in the north (Levin and Miller, 2005). Consuming wolfberry could promote human health by nourishing the liver and kidneys, enhancing vision and regulating the immune system (Vidović et al., 2022). As a result, a large number of studies have assessed the nutritional value of wolfberry from various perspectives. For example, one of the key components responsible for the antioxidant, immunomodulatory, and anticancer effects in *Lycium barbarum* L. is *Lycium barbarum* polysaccharide (LBP) (Jin et al., 2013; Masci et al., 2018). *Lycium barbarum* fruits (LBFs) flavonoids are involved in prominent antioxidant, hypolipidemic, hypoglycemic, immunity-enhancing, and antitumor activities (Yang et al., 2022). The anthocyanins from *Lycium ruthenicum* Murray have a positive role in maintaining intestinal health and play an antioxidant role (Yan et al., 2018). Ascorbic acid and its derivatives in LBFs can regulate the intestinal flora in mice (Huang et al., 2020). However, except for these known active substances, wolfberry-specific metabolites are still poorly known.

To date, various metabolic methods have been applied to the determination of metabolites in wolfberry. Fifty-six volatile compounds in Ningxia goji berries were characterized by gas chromatography-spectrometry (GC-MS) and identified by gas chromatography-olfactometry (GC-O) and aroma dilution analysis (AEDA) (Lu et al., 2017). Using ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-Q-TOF/MS), 41 spermidine derivatives were tentatively characterized from LBFs (Ahad et al., 2020). Nine alkaloids were yielded in LBFs by spectroscopic analyses and chemical methods (Chen et al., 2021). Thirteen flavonoid compounds were identified in LBFs using LC-MS (Yang et al., 2022). Based on these metabolic methods, the comparative metabolome of different varieties (Zhang et al., 2016), tissues (Xiao et al., 2021) or development stages of wolfberry (Zhao et al., 2015) under various environmental conditions (Poggioni et al., 2022) has been extensively studied. For example, by comparing the metabolic groups of *Lycium barbarum*, *Lycium chinense*, and *Solanum lycopersicum*, the metabolic markers distinguishing *Lycium* and *Solanum* fruits were revealed (Dumont et al., 2020). However, there is still a large gap in metabolome comparison between wolfberry and different species, which needs further exploration.

To study the metabolic mechanism of wolfberry-specific substances, many studies have been carried out on the regulatory mechanism of metabolites during fruit ripening. The *LbNCED1* transcript level was transcriptionally activated by the developmental cues of *Lycium* fruit, enhancing the accumulation of abscisic acid (ABA), thereby promoting anthocyanin production and leading to fruit coloration (Li et al., 2019). Distinction in the expression patterns of 22 transcription regulators may be the main reason for the morphological and phytochemical differences between *L. barbarum* (LB) and *L. ruthenicum* (LR) fruits at five developmental stages (Zhao et al., 2020b). Recent research has shown that *LbNR* (nitric reductase (NR) from *L. barbarum*) inhibited anthocyanin biosynthesis and enhanced proanthocyanidin (PA) accumulation by regulating nitric oxide (NO) (Li et al., 2020a). Moreover, based on metabolome and transcriptome analysis, many key genes involved in metabolite synthesis were identified in wolfberry. For example, candidate genes for flavonoid biosynthesis were identified by conducting transcriptome and flavonoid metabolic profiling, and the molecular regulatory mechanism of *LrAN1b* on anthocyanins and fruit color was verified, which provided a new understanding of the potential mechanism of action of flavonoids (Li et al., 2020b). Through competitive transcriptome analysis between LB and LR, 38 MYB transcription factors that may regulate the fruit development of wolfberry were identified (Wang et al., 2020). Recently, the first reference genome of wolfberry was published (Cao et al., 2021), indicating that the genetic basis of metabolites in wolfberry can be analyzed at the genomic level.

To comprehensively explore the specific metabolites and their causes in wolfberry, we compared the metabolome of wolfberry, rice, maize, wheat, soybean, grape, and tomato. A total of 1043 distinct metabolic features were detected and were divided into 10 categories, 16 of which were identified as wolfberry-specific metabolites. Our results showed that metabolites could reflect the evolutionary relationship among different species. We further showed that the copy numbers of *RFK*, *HPPR*, and *UGT1* may be the main reasons for the specific accumulation of riboflavin and phenyllactate in wolfberry.

2 Materials and methods

2.1 Plant materials

To study the differences in metabolites between wolfberry (*Lycium*, of the family Solanaceae) and other species, we selected three monocot crops and three dicot crops as reference objects. The monocot crops included three major staple crops maize (*Zea mays* L.), rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.), with rice being the model plant for monocots. The dicot crops included legume crop soybean (*Glycine max* (L.) Merr.) and the fruit crops tomato (*Solanum lycopersicum* L.) and grape (*Vitis vinifera* L.). Wolfberry (*L.barbarum* 'Ningqi No.1', *L. barbarum* var. *auranticarpum*, *L. ruthenicum*) was obtained from Yinchuan, Ningxia (E113°42′, N34°48′). Maize (Waxy maize, Red Waxy maize, Fruit maize, Sweet maize), Grape (Red grape, Green grape, Seedless red grape, Jufeng grape), Tomato (Pink tomato, Tomato, Cherry tomato, Millennium cherry tomato) were obtained from Haikou, Hainan (E110°20′, N20°02′). Soybeans (GDC058, GDC062, GDC063) were obtained from Zhengzhou, Henan (E113°42′, N34°48′). Wheat (Lumai 21, Heng 7228, Linmai No. 2, Xinong 529) was obtained from Zhaoxian, Hebei (E114°28′, N38°02′). Rice (Huanghuazhan, Nipponbare, Mimghui 63, Zhengshan 97, Zhonghua 11) was obtained from Hainan University (E110°20′, N20°02′).

2.2 Sample preparation and extraction

The samples were put in a lyophilizer for vacuum freeze-drying, and they were crushed in a mixer mill (MM 400; Retsch, Haan, Germany) for 1 min at 30 Hz. Next, 80 mg of the powdered sample was weighed into a 2 mL centrifuge tube, and 70% aqueous methanol (v/v) with lidocaine internal standard was added to extract the water-soluble metabolites (pure methanol was used to extract the fat-soluble metabolites). Then, the tube was vortexed for 10 seconds, allowed to stand for 10 min, repeated three times, and placed in a refrigerator at 4°C for 10-12 hours. Then, the sample was centrifuged (4°C, 10000 rpm, 10 min), the supernatant was pipetted, the water-soluble and fat-soluble metabolites were mixed 1:1 and filtered by a microporous filtration membrane (SCAA-104, 13 mm, 0.22 μm, Shanghai Anpu Experimental Technology Co., Ltd., http://www.anpel.com.cn/), and the sample was filtered into an injection bottle for storage for UPLC-MS analysis.

2.3 Detection of metabolites

The instruments used for LC-MS/MS analysis included UPLC-Q Exactive Plus Orbitrap HRMS and UPLC-Q-Trap 6500+ MS. The analytical column used was a C18 column (Shim-pack GLSS C18, 1.9 μm, 2.1*100, Shimadzu).

UPLC chromatographic conditions: mobile phase A is an aqueous solution containing 0.04% glacial acetic acid, and mobile phase B is a methanol solution containing 0.04% glacial acetic acid. Elution gradient: At 0 min, $V_{phase\ A}$: $V_{phase\ B}$ = 95:5; At 10 min, $V_{phase\ A}$: $V_{phase\ B}$ = 5:95; At 11 min, $V_{phase\ A}$: $V_{phase\ B}$ = 5:95; At 11.1 min, $V_{phase\ A}$: $V_{phase\ B}$ = 95:5, this ratio was continued until the end (duration is 14 minutes). The column temperature was set to 40°C, the injection volume of the injector was 2 μL, and the flow rate was 0.35 mL/min.

Orbitrap HRMS mass spectrometry conditions: ESI ionization method, mass spectrometry scanning mode is Full MS/ddMS2, ion collection mass range is 100~1200 m/z, and the lysis voltage is set to 20, 40, 60 eV; The spray voltage in positive ion mode is 3.5 kV, the capillary temperature is 350°C, the heater temperature is 350°C, the sheath gas (nitrogen) flow rate is 40 arb, and the auxiliary gas (nitrogen) flow rate is 10 arb. The spray voltage in negative ion mode is -3.0 kV, the capillary temperature is 350°C, the heater temperature is 350°C, the sheath flow rate is 30 arb, and the auxiliary air flow rate is 10 arb.

For wide-target detection of metabolites by UPLC-Q-Trap 6500 + multiplex reaction monitoring mode, the MRM detection window was set to 60 s, and the target cycle time was set to 0.8 s. The raw data were integrated by Multi Quant 3.0.3 to accurately obtain the relative content of each substance.

2.4 Metabolome analyses

All statistical analyses were carried out using R (4.1.1, http://www.r-project.org). (PCA) was performed using the R package "FactoMineR" (Lê et al., 2008) with 587 metabolite data from 3 biological replicates of mixed samples of each species to evaluate the metabolome differences across seven species. Circular plots were constructed using the R package Circlize, with the raw data normalized and scaled in the R program; the y-axes of the seven circles were set on the same scale. The hierarchical clustering tree using metabolome data of the seven species was constructed using the R package "hclust". Orthogonal partial least squares discriminant analysis (OPLS-DA) was conducted by the R package "ropls" (Thévenot et al., 2015) to identify the major discriminant metabolite features among different species. A metabolite feature was considered a species-specific metabolic trait when matching the following criteria compared with the other six species: (i) *P* value of paired *t-test* ≤ 0.05; (ii) Fold Change ≥ 3. Venn diagrams were generated on the online website jvenn (http://jvenn.toulouse.inra.fr/app/index.html).

2.5 Enrichment analysis

Heatmaps were generated using the R package "pheatmap" with data normalization to divide the metabolome into eight clusters, and comparisons were performed via an average linkage method based on the Manhattan distance. Metabolic traits of different clusters were used for the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis to elucidate the differential metabolic pathways among these seven different species. Enrichment analysis was carried out using the online platform MetaboAnalyst (https://www.metaboanalyst.ca/) selecting *Oryza sativa japonica* (Japanese rice) (KEGG) as a reference database. The bubble plot was drawn by the R package "ggplot2" (Ginestet, 2011).

3 Results

3.1 Metabolomic profiling of seven species

To dissect the metabolome differences between wolfberry and diverse species, six popular species were selected for comparison with wolfberry for analysis in this study. These species include four cereal crops maize, rice, wheat, legume crop soybean and two fruit crops tomato and grape. To reduce the specific error associated with a variety, we collected commercially available varieties of these seven species and mixed them to prepare mixed samples that represented the metabolome of each species. Collectively, a total of 1043 distinct metabolic features were detected and quantified in these seven species, the identification class is shown in Supplementary Figure S1A. Based on the comparison of our local metabolite database and commercial standards (Chen et al., 2014; Li et al., 2022b), 587 metabolites were identified, which could be divided into 10 categories, including 127 phenylpropanoids, 125 amino acids and their derivatives, 116 organic acids and sugars, 66 lipids, 40 nucleic acids and their derivatives, 35 vitamins and cofactor derivatives, 20 terpenoids, 15 alkaloids, 10 phytohormones and 33 other metabolites (Figure 1A, Supplementary Table S1).

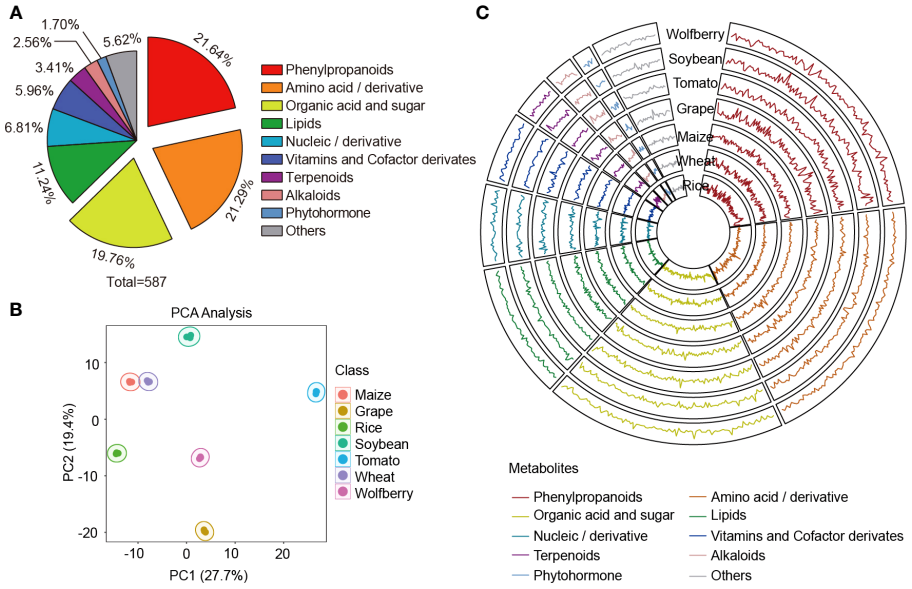


FIGURE 1
Multispecies metabolic spectrum composition analysis. **(A)** Classification of metabolites that have been speculated or verified. A total of 587 metabolites in wolfberry, soybean, tomato, grape, maize, wheat and rice were detected in this study. **(B)** Principal component analysis of the 587 metabolites among the seven species. **(C)** Features of the 587 metabolites in the seven species. Raw data were normalized and scaled using the R program, and data points show the average values of three replicates of each metabolite. The y-axes of the seven circles are on the same scale, from 0.508 to 1.556. Substance classes are represented by the line colors.

Principal component analysis (PCA) was performed to assess the overall metabolome differences under the unsupervised model. The first principal component (PC1) explained the greatest variance (27.7%) of the variation among the seven species. Among the top 50 contributors of PC1, the proportion of amino acid and their derivative was the highest, followed by organic acids and sugars (Supplementary Table S2). The second principal component (PC2) was orthogonal to PC1 and was the largest in the remaining variance, accounting for 19.4% of the variance. Among the top 50 contributors in PC2, phenylpropanoids had the highest proportion, followed by amino acid and their derivative (Supplementary Table S2). For example, ornithine hydrochloride (ms025), the first-ranked identified metabolite, was found to accumulate more in the species with negative PC2 (grape, wolfberry and rice) than in those with positive PC2 (maize, wheat, soybean and tomato) (Supplementary Figure S1B). The results showed that the seven species were clearly separated, and three biological replicates of each species were compactly grouped together. This discrete clustering of the respective species indicates the distinct attributes of each species and the high repeatability and reliability of the experimental results (Figure 1B).

To visualize the variation in the metabolome among different species, we generated a circular plot for the seven species (Figure 1C). In the circular plot, strong variation in metabolic accumulation across seven species could be observed, with phenylpropanoids, nucleic acids and derivatives as well as phytohormones showing the most pronounced variation. To quantify the degree of variation, we used the coefficient of variation (CV) values of metabolites across the seven species. Phenylpropanoids were found to be the most variable metabolites, with a CV range from 11.56% to 244.93%, which is consistent with the results shown in the circular plot (Supplementary Figure S1C, Supplementary Table S3). In contrast, nucleic acids and derivatives have the least variation, probably because they are crucial components of genetic material in all species, some of which play vital roles in influencing the structure and function of RNA or in post-transcriptional gene regulation (Dominissini et al., 2016). Therefore, the conservation of these compounds, which represented the minimum variation in our data, implied their importance in living organisms, illustrating the role of nucleic acids and derivatives as building blocks in the life of organisms.

3.2 Differential accumulation patterns of metabolites among seven species

The overall profile of all identified metabolites in the seven species was analyzed by hierarchical cluster analysis (HCA). As shown in the heatmap, the metabolic diversity in different species was further indicated. Furthermore, all the compounds were hierarchically clustered into eight main clades, the first seven of which specifically accumulated in wolfberry, rice, maize, wheat, tomato, grape and soybean, indicating that each species has its preferential metabolites (Figure 2A, Supplementary Table S4).

To explore the preferred metabolic pathway of each species, we performed pathway enrichment analysis of the metabolites in clade 1 to clade 7 by the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Figure 2B, Supplementary Figure S2). Here, we focused on the preferred metabolite of wolfberry in clade 1. The 43

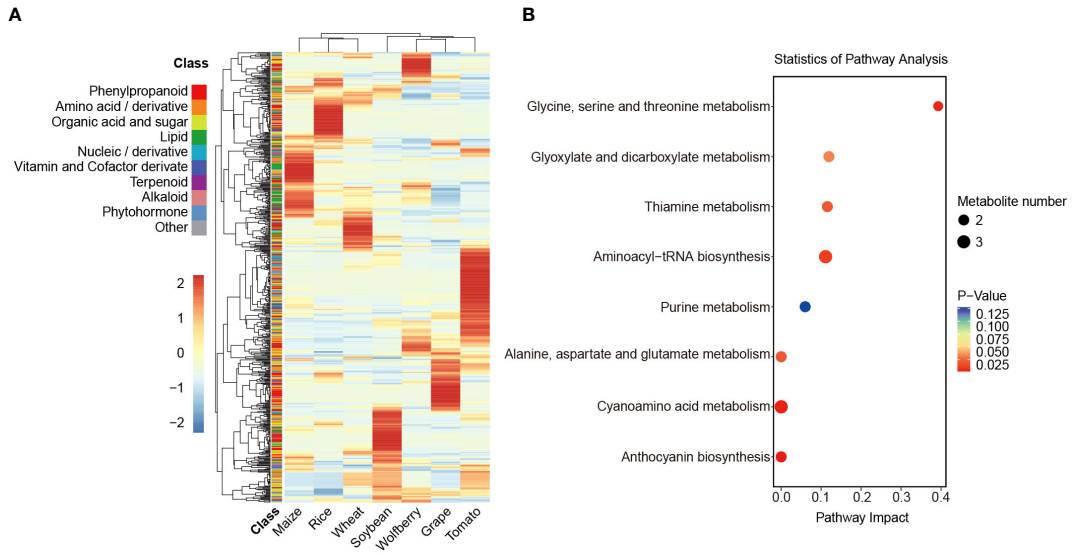


FIGURE 2
Analysis of differences in the relative content of metabolites. **(A)** Cluster analysis of the relative content of annotated metabolites in 7 species. The relative content of the annotated metabolites was represented by the mean of three biological replicates and normalized by z score standardization. Metabolites were clustered by hierarchical cluster analysis. Shades from blue to red in the figure represent increasing levels of metabolites, and the 10 color blocks in the class column indicate the classification of metabolites. **(B)** KEGG enrichment analysis of metabolites specifically accumulated in wolfberry. The *P* value ranges from 0.125 to 0.025, as seen by the color scheme of red to blue. The size of the dots represents the number of metabolites.

substances in clade 1 showed specific accumulation in wolfberry and were classified into 9 different categories. The two largest metabolites among them are organic acids and sugars and amino acids and their derivatives, accounting for 25.58% and 23.26%, respectively. Organic acids and sugars in wolfberry fruit are not only important nutrients but also the main influencing factors of flavor quality, while amino acids and derivatives are the main nutritional and medicinal components in LBFs (Zhao et al., 2020a). In the KEGG pathway enrichment analysis, the 43 metabolites were involved in 19 pathways. The major pathways are presented in the bubble plot (Figure 2B). In addition to amino acid metabolism, they also include glyoxylate and dicarboxylate metabolism, thiamine metabolism, aminoacyl-tRNA metabolism, purine metabolism, cyanoamino acid metabolism and anthocyanin biosynthesis. Similarly, 48 metabolites in clade 6 showed more accumulation in grapes and were enriched in the glutathione metabolic pathway. This could explain the antioxidant activity conferred by glutathione in grapes (Supplementary Figure S2F). Collectively, these results suggest that the specific accumulation of metabolites in different species can represent and determine their specific nutritional value.

3.3 Metabolic profile reflects the evolutionary relationship between monocotyledons and dicotyledons

To explore the affinities among the seven species, a neighbor-joining tree was constructed using metabolome data of the seven species (Figure 3A), and we also created a phylogenetic tree using the whole genome protein sequence of the seven species Supplementary (Figure S3). The dicotyledons wolfberry, tomato, soybean and grape were clustered in both metabolome-based and protein-based trees, while the monocotyledons wheat, rice and maize were clustered together. These results indicate that there are differences in the metabolome between monocots and dicots, and the metabolomes could reflect the evolutionary relationship among different species.

To estimate the classification performance of monocots and dicots, we conducted supervised PCA and partial least-squares discriminant analysis (PLS-DA) among the seven species. The R^2X, R^2Y and Q^2 are 0.436, 0.998 and 0.849, respectively, which are the prediction parameters of the PLS-DA model. The R^2X value reflects the degree to which the model explains the variability of the input data, while the R^2Y and Q^2 values represent the model's explanatory power for the output variable and its predictive accuracy, respectively. Here, R^2Y and Q^2 were both close to 0.9, with the same order of magnitude, indicating the stability and trustworthiness of this model (Golbraikh and Tropsha, 2002; Blasco et al., 2015; Su et al., 2022). The results showed that the main principal components (PC1 and PC2) explained 46.6% of the variability, with monocots and dicots being well discriminated from each other (Figure 3B).

The PLS-DA loading values of these 587 compounds are listed in Supplementary Table S5 to quantify the contribution to the classification of the metabolites. On the basis of these data, two heatmaps were generated to organize the accumulation levels of the most effective contributors within monocots and dicots. We found that the top 50 contributors preferentially accumulated in dicots

10.3389/fpls.2024.1392175

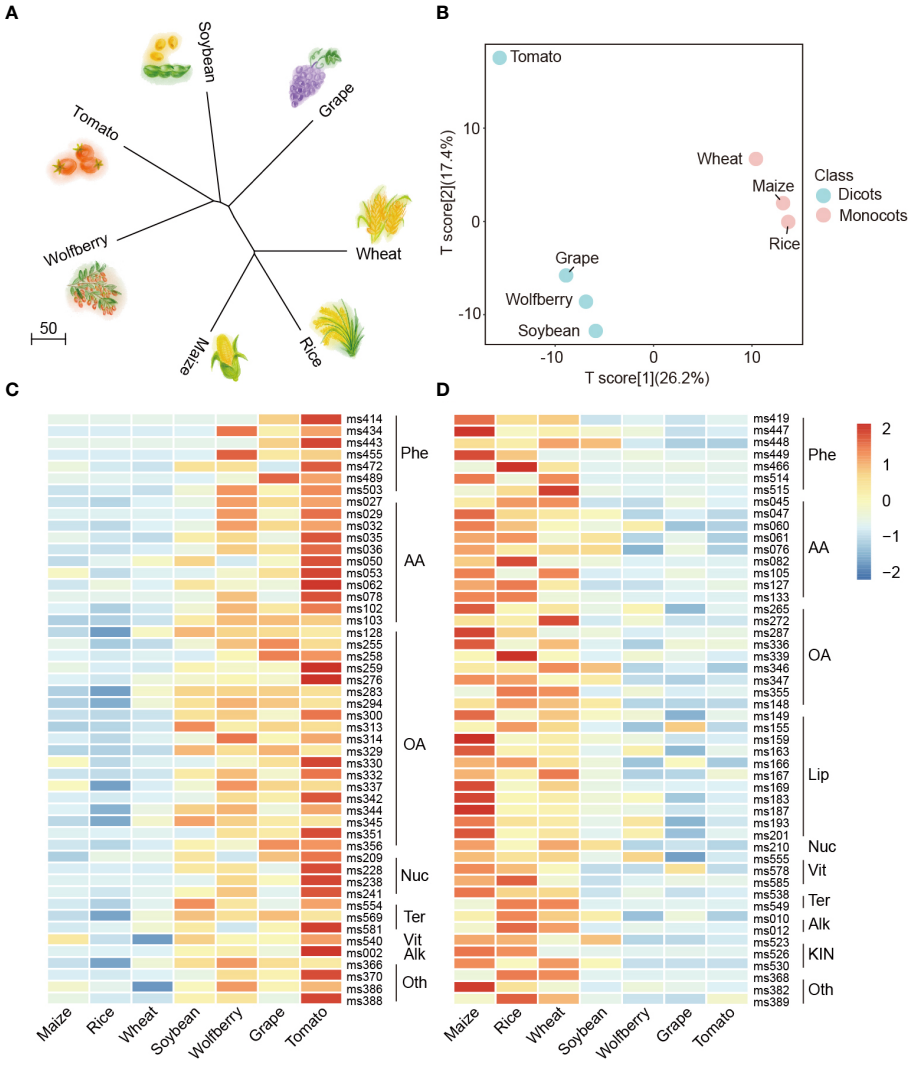


FIGURE 3
Metabolome analysis between monocotyledons and dicotyledons. **(A)** Neighbor-joining tree of the seven species with metabolome data. **(B)** PLS-DA analysis of the metabolite profiling of seven species. The data represent the mean values obtained from three biological replicates of each species. Powder blue and baby pink indicate dicots and monocots, respectively. **(C, D)** Comparisons of the relative accumulation levels of the top 50 **(C)** and bottom 50 **(D)** contributors identified in the PLS-DA distribution in the dicots and monocots. The relative content of metabolites was represented by the mean of three biological replicates and normalized by z score standardization. Metabolites were clustered by hierarchical cluster analysis. Shades from blue to red in the figure represent increasing levels of metabolites. Phe, Phenylpropanoids; AA, Amino acids/derivatives; OA, Organic acids and sugars; Nuc, Nucleics/derivatives; Ter, Terpenoids; Vit, Vitamins and Cofactor derivates; Alk, Alkaloids; Oth, Others; Lip, Lipids; KIN, Phytohormones.

(Figure 3C, according to Supplementary Table S5, indicated in red), and organic acids and sugars (38%) as well as amino acids and derivatives (22%) were predominant among them, which might be identified as the specific accumulated components of dicots. In contrast, the bottom 50 contributors tended to have a preferential accumulation pattern in monocots (Figure 3D, Supplementary Table S5, indicated in blue), and it can be inferred that lipids and phytohormones accumulated specifically only in monocots. The aforementioned results reveal that the differentiation between monocots and dicots can mainly be ascribed to the considerable variation in compounds such as organic acids and sugars, amino acids and derivatives, phenylpropanoids and lipids. Together, these results indicate that the metabolome reflects evolutionary relationships between different crops.

3.4 Species-specific metabolites of the seven species

To identify those metabolites that could be used to split and distinguish the seven species, we calculated the fold change for each compound among the species. We defined the metabolites in a certain species whose content was more than 3 times higher than that of other species as species-specific metabolites. There were 16, 29, 14, 21, 15, 52 and 23 species-specific metabolites in wolfberry, rice, wheat, maize, soybean, tomato, and grape, respectively (Figure 4A). Seven phenylpropanoids (cyanidin chloride, echinacoside, esculoside, isorhamnetin 3-O-neohesperidoside, methyl p-coumarate, narcissoside and scopoletin), two amino acids and their derivatives (L-asparagine and N-acetylneuraminic acid),

10.3389/fpls.2024.1392175

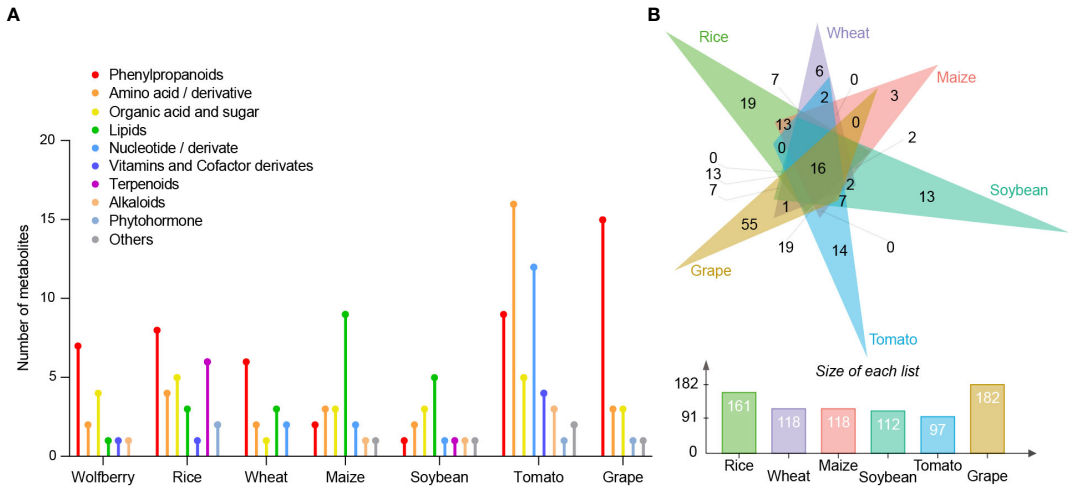


FIGURE 4
Comparison of differential metabolites between different species. **(A)** The number of known classified metabolites in each species showing more than threefold higher content than those in the other six species. The Y-axis refers to the number of metabolites. **(B)** Venn diagram analysis of wolfberry-specific metabolites relative to each species. Each color triangle denotes the wolfberry-specific metabolites relative to each species (top). Each color bar represents the number of wolfberry-specific metabolites relative to each species (bottom). The Y-axis represents the number of metabolites. Green: rice; purple: wheat; pink: maize; blue-green: soybean; blue: tomato; yellow: grape.

four organic acids and sugars (2-hydroxyisocaproic acid, 5-hydroxyhexanoic acid, phenyllactate and shikimic acid), one lipid (cholesterol), one vitamin and coenzyme derivative (riboflavin), and one alkaloid (4-amino-5-hydroxymethyl-2-methylpyrimidine) were among the 16 wolfberry-specific metabolites. Twenty-nine rice-specific metabolites comprised six terpenoids, two phytohormones, four amino acids and their derivatives, five organic acids and sugars, three lipids, eight phenylpropanoids and one vitamin and coenzyme derivative. Six phenylpropanoids, two amino acids and their derivatives, one organic acid and sugar, three lipids, and two nucleic acids and their derivatives are among the 14 wheat-specific metabolites. Among the 21 maize-specific metabolites were two phenylpropanoids, three amino acids and their derivatives, three organic acids and sugars, nine lipids, two nucleic acids and their derivatives, one alkaloid and one other compound. Fifteen soybean-specific metabolites included one phenylpropanoid, two amino acids and their derivatives, three organic acids and sugars, five lipids, one nucleic acid and its derivative, one terpenoid, one alkaloid and one other substance. Fifty-two tomato-specific metabolites included nine phenylpropanoids, 16 amino acids and their derivatives, five organic acids and sugars, 12 nucleic acids and their derivatives, four vitamins and coenzyme derivatives, three alkaloids, one phytohormone and two other compounds. In addition, 23 grape-specific metabolites were identified, comprising 15 phenylpropanoids, three amino acids and their derivatives, three organic acids and sugars, one phytohormone and one other compound.

We found that phenylpropanoids, organic acids and sugars accounted for the largest proportion of specific metabolites in wolfberry, which may determine the special nutritional value of wolfberry. To further compare the metabolites specifically accumulated in wolfberry compared to each species, we performed a Venn diagram analysis. The Venn diagram shows the number of wolfberry-specific metabolites compared with other species. Wolfberry had the most differential metabolites compared with grapes, followed by rice, corn and wheat, the three monocotyledon crops (Figure 4B). Venn diagrams of specific metabolites of other species are shown in Supplementary Figure S4. We found that all species had the highest number of specific metabolites relative to grape, while grape had the most specific metabolites compared to wheat.

3.5 Variation in the number of gene copies causes specific metabolites of Lycium barbarum to accumulate

The copy numbers of the riboflavin and phenyllactate degradation genes *riboflavin kinase* (*RFK*) and *phenyllactate UDP-glycosyltransferase* (*UGT1*) were lower in wolfberry than in other species, while the copy number of the phenyllactate synthesis gene *hydroxyphenyl-pyruvate reductase* (*HPPR*) were higher in wolfberry, suggesting that the copy number variation of these genes among species may be the main reason for the specific accumulation of riboflavin and phenyllactate in wolfberry. To illuminate the genetic basis of the high metabolite content in wolfberry, we investigated the metabolic pathway of riboflavin and phenyllactate, which are more than five times more abundant in wolfberry than in other species. *RFK*, encoding riboflavin kinase, which is responsible for the degradation of riboflavin, had the lowest number of copies in wolfberry (Figure 5). This could lead to the weakening of riboflavin's ability to turn into flavin mononucleotide (FMN), thus increasing the accumulation of riboflavin in wolfberry. *HPPR*, encoding hydroxyphenyl-pyruvate reductase, which is responsible for synthesizing phenyllactate, has the highest copy number in wolfberry when compared with the

10.3389/fpls.2024.1392175

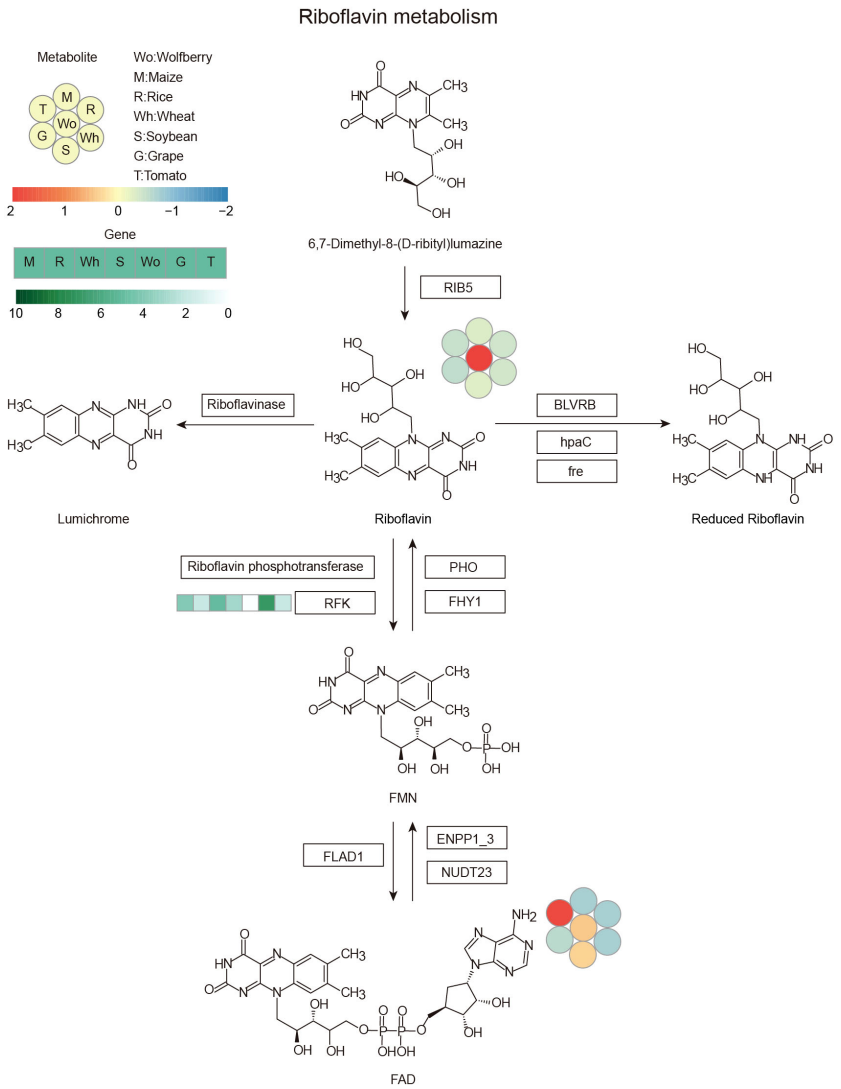


FIGURE 5
Riboflavin metabolic pathway. The heatmap shows the accumulation of metabolites and the copy number of genes among the seven species. Left corner there is a scale bar. The circles represent the accumulation levels of metabolites, and the data are represented by the average values of three biological replicates and normalized with z score standardization. Shades from blue to red represent increasing metabolite levels. The rectangular heatmap represents the copy numbers of genes, with green representing a high copy number and white representing a low copy number. The rectangle box indicates the key enzyme, and the abbreviation: RIB5, riboflavin synthase; BLVRB, biliverdin reductase/flavin reductase; hpaC, flavin reductase (NADH); fre, NAD(P)H-flavin reductase; PHO, acid phosphatase; FHY1, FMN hydrolase/5-amino-6-(5-phospho-D-ribitylamino) uracil phosphatase; RFK, riboflavin kinase; FLAD1, FAD synthetase; ENPP1_3, ectonucleotide pyrophosphatase/phosphodiesterase family member 1/3; NUDT23, ADP-ribose/FAD diphosphatase. Wo, M, R, Wh, S, G and T represent wolfberry, maize, rice, wheat, soybean, grape and tomato, respectively.

other six species (Figure 6). As a result, the amount of phenyllactate in wolfberry may dramatically increase. Moreover, *UGT1*, encoding a key enzyme in the initial step of phenyllactate degradation, had the fewest copies in wolfberry (Figure 6). This will slow down the reaction of phenyllactate to phenyllactylglucose in wolfberry. Therefore, the increased synthesis and decreased degradation of phenyllactate are jointly responsible for the specific accumulation of phenyllactate in wolfberry (Figure 6). Taken together, our results showed that specific accumulation of riboflavin and phenyllactate in wolfberry may be caused by variation in the number of genes involved in riboflavin and phenyllactate synthesis or degradation.

4 Discussion

In this study, we discovered 16 metabolites that preferentially accumulate in wolfberry and showed that the copy number of genes may be one of the main factors contributing to the accumulation of these metabolites. We also demonstrate that metabolite contents can be utilized as indicators to determine the evolutionary relationship between various species.

In recent years, the nutritional components in plants have aroused broad concern. (Sun et al., 2024) Wolfberry has special nutritional value, and scholars have conducted extensive research on it. Over the

10.3389/fpls.2024.1392175

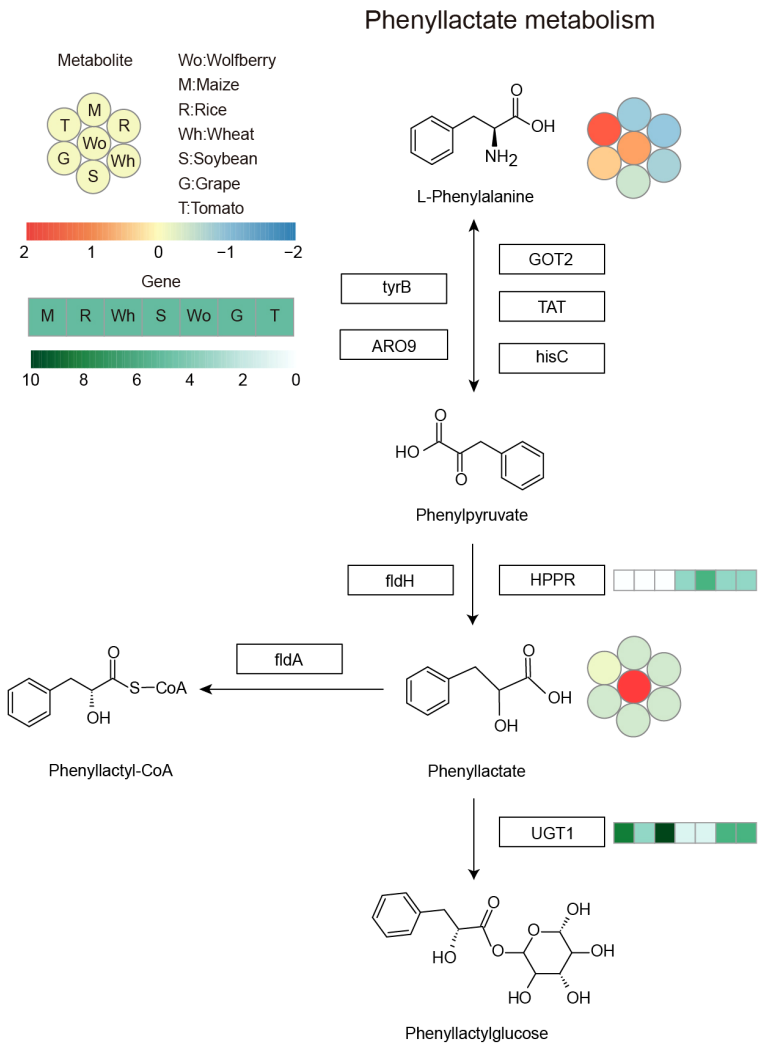


FIGURE 6
Phenyllactate metabolism pathway. The heatmap shows the accumulation of metabolites and the copy number of genes among the seven species. Left corner there is a scale bar. The circles represent the accumulation levels of metabolites, and the data are represented by the average values of three biological replicates and normalized with z score standardization. Shades from blue to red represent increasing metabolite levels. The rectangular heatmap represents the copy numbers of genes, with green representing a high copy number and white representing a low copy number. The rectangular box indicates the key enzyme, and the abbreviations are as follows: GOT2, aspartate aminotransferase, mitochondrial; TAT, tyrosine aminotransferase; hisC, histidinol-phosphate aminotransferase; tyrB, aromatic-amino-acid transaminase; ARO9, aromatic amino acid aminotransferase II; fldH, aromatic 2-oxoacid reductase; HPPR, hydroxyphenyl-pyruvate reductase; fldA, cinnamoyl-CoA:phenyllactate CoA-transferase; UGT1, phenyllactate UDP-glycosyltransferase. Wo, M, R, Wh, S, G and T represent wolfberry, maize, rice, wheat, soybean, grape and tomato, respectively.

past few years, with the advancement of metabolomic methods, metabolomic studies of various plants have been extensively studied, including wolfberry. Using widely targeted LC-MS/MS, the metabolites and their spatial distribution in dried kernels of six representative bread wheat cultivars in China were determined. Flavonoids varied the most in different varieties, and the concentration was higher in the outer layer of the grain but lower only in the kernel (Zhu et al., 2022). The metabolome study of maize and rice showed significant interspecific differences in their metabolic variation and identified flavonoids as the key constituent of interspecific metabolic divergence (Deng et al., 2020). A study with widely targeted metabolomics in three major food crops (wheat, maize and rice) and three fruits (banana, mango and grape) revealed that the main differential metabolites in crops and fruits were vitamins, amino acids, flavonoids and lipids and identified complementary patterns of essential nutrients in crops and fruits (Shi et al., 2022). Through the analysis of metabolites in wolfberry and tomato, it was found that the typical markers of tomato were lycopene, carotene, glutamate and GABA, while the characteristic metabolites of wolfberry were lycibarbarphenylpropanoids and zeaxanthin esters (Dumont et al., 2020). However, few direct comparisons have been made between the metabolomes of wolfberry and other common species. In this work, we demonstrated that the metabolomes of monocots and dicots differed significantly. By comparing wolfberry with other common species, we identified more undiscovered specific metabolites in wolfberry, which may be closely linked to wolfberry's biological functions.

Numerous active ingredients have been identified in wolfberry. Lutein, zeaxanthin and carotene in wolfberry can reduce the risk of age-related macular degeneration (AMD) (Bertoldi et al., 2019). Together with vitamin B, they are responsible for the vision-improving effects of wolfberry (Kocyigit and Sanlier, 2017). Phenolic substances in wolfberry have been widely reported, and wolfberry has antioxidant, anti-inflammatory, neuroprotective, anticancer properties and intestinal microbiome regulatory effects (Ilić et al., 2020). In addition, LBPs, organic acids and carotenoids can also enhance the antioxidant capacity of wolfberry (Kocyigit and Sanlier, 2017). Changes in anthocyanin content are associated with the difference in fruit color of different wolfberry varieties and with the antioxidant capacity of the fruit (Zheng et al., 2011). In addition, wolfberry fruit also lowers blood sugar, blood fat and blood pressure, which is mainly due to the accumulation of amino acids, various monosaccharides and LBP in wolfberry (Potterat, 2010). In this study, we identified important active components in wolfberry. For example, betaine has been widely reported as an important active substance in wolfberry due to its antioxidant activity (Vidović et al., 2022). Our findings revealed that the concentration of betaine in wolfberry exceeds that of rice, maize, soybean, tomato, and grape by over 30 times, while its content in wheat is comparable to that in wolfberry (Supplementary Table S2). Besides, compared to those previously described in wolfberry, we discovered more metabolites that accumulate specifically in wolfberry, including N-acetylneuraminic acid, asparagine, 2-hydroxyisocaproic acid, 5-hydroxyhexanoic acid, phenyllactate, shikimic acid, echinacoside, esculin, isorhamnetin 3-O-neohesperidoside, narcissoside, scopoletin, cyanidin chloride, methyl p-coumarate, riboflavin, 4-amino-5-hydroxymethyl-2-methylpyrimidine and cholesterol. Of these, phenyllactate, N-acetylneuraminic acid, echinacoside, esculin, isorhamnetin 3-O-neohesperidoside, narcissoside and 4-amino-5-hydroxymethyl-2-methylpyrimidine were first identified as wolfberry-specific metabolites. Previous studies have shown that wolfberry can improve people's vision (Kocyigit and Sanlier, 2017). Riboflavin, often known as vitamin B2, has been shown to sustain proper visual capabilities in living beings (Kocyigit and Sanlier, 2017). Wolfberry contains high levels of riboflavin, which suggests that riboflavin is responsible for the fruit's ability to improve vision. Anthocyanins and organic acids in wolfberry have been reported to be associated with conferring antioxidant activity to wolfberry (Yan et al., 2018; Oğuz et al., 2021). In this study, we identified cyanidin chloride and four organic acids and sugars as wolfberry-specific metabolites, which may be significant anthocyanins and organic acids that confer antioxidant activity to wolfberry. Additionally, it has been reported that scopoletin has antioxidant activity in rats (Panda and Kar, 2006), and its specific accumulation in wolfberry may also be responsible for the antioxidant capacity of wolfberry.

The dissection of the genetic mechanism of important nutrient production in wolfberry is conducive to accelerating the process of plant breeding. Combining transcription and metabolism has become a widely used method to analyze the formation mechanisms of important metabolites in wolfberry. Using this research method, researchers have identified several key genes that regulate important metabolites of wolfberry through differences in transcription levels. For instance, the transcript level of *LbNCED1* positively regulates anthocyanin accumulation in wolfberry, thereby promoting fruit coloration. Nowadays, the genome has become a powerful tool for identifying functional genes (Zhao and Shi, 2022; Li et al., 2022a). The wolfberry reference genome is the first published reference genome of woody Solanaceae, which is beneficial for analyzing the genetic basis of wolfberry metabolites at the genome level. However, there is still a gap in the research analyzing the formation mechanism of wolfberry-specific metabolites at the genome level. In this work, we identified 16 wolfberry-specific metabolites, among which 10 exhibited a more distinct accumulation pattern in wolfberry. The concentration of these 10 metabolites in wolfberry was more than five times higher than in other species. Specifically, these metabolites include N-acetylneuraminic acid, 2-hydroxyisocaproic acid, 5-hydroxyhexanoic acid, phenyllactate, echinacoside, esculoside, isorhamnetin 3-O-neohesperidoside, narcissoside, scopoletin and riboflavin. Among them, the metabolic pathways of phenyllactate and riboflavin have been clearly analyzed in KEGG, we compared the copy number of the genes encoding key enzymes for the synthesis and degradation of specific metabolites in wolfberry and other species. Compared with other species, riboflavin and phenyllactate have more copies of key enzymes for synthesis and fewer copies of key enzymes for degradation in wolfberry, which may be the reason for their specific accumulation in wolfberry. Our work demonstrates that the formation of wolfberry-specific metabolites is controlled not only by the level of gene expression but also by the copy number of key genes that may lead to differences in metabolite synthesis pathways.

5 Conclusion

In this study, we compared the metabolome of wolfberry with that of six species, including the cereal crops maize, rice, wheat, legume crop soybean, and the fruit crops tomato and grape, and identified metabolites that accumulate specifically in wolfberry. Through high-throughput metabolomic analysis with widely targeted liquid chromatography-tandem mass spectrometry (LC-MS/MS), a total of 16 wolfberry-specific metabolites were identified, including seven phenylpropanoids, two amino acids and their derivatives, four organic acids and sugars, one lipid, one vitamin and coenzyme derivative, and one alkaloid. The phenyllactate degradation gene *UGT1* had the lowest copy number of the six species, whereas the riboflavin and phenyllactate synthesis genes *RFK* and *HPPR* had higher copy numbers than those of the other six species. This suggests that the copy numbers of *RFK*, *HPPR*, and *UGT1* may be the main reasons for the specific accumulation of riboflavin and phenyllactate in wolfberry. Moreover, the metabolome-based neighbor-joining tree showed that monocots and dicots clustered together separately, suggesting that metabolites could reflect the evolutionary relationship among different species. Taken together, we identified specific metabolites in wolfberry and provided new insight into the accumulation mechanism of species-specific metabolites at the genomic level.

Data availability statement

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

Author contributions

QL: Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Methodology. CZ: Writing – original draft, Methodology, Visualization, Writing – review & editing. HZ: Writing – original draft, Visualization. YZho: Writing – review & editing, Methodology. SL: Writing – review & editing, Visualization. YL: Writing – review & editing, Methodology. XM: Writing – review & editing. WA: Writing – review & editing, Resources. JuZ: Writing – review & editing. JiZ: Writing – review & editing, Funding acquisition. YZha: Writing – review & editing, Funding acquisition, Project administration, Supervision. CJ: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work is supported by the Key Research & Development Program of Ningxia Hui Autonomous Region [grant number 2022BBF01001 and 2021BEF02002], the National Natural Science Foundation of China [grant number U23A20221], the Innovative Research Group Project of Ningxia Hui Autonomous Region [grant number 2021AAC01001], and the Innovation Team for Genetic Improvement of Economic Forests [grant number 2022QCXTD04].

Acknowledgments

We are grateful for the reviewers' hard work and constructive comments, which allowed us to improve the quality of this manuscript.

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.2024.1392175/full#supplementary-material

References

Ahad, H., Jin, H., Liu, Y., Wang, J., Sun, G., Liang, X., et al. (2020). Chemical profiling of spermidines in goji berry by strong cation exchange solid-phase extraction (SCX-SPE) combined with ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-Q-TOF/MS/MS). *J. Chromatogr. B* 1137, 121923. doi: 10.1016/j.jchromb.2019.121923

Bertoldi, D., Cossignani, L., Blasi, F., Perini, M., Barbero, A., Pianezze, S., et al. (2019). Characterisation and geographical traceability of Italian goji berries. *Food Chem.* 275, 585–593. doi: 10.1016/j.foodchem.2018.09.098

Blasco, H., Błaszczyński, J., Billaut, J. C., Nadal-Desbarats, L., Pradat, P. F., Devos, D., et al. (2015). Comparative analysis of targeted metabolomics: Dominance-based rough set approach versus orthogonal partial least square-discriminant analysis. *J. Biomed. Inf.* 53, 291–299. doi: 10.1016/j.jbi.2014.12.001

Cao, Y., Li, Y., Fan, Y., Li, Z., Yoshida, K., Wang, J., et al. (2021). Wolfberry genomes and the evolution of *Lycium* (Solanaceae). *Commun. Biol.* 4, 671. doi: 10.1038/s42003-021-02152-8

Chen, D., Guo, S., Zhou, J., Zhu, Y., Zhang, F., Zeng, F., et al. (2021). Chemical constituents from *Lycium barbarum* (Solanaceae) and their chemophenetic significance. *Biochem. Syst. Ecol.* 97, 104292. doi: 10.1016/j.bse.2021.104292

Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., et al. (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* 46, 714–721. doi: 10.1038/ng.3007

Deng, M., Zhang, X., Luo, J., Liu, H., Wen, W., Luo, H., et al. (2020). Metabolomics analysis reveals differences in evolution between maize and rice. *Plant J.* 103, 1710–1722. doi: 10.1111/tpj.14856

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., et al. (2016). The dynamic N^1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530, 441–446. doi: 10.1038/nature16998

Dumont, D., Danielato, G., Chastellier, A., Hibrand Saint Oyant, L., Fanciullino, A., and Lugan, R. (2020). Multi-targeted metabolic profiling of carotenoids, phenolic compounds and primary metabolites in goji (*Lycium* spp.) berry and tomato (*Solanum lycopersicum*) reveals inter and intra genus biomarkers. *Metabolites* 10, 422. doi: 10.3390/metabo10100422

Ginestet, C. (2011). ggplot2: elegant graphics for data analysis. *J. R Stat. Soc. A Stat.* 174, 245–245. doi: 10.1111/j.1467-985X.2010.00676_9.x

Golbraikh, A., and Tropsha, A. (2002). Beware of q2! *J. Mol. Graph. Model.* 20, 269–276. doi: 10.1016/S1093-3263(01)00123-1

Huang, K., Yan, Y., Chen, D., Zhao, Y., Dong, W., Zeng, X., et al. (2020). Ascorbic acid derivative 2-*O*-β-$_D$-glucopyranosyl-$_L$-ascorbic acid from the fruit of *Lycium barbarum* modulates microbiota in the small intestine and colon and exerts an immunomodulatory effect on cyclophosphamide-treated BALB/c mice. *J. Agric. Food Chem.* 68, 11128–11143. doi: 10.1021/acs.jafc.0c04253

Ilić, T., Dodevska, M., Marčetić, M., Božić, D., Kodranov, I., and Vidović, B. (2020). Chemical characterization, antioxidant and antimicrobial properties of goji berries cultivated in Serbia. *Foods* 9, 1614. doi: 10.3390/foods9111614

Jin, M., Huang, Q., Zhao, K., and Shang, P. (2013). Biological activities and potential health benefit effects of polysaccharides isolated from *Lycium barbarum* L. *Int. J. Biol. Macromol.* 54, 16–23. doi: 10.1016/j.ijbiomac.2012.11.023

Kocyigit, E., and Sanlier, N. (2017). A review of composition and health effects of *Lycium barbarum*. *Int. J. Chin. Med.* 1, 1–9. doi: 10.11648/J.IJCM.20170101.11

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Software* 25, 1–18. doi: 10.18637/jss.v025.i01

Levin, R. A., and Miller, J. S. (2005). Relationships within tribe Lycieae (Solanaceae): paraphyly of *Lycium* and multiple origins of gender dimorphism. *Am. J. Bot.* 92, 2044–2053. doi: 10.3732/ajb.92.12.2044

Li, T., Fan, Y., Qin, H., Dai, G., Li, G., Li, Y., et al. (2020b). Transcriptome and flavonoids metabolomic analysis identifies regulatory networks and hub genes in black and white fruits of *Lycium ruthenicum* Murray. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.01256

Li, W., Liu, J., Zhang, H., Liu, Z., Wang, Y., Xing, L., et al. (2022a). Plant pan-genomics: recent advances, new challenges, and roads ahead. *J. Genet. Genomics* 49, 833–846. doi: 10.1016/j.jgg.2022.06.004

Li, G., Qin, B., Li, S., Yin, Y., Zhao, J., An, W., et al. (2020a). *LbNR*-derived nitric oxide delays lycium fruit coloration by transcriptionally modifying flavonoid biosynthetic pathway. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.01215

Li, Y., Yang, Z., Yang, C., Liu, Z., Shen, S., Zhan, C., et al. (2022b). The *NET* locus determines the food taste, cooking and nutrition quality of rice. *Sci. Bull.* 67, 2045–2049. doi: 10.1016/j.scib.2022.09.023

Li, G., Zhao, J., Qin, B., Yin, Y., An, W., Mu, Z., et al. (2019). ABA mediates development-dependent anthocyanin biosynthesis and fruit coloration in *Lycium* plants. *BMC Plant Biol.* 19, 317. doi: 10.1186/s12870-019-1931-7

Lu, J., Li, H., Quan, J., An, W., Zhao, J., and Xi, W. (2017). Identification of characteristic aroma volatiles of Ningxia goji berries (*Lycium barbarum* L.) and their developmental changes. *Int. J. Food Prop.* 20, S214–S227. doi: 10.1080/10942912.2017.1295254

Masci, A., Carradori, S., Casadei, M. A., Paolicelli, P., Petralito, S., Ragno, R., et al. (2018). *Lycium barbarum* polysaccharides: Extraction, purification, structural characterisation and evidence about hypoglycaemic and hypolipidaemic effects. A review. *Food Chem.* 254, 377–389. doi: 10.1016/j.foodchem.2018.01.176

Oğuz, İ., Oğuz, H.İ., and Kafkas, N. E. (2021). Evaluation of fruit characteristics of various organically-grown goji berry (*Lycium barbarum* L., *Lycium chinense* Miller) species during ripening stages. *J. Food Compost. Anal.* 101, 103846. doi: 10.1016/j.jfca.2021.103846

Panda, S., and Kar, A. (2006). Evaluation of the antithyroid, antioxidative and antihyperglycemic activity of scopoletin from *Aegle marmelos* leaves in hyperthyroid rats. *Phytother. Res.* 20, 1103–1105. doi: 10.1002/(ISSN)1099-1573

Poggioni, L., Romi, M., Guarnieri, M., Cai, G., and Cantini, C. (2022). Nutraceutical profile of goji (*Lycium barbarum* L.) berries in relation to environmental conditions and harvesting period. *Food Biosci.* 49, 101954. doi: 10.1016/j.fbio.2022.101954

Potterat, O. (2010). Goji (*Lycium barbarum* and *L.* chinense): phytochemistry, pharmacology and safety in the perspective of traditional uses and recent popularity. *Planta Med.* 76, 7–19. doi: 10.1055/s-0029-1186218

Shi, Y., Guo, Y., Wang, Y., Li, M., Li, K., Liu, X., et al. (2022). Metabolomic analysis reveals nutritional diversity among three staple crops and three fruits. *Foods* 11, 550. doi: 10.3390/foods11040550

Su, Q., Zhang, F., Xiao, Y., Zhang, P., Xing, H., and Chen, F. (2022). An efficient screening system to identify protein-protein or protein-DNA interaction partners of rice transcription factors. *J. Genet. Genomics* 49, 979–981. doi: 10.1016/j.jgg.2022.02.007

Sun, Y., Zhou, Y., Long, Q., Xing, J., Guo, P., Liu, Y., et al. (2024). *OsBCAT2*, a gene responsible for the degradation of branched-chain amino acids, positively regulates salt tolerance by promoting the synthesis of vitamin B5. *New Phytol.* 241, 2558–2574. doi: 10.1111/nph.19551

Thévenot, E. A., Roux, A., Xu, Y., Ezan, E., and Junot, C. (2015). Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. *J. Proteome Res.* 14, 3322–3335. doi: 10.1021/acs.jproteome.5b00354

Vidović, B. B., Milinčić, D. D., Marčetić, M. D., Djuriš, J. D., Ilić, T. D., Kostić, A.Ž., et al. (2022). Health benefits and applications of goji berries in functional food products development: A Review. *Antioxidants* 11, 248. doi: 10.3390/antiox11020248

Wang, C., Dong, Y., Zhu, L., Wang, L., Yan, L., Wang, M., et al. (2020). Comparative transcriptome analysis of two contrasting wolfberry genotypes during fruit development and ripening and characterization of the *LrMYB1* transcription factor that regulates flavonoid biosynthesis. *BMC Genomics* 21, 295. doi: 10.1186/s12864-020-6663-4

Xiao, J., Wei, X., Zhou, Y., Xin, Z., Miao, Y., Hou, H., et al. (2021). Genomes of 12 fig wasps provide insights into the adaptation of pollinators to fig syconia. *J. Genet. Genomics* 48, 225–236. doi: 10.1016/j.jgg.2021.02.010

Yan, Y., Peng, Y., Tang, J., Mi, J., Lu, L., Li, X., et al. (2018). Effects of anthocyanins from the fruit of *Lycium ruthenicum* Murray on intestinal microbiota. *J. Funct. Foods* 48, 533–541. doi: 10.1016/j.jff.2018.07.053

Yang, T., Hu, Y., Yan, Y., Zhou, W., Chen, G., Zeng, X., et al. (2022). Characterization and evaluation of antioxidant and anti-inflammatory activities of flavonoids from the fruits of *Lycium barbarum*. *Foods* 11, 306. doi: 10.3390/foods11030306

Zhang, Q., Chen, W., Zhao, J., and Xi, W. (2016). Functional constituents and antioxidant activities of eight Chinese native goji genotypes. *Food Chem.* 200, 230–236. doi: 10.1016/j.foodchem.2016.01.046

Zhao, D., Li, S., Han, X., Li, C., Ni, Y., and Hao, J. (2020a). Physico-chemical properties and free amino acids profiles of six wolfberry cultivars in Zhongning. *J. Food Compost. Anal.* 88, 103460. doi: 10.1016/j.jfca.2020.103460

Zhao, J., Li, H., Xi, W., An, W., Niu, L., Cao, Y., et al. (2015). Changes in sugars and organic acids in wolfberry (*Lycium barbarum* L.) fruit during development and maturation. *Food Chem.* 173, 718–724. doi: 10.1016/j.foodchem.2014.10.082

Zhao, J., Li, H., Yin, Y., An, W., Qin, X., Wang, Y., et al. (2020b). Fruit ripening in *Lycium barbarum* and *Lycium ruthenicum* is associated with distinct gene expression patterns. *Febs. Open Bio.* 10, 1550–1567. doi: 10.1002/2211-5463.12910

Zhao, W., and Shi, Y. (2022). Comprehensive analysis of phenolic compounds in four varieties of goji berries at different ripening stages by UPLC-MS/MS. *J. Food Compost. Anal.* 106, 104279. doi: 10.1016/j.jfca.2021.104279

Zheng, J., Ding, C., Wang, L., Li, G., Shi, J., Li, H., et al. (2011). Anthocyanins composition and antioxidant activity of wild *Lycium ruthenicum* Murr. from Qinghai-Tibet Plateau. *Food Chem.* 126, 859–865. doi: 10.1016/j.foodchem.2010.11.052

Zhu, A., Zhou, Q., Hu, S., Wang, F., Tian, Z., Hu, X., et al. (2022). Metabolomic analysis of the grain pearling fractions of six bread wheat varieties. *Food Chem.* 369, 130881. doi: 10.1016/j.foodchem.2021.130881

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 June 2024
DOI 10.3389/fpls.2024.1408202

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Zhengkun Qiu,
South China Agricultural University, China
Ping Chen,
Chinese Academy of Agricultural Sciences, China
Feng Xinkang,
Chinese Academy of Agricultural Sciences, China, in collaboration with reviewer PC

*CORRESPONDENCE
Andrew C. Allan
andrew.allan@plantandfood.co.nz

RECEIVED 27 March 2024
ACCEPTED 03 June 2024
PUBLISHED 20 June 2024

CITATION
Martinez-Sanchez M, Hunter DA, Saei A, Andre CM, Varkonyi-Gasic E, Clark G, Barry E and Allan AC (2024) *SmuMYB113* is the determinant of fruit color in pepino (*Solanum muricatum*). *Front. Plant Sci.* 15:1408202. doi: 10.3389/fpls.2024.1408202

SmuMYB113 is the determinant of fruit color in pepino (*Solanum muricatum*)

Marcela Martinez-Sanchez[1,2], Donald A. Hunter[3], Ali Saei[4], Christelle M. Andre[1], Erika Varkonyi-Gasic[1], Glen Clark[1], Emma Barry[1] and Andrew C. Allan[1,2]*

[1]The New Zealand Institute for Plant and Food Research Limited (Plant & Food Research) Mt Albert, Auckland, New Zealand, [2]School of Biological Sciences, University of Auckland, Auckland, New Zealand, [3]The New Zealand Institute for Plant and Food Research Limited (Plant & Food Research), Palmerston North, New Zealand, [4]Grasslands Research Centre, AgResearch Limited, Palmerston North, New Zealand

Pepino (*Solanum muricatum*) is an herbaceous crop phylogenetically related to tomato and potato. Pepino fruit vary in color, size and shape, and are eaten fresh. In this study, we use pepino as a fruit model to understand the transcriptional regulatory mechanisms controlling fruit quality. To identify the key genes involved in anthocyanin biosynthesis in pepino, two genotypes were studied that contrasted in foliar and fruit pigmentation. Anthocyanin profiles were analyzed, as well as the expression of genes that encode enzymes for anthocyanin biosynthesis and transcriptional regulators using both RNA-seq and quantitative PCR. The differential expression of the transcription factor genes R2R3 MYB *SmuMYB113* and R3MYB *SmuATV* suggested their association with purple skin and foliage phenotype. Functional analysis of these genes in both tobacco and pepino showed that SmuMYB113 activates anthocyanins, while SmuATV suppresses anthocyanin accumulation. However, despite elevated expression in all tissues, *SmuMYB113* does not significantly elevate flesh pigmentation, suggesting a strong repressive background in fruit flesh tissue. These results will aid understanding of the differential regulation controlling fruit quality aspects between skin and flesh in other fruiting species.

KEYWORDS

plant model, *Solanum muricatum*, anthocyanins, MYB, transcriptomics

Introduction

Pepino (*Solanum muricatum* Aiton, 2n=24), is an herbaceous crop native to the High Andes region of South America (Anderson et al., 1996). This crop is in the same genus as tomato (*S. lycopersicum* L.) and potato (*S. tuberosum* L.) and is phylogenetically closely related to other major solanaceous fruit crops such as pepper (*Capsicum annuum* L.) and eggplant (*Solanum melongena* L.) (Spooner et al., 1993; Rodríguez-Burruezo et al., 2011;

Sarkinen et al., 2013). However, it has a number of advantageous characteristics, including propagation and growing practices. Pepino plants are easy to grow and propagate, have a relatively short life cycle, simple genetics (diploid), and a recently sequenced genome (Song et al., 2022). Pepino has a tendency to produce parthenocarpic fruit (Prohens et al., 2005), and the fruit need over 90 days to reach rich full ripeness. Pepino plants are sensitive to high temperatures during pollination, which can significantly reduce fruit set (Burge, 1989) and affect ripening, fruit quality and taste (Rodríguez-Burruezo et al., 2011).

Pepino fruit can be round, ellipsoid or elongated (Herraiz et al., 2015), and have attractive characteristics for consumers, such as brightly colored skin with stripes in some genotypes, intense aroma and a yellow juicy flesh with a mild sweet taste (Rodríguez-Burruezo et al., 2011; Herraiz et al., 2016). Fruit are generally consumed when fully ripe in the same way as melon, as a dessert fruit, although they are notably less sweet (Prohens et al., 2005). Another use for the unripe fruit is to cut and use in a similar way to cucumber. The pepino fruit is recognized for its beneficial attributes for human health (Hsu et al., 2011; (Shathish and Guruvayoorappan, 2014). The relatively large variation in appearance, anthocyanin concentration, soluble solids content, acidity, and vitamin C content (Rodríguez-Burruezo et al., 2011 (Prohens et al., 2003), Herraiz et al., 2016), suggests that improvement can be made via selection and breeding (Rodríguez-Burruezo et al., 2011), with breeding goals such as enhanced yield, reduced ripening time, improved heat tolerance and increased fruit quality.

Anthocyanins are flavonoids synthesized by plants and along with other pigments such as betalains and carotenoids, are responsible for the red to blue color in leaves, flowers and fruits (Holton and Cornish, 1995; Grotewold, 2006; Petroni and Tonelli, 2011). They help attract pollinators and seed distributors (Davies et al., 2012) and also play a critical role in protecting the plant against abiotic stresses including drought (Castellarin et al., 2007; André et al., 2009), UV radiation and cold temperatures (Christie et al., 1994; Sarma and Sharma, 1999). In addition, anthocyanins have been identified as powerful antioxidants and anti-inflammatory agents in the human body (Kähkönen and Heinonen, 2003; Miguel, 2011).

Anthocyanin biosynthesis and regulatory pathways have been extensively characterized in many plant species (Holton and Cornish, 1995; Petroni and Tonelli, 2011). The co-ordinated expression of what has been termed 'early' biosynthetic genes: chalcone synthase (CHS), chalcone isomerase (CHI) and flavone3-hydroxylase (F3H); and what has been termed 'late' biosynthetic genes encoding flavonoid 3'5'-hydroxylase ((F3'5'H), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), flavonoid 3-*o*-glucosyltransferase (3GT), anthocyanin rhamnosyltransferase (RT), anthocyanin acyltransferase (AAC), flavonoid 5-O-glucosyltransferasese (5GT), and glutathione *S*-transferase (GST) generate anthocyanins. This transcriptional regulation is by DNA-binding R2R3 MYB transcription factors (TFs), MYC-like basic helix-loop-helix (bHLH) TFs, and WD40-repeat proteins in the MYB-bHLH-WD40 (MBW) complex (Baudry et al., 2004; Allan et al., 2008; Wang et al., 2020).

Expression of the genes encoding the MBW complex is activated or repressed by environmental and developmental factors which then drive expression of downstream targets (Allan and Espley, 2018). It is often a MYB transcription factor (TF) that is the limiting factor in this response. The R2R3 MYBs participate in regulating both early and late genes of the anthocyanin biosynthetic pathway (Schwinn et al., 2006; Dubos et al., 2010). These activators (often MYB TFs sub-group 6) have been characterized in many species including *Arabidopsis* (Borevitz et al., 2000; Gonzalez et al., 2008), and many others including maize (Paz-Ares et al., 1987), potato (Jung et al., 2005, 2009; Zhang et al., 2009), tomato (Mathews et al., 2003), pepper (Borovsky et al., 2004), sweet potato (Chu et al., 2013), petunia (Quattrocchio et al., 1999), grapevine (Kobayashi et al., 2005; Walker et al., 2007; Deluc et al., 2008), Bayberry (Liu et al., 2013), and apple (Takos et al., 2006; Ban et al., 2007; Espley et al., 2007). However, there is apparently strong repression or restrictions on these activators, as certain tissues accumulate anthocyanins while other tissues do not.

In contrast, R2R3 MYB TFs of subgroup 4 have been identified as negative regulators of the anthocyanin biosynthetic pathway (LaFountain and Yuan, 2021) in several species, including strawberry (Aharoni et al., 2001), snapdragon (Tamagnone et al., 1998), apple (Lin-Wang et al., 2011), grapevine (Cavallini et al., 2015), Arabidopsis (Jin et al., 2000), petunia (Albert et al., 2011; Albert et al., 2014a) and ginkgo (Xu et al., 2014). Subgroup 4 repressors actively repress transcription of their target genes, mediated by repressive motifs (e.g. ERF Amphiphilic Repression – EAR, and TLLLFR) within their C-termini (Tamagnone et al., 1998; Aharoni et al., 2001; Matsui et al., 2008). These R2R3 MYB repressors are co-repressors acting within the MBW complex, binding to the bHLH proteins, and can convert an MBW complex into one that inhibits transcription (Matsui et al., 2008; Albert et al., 2014a, 2014b).

R3 MYBs contain a single imperfect repeat and are therefore classed within the 1R-MYB-like subclass. However, these MYBs are unable to bind DNA directly, as they lack a transcriptional activation domain present in R2R3 MYB TFs, and therefore these R3 MYBs affect transcription of their target genes with other partners such as bHLHs (Du et al., 2015). The R3 repeat allows the R3 MYB to interact with bHLH partners, and several have been shown to act as inhibitors of transcriptional regulation, by competing with R2R3 MYBs for the MBW complex (Matsui et al., 2008; Xu et al., 2015). In Arabidopsis, AtCPC competes with the activator MYB, AtPAP1, for the formation of the MBW complex (Zhu et al., 2009). In tomato, the R3 MYB SlATV inhibits the expression of the structural genes responsible for anthocyanin biosynthesis, leading to a repression of anthocyanin production (Cao et al., 2017).

Anthocyanin biosynthesis and regulation has been widely characterized in solanaceous species (Albert et al., 2011; Albert et al., 2014a; Kiferle et al., 2015; Montefiori et al., 2015; Liu et al., 2016, 2018; Colanero et al., 2020; Sun et al., 2020; Yan et al., 2020; Zhou et al., 2022). However, less is known about anthocyanin production in pepino, although key genes have been reported in the recently published genome (Song et al., 2022). The profile of

phenolic acids, flavonoids, and anthocyanins in aqueous extracts of pepino fruit was reported (Hsu et al., 2011) with (Wu et al., 2013) identifying eight phenolic compounds and one flavonoid in the fruit. The diversity of plant color within pepino fruit, including green, yellow and purple-skinned varieties, as well as its close phylogenetic relationship to potato and tomato, makes pepino an ideal candidate for further investigation and modeling of anthocyanin biosynthesis in the Solanaceae, as well as fruiting plant crops outside the Solanaceae family. In pepino, the variation in phenolic and polyphenolic compounds (including flavonoids and total anthocyanins) has been studied (Anderson et al., 1987; Hsu et al., 2011; Herraiz et al., 2016; Hsu et al., 2018). Also, by comparing wild pepino and a cultivated pepino, Song et al. (2022) used RNAseq aligned to their constructed genome, to study the anthocyanin pathway and its regulation. Here, we studied pepino lines with different fruit and foliar pigmentation, identified the anthocyanin profile present in pepino skin, and used transcriptomics to determine the controlling transcription factors for anthocyanin biosynthesis. Transient and stable transformation showed that pepino R2R3 MYB113 (AN1-like) is a key transcriptional activator for pepino anthocyanin accumulation. However, a strong repressive background is suggested in the fruit flesh, which prevents anthocyanin accumulation in this tissue.

Materials and methods

Plant material

A purple-skinned pepino (*S. muricatum)* genotype, termed Purple Selection (PS, purple skin and yellow flesh) and a yellow-skinned pepino selection, Yellow Selection (YS, yellow skin and yellow flesh) were propagated by cuttings in a greenhouse at Plant & Food Research, Auckland, New Zealand. After 56 days, rooted cuttings were transferred into 30-L pots and grown in a greenhouse at Plant & Food Research, Pukekohe, New Zealand, at ambient conditions. Ten to thirteen pepino fruit from separate plants were collected at 20, 34, 48 and 73 days after flowering (DAF), and pooled from at least three plants (each stage of fruit development had three biological replicates). Skin tissue was carefully separated from flesh tissue, with any residual flesh tissue removed (pepino has a very strong skin layer) and frozen separately in liquid nitrogen and stored at −80°C.

For functional gene testing, *Nicotiana tabacum* (tobacco) and pepino plants were grown under glasshouse conditions between 20 to 24°C using natural daylight with extension to 16 h at Plant & Food Research, Auckland, New Zealand.

Anthocyanin identification and quantification

Pepino extracts were analyzed with a Waters Acquity® Ultra Performance Liquid Chromatography (UPLC) system (Milford, MA, USA) equipped with a photodiode array detector (PDA) and a single quadrupole mass spectrometer (QDa, Waters, MA, USA). The separation of the 5-µL aliquot was performed on a reverse-phase Acquity UPLC HSS T3 column (2.1 × 100 mm, 1.8 µm particle size, Waters, MA, USA). The eluents were 0.1% formic acid in water (A), and 0.1% formic acid in acetonitrile (B). The gradient was as follows: 0 min, 5% B; 5 min, 10% B; 10 min, 20% B; 12 min, 100% B; 14 min, 100% B; 14.5 min, 10% B; 18 min, 10% B. The flow rate was 0.5 mL min^{-1} and the column temperature was 50°C. Anthocyanins were detected at 520 nm and quantified as petunidin-3-glucoside equivalents using a six-point calibration curve. Furthermore, a validation standard was injected after every tenth injection.

Transcriptome analysis using RNA-sequencing

Total RNA was extracted using the Spectrum Plant Total RNA kit® (Sigma-Aldrich). Three biological replicates were used. Library construction, sequencing, mapping to reference genome, and quality control were performed at Australian Genome Research Facility (AGRF; www.agrf.org.au) (Illumina Stranded mRNA with 150 paired end reads and sequencing to 20M depth). The primary sequence data were generated using the Illumina DRAGEN BCL Convert 07.021.624.3.10.8 pipeline. Data yield for the 48 samples from the pipeline ranged from 6.91Gb to 12.7 Gb, with mean yield of 9.3 Gb. Methods at AGRF followed the strict requirements of the International standard ISO17025 for quality control, testing and calibration. The per-base sequence quality for the 48 samples was excellent, with >87% bases above Q30 across all samples. Reads were mapped using the STAR aligner (v2.5.3a) to the *Solanum muricatum* genome downloaded from http://songlab.bio2db.com/pepino.html. Mapping QC results are presented in Supplementary Figure 1. Counts were normalized using the Relative Log Expression method in the DESeq2 package (Love et al., 2014). Principal component analysis was performed using common variance on the top 500 most variable genes, to check consistency of the biological replicates. The transcriptome response of the three biological replicates was similar, allowing differences in genotype, tissue type and developmental stage to be examined (Supplementary Figure 2).

Phylogenetic analysis

Sequence alignment was performed using Geneious Prime® (version 2022.0.1) Muscle alignment with eight iterations and the phylogenetic tree was constructed using the Neighbor-Joining method with 1000 bootstrap replicates.

Real time qPCR expression analysis

Total RNA of skin and flesh of the two pepino genotypes, PS and YS, was extracted using the Spectrum Plant Total RNA Kit (Sigma -Aldrich, St. Louis, MO, USA). Removal of genomic DNA and first strand cDNA synthesis was carried out using the

QuantiTect Reverse Transcription kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Real-time qPCR DNA amplification and analysis was carried out using the LightCycler 480 Real-Time PCR System (Roche Diagnostics, Switzerland), with LightCycler 480 software version 1.5. The LightCycler 480 SYBR Green I Master Mix (Roche) was used. Each reaction volume was 10 µL; reactions were run with four replicates and a water control was included in each run. qPCR conditions were as follows: 5 min at 95°C, followed by 40 cycles of 5 s at 95°C, 5 s at 60°C, and 10 s at 72°C, followed by 65–95°C melting curve detection. The qPCR efficiency of each gene was obtained by analyzing the raw fluorescence data using the LinRegPCR (2014.x) software (Ruijter et al., 2009). Used primer sequences are listed in Supplementary Table 1. These were normalized to elongation factor-1 of pepino (Sm06G02378).

Gene cloning and sequence analysis

Sequences of candidate genes were identified by BLAST (Altschul et al., 1997) from the *S. muricatum* pepino genome (Song et al., 2022) and PCR-amplified from leaf cDNA using specific primers and Phusion® High Fidelity DNA polymerase (Thermo Fisher Scientific, USA). PCR products of anthocyanin-related regulatory genes (*SmMYB113* and *SmATV*) and genes that encode enzymes of the anthocyanin pathway (*SmDFR* and *SmF3'5'H*) were cloned into pENTR-TOPO (ThermoFisher) followed by In-Fusion cloning (Takara Bio) into pHEX2. Clones were confirmed by sequencing (Macrogen).

Transient assays of gene function

Transient assays were performed in tobacco as previously reported (Hellens et al., 2005). Approximately 300 µL of *Agrobacterium tumefaciens* GV3101 culture containing genes of interest were infiltrated into a young *N. tabacum* leaf. Final color development was evaluated by taking digital photographs four days after the infiltration.

Stable transformation of pepino

SmMYB113 was transformed into etiolated hypocotyls of pepino (*S. muricatum*) yellow-skinned selection (YS) following a previously described transformation protocol (Horsch et al., 1985) adapted to use hypocotyls as explants. Briefly, etiolated hypocotyls were cut into 5-mm long pieces. Explants were inoculated with 10 mL *Agrobacterium* culture for 10 min and blotted dry with sterile filter paper. Inoculated explants were transferred onto co-cultivation medium for 3 days in the dark. After co-cultivation, explants were transferred onto regeneration medium (MS basal salts and vitamins (Duchefa, Haarlem, The Netherlands), 3% sucrose, 0.7% agar (Invitrogen, Waltham, MA, USA), 1 mg/L BAP + 0.1 mg/L NAA and 300 mg/L Timentin) and selection media (MS basal salts and vitamins (Duchefa), 3% sucrose, 0.7% agar (Germantown, 0.1 mg/L NAA and 300 mg/L Timentin) to induce adventitious shoots from callus tissue. Selection was by the inclusion of 150 mg/L kanamycin in the medium. A single shoot was selected from each explant and transferred onto elongation medium which also included 150 mg/L kanamycin. Three to four weeks later shoots started to develop roots. Once roots were established, rooted plants were transferred into potting mix and moved to glasshouse conditions between 20 to 24°C using natural daylight with extension to 16 h.

Results

Anthocyanins are responsible for differences in pepino fruit pigmentation

Two pepino selections that differ in anthocyanin pigmentation of the fruit skin, denoted purple skin (PS) or yellow skin (YS), were chosen for analysis. They produce fruit of comparable size, shape, growth rate and yellow flesh color (Figures 1A, B). Anthocyanins were absent from the skin and flesh of the YS genotype, throughout fruit development. Both YS and PS genotypes can produce anthocyanins in leaves and petals, although PS has much darker purple foliage and petals (Supplementary Figure 3). Total anthocyanin concentrations were evaluated throughout the fruit development in the PS genotype (Figure 1C). The total amount of anthocyanin was significantly higher at stage 4 than in stages 1 to 3, reaching 1.7 mg g^{-1} (dry weight or DW) at stage 4.

Anthocyanin composition was first determined in the skin of the PS genotype by Ultra Performance Liquid Chromatography-High-Resolution Mass Spectrometry (UPLC-HR/MS) and Ultra Performance Liquid Chromatography- diode Array Detector (UPLC-DAD) and compared with available data from the literature. Four predominant anthocyanin compounds were detected in the profiles (Figure 1D). Peak 2 represented about 60% of total anthocyanins and was identified as petunidin-3-*p*-coumaroyl-rutinoside-5-glucoside (Figure 1D; Supplementary Figure 4). The presence of this anthocyanin has previously been reported in other fruits of the Solanaceae family, such as purple tomatoes (*Solanum lycopersicum* L.) (Silva Souza et al., 2020), and purple-fleshed potato tubers (*Solanum tuberosum* L.) (Andre et al., 2007). When anthocyanins are acylated, both cis- and trans-isomers can be found (Saha et al., 2020). Considering the predominance of trans isomers in *Solanum* species, it is likely that peak 2 is petunidin-3-trans-coumaroyl-rutinoside-5-glucoside (also known as petanin), whereas peak 3 is its cis- stereoisomer.

The other anthocyanin compounds followed the same substitution pattern: -3-acylrutinoside-5-glucoside and exhibited a similar anthocyanidin fragment at m/z 317, indicative of petunidin derivatives. Consequently, peaks 1 and 4 were identified as petunidin-3-caffeoyl-rutinoside-5-glucoside and petunidin-3-feruloyl-rutinoside-5-glucoside, respectively.

There were no significant differences in the proportions of the different anthocyanins at the four developmental stages investigated, suggesting that there is a common regulatory system for all four anthocyanin compounds.

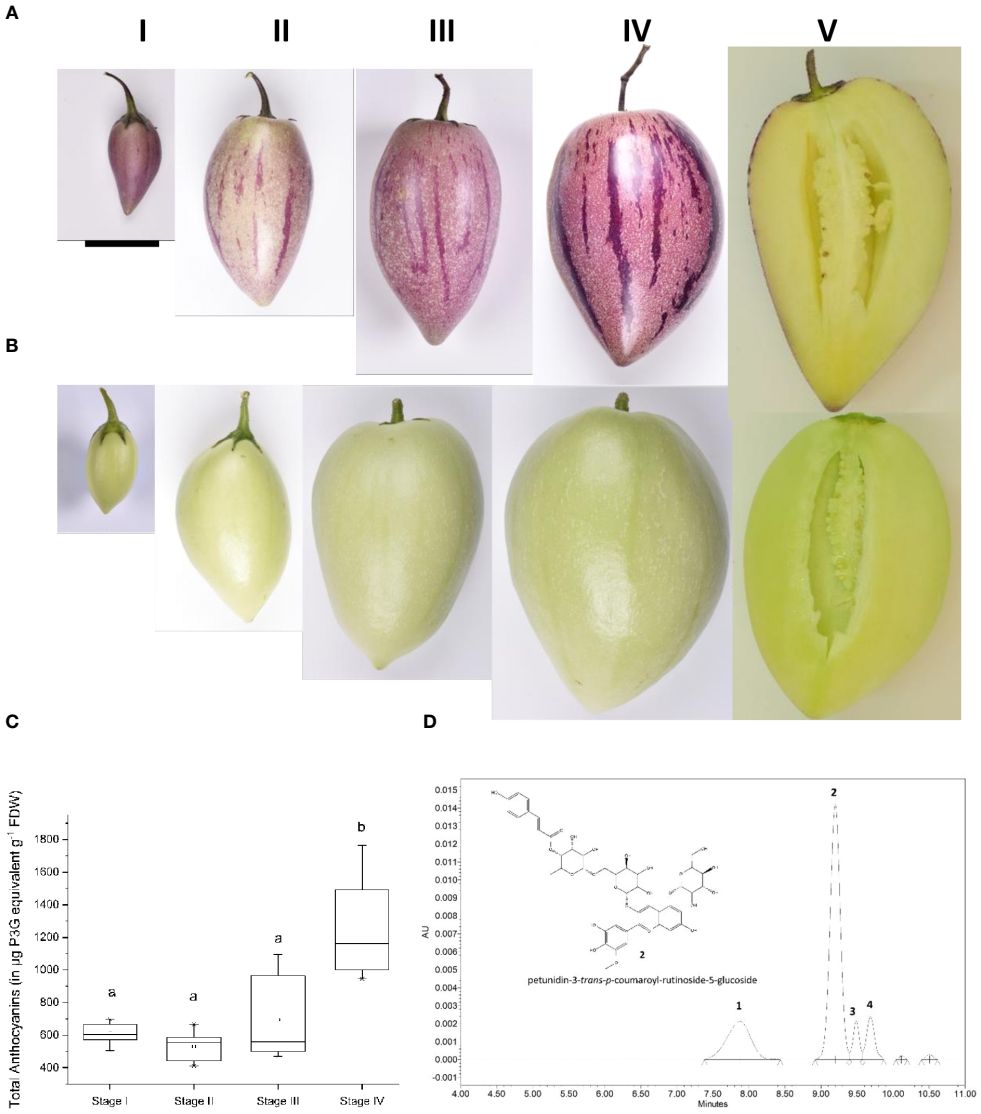


FIGURE 1
Stages of pepino fruit development. I (20 Days after flowering, DAF), II (34 DAF), III (48 DAF), IV (73 DAF) for the purple-skinned **(A)** selection and the yellow-skinned **(B)** selection. **(C)** Total anthocyanin content in the purple-skinned selection of pepino at four different stages of development (n =9). Data were expressed in µg of petunidin-3-glucoside equivalent (P3G) per gram of freeze-dried weight (FDW). Stages with different letters are significantly different at the p <0.05 level. Anthocyanins are absent in the yellow-skinned cultivar. **(D)** UPLC-PDA chromatogram of pepino skin (Stage IV) recorded at 520 nm, with molecular structure of the predominant anthocyanin 2 identified as petunidin-3-*trans*-*p*-coumaroyl-rutinoside-5-glucoside. Scale bar on **(A)** is equivalent to 4 cm.

RNAseq analysis of the anthocyanin biosynthetic and regulator genes

To study the transcripts associated with anthocyanin accumulation in pepino, skin and flesh samples were collected from PS and YS genotypes over the four developmental stages and subjected to RNA-seq. A comparative analysis between purple and yellow tissues at the same developmental stage was performed and differentially expressed genes (DEGs) with the minimum of $\log_2$-fold change in expression (DESeq, adjusted p-value<0.05, mean count 100) between the purple and yellow cultivars were identified. This analysis revealed 1540 up-regulated DEGs and 1467 down-regulated DEGs in skin, and 1524 up-regulated DEGs and 1269 down-regulated DEGs in flesh, when PS was compared with YS. Of these genes, there was an enrichment for candidate genes related to anthocyanin biosynthesis or regulation of this pathway.

Previously, 17 pepino regulatory genes and 11 biosynthetic enzymes have been identified as participating in anthocyanin biosynthesis (Song et al., 2022). Most of these gene models were differentially expressed between YS and PS (Figure 2). In our study, a total of 25 anthocyanin-related DEGs were consistently up-regulated, with four down-regulated in PS skin samples compared with YS (Figure 2). Four genes were DEGs but had variable expression across skin developmental stages. In flesh, 13 and seven DEGs were consistently up- or down-regulated, respectively and when comparing the DEGs in skin and flesh, eight biosynthetic genes were up-regulated (*SmuPAL6, SmuPAL7, SmuPAL10, Smu4CL3, Smu4CL4, SmuCHI3, SmuF3'H and SmuUGT*). Also,

10.3389/fpls.2024.1408202

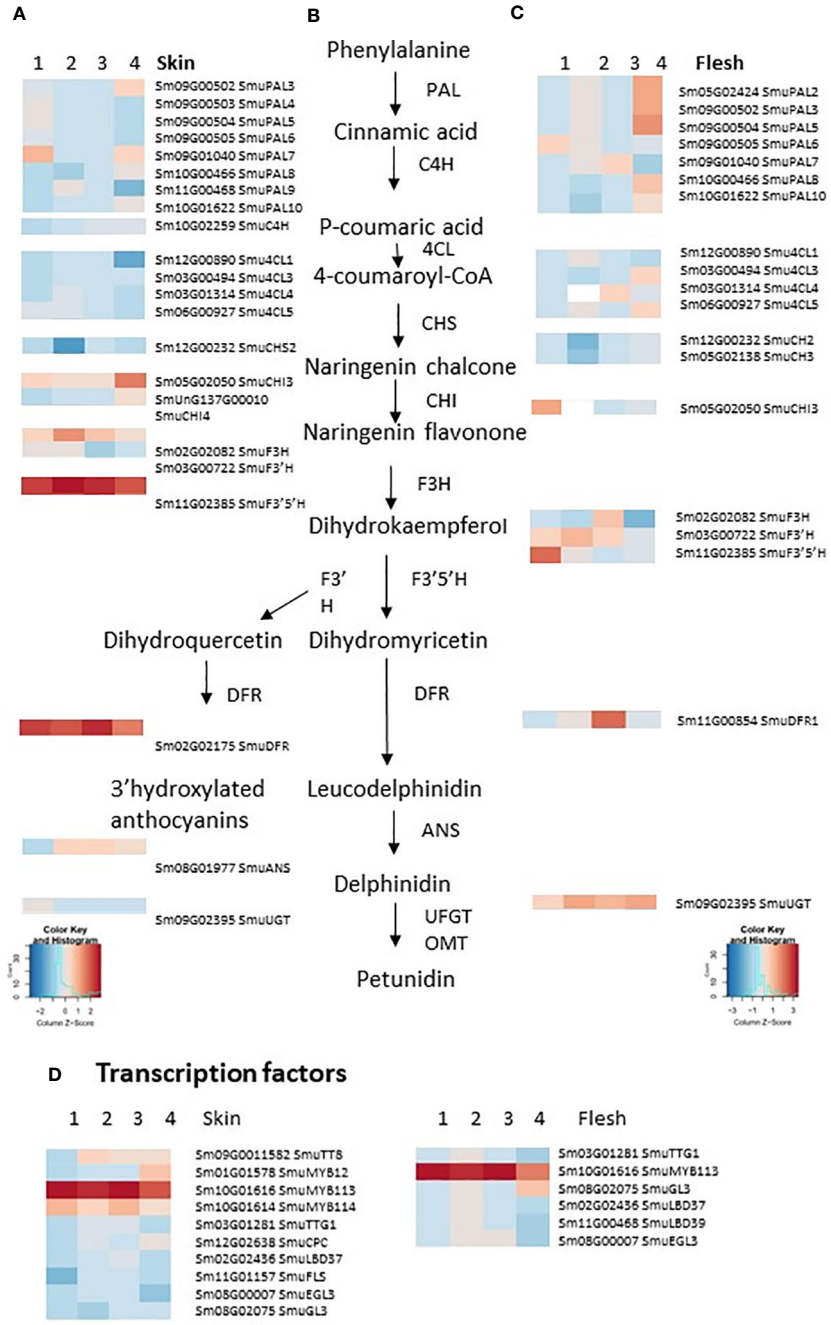


FIGURE 2
Heatmaps showing the differential expression of anthocyanin biosynthesis, regulatory and transport genes in the purple-skinned selection vs yellow-skinned selection pepino skin **(A)** and flesh **(C)**, in four stages of development. An abbreviated version of the anthocyanin biosynthetic pathway is provided **(B)**. **(D)** Transcription factors implicated in regulating this pathway. The color of any given cell represents the normalized read values, with red representing high expression levels and blue representing low expression levels, with two color keys separate for skin and flesh. Genes are annotated according to Song et al. (2022).

the anthocyanin-related R2R3 MYB regulator *SmuMYB113* was up-regulated in both PS tissues, whereas the bHLH SmuTT8 was up-regulated in skin only. In contrast, *Smu4CL1* and *SmuCHS2* (both anthocyanin biosynthesis-related genes) were down-regulated in both skin and flesh tissues. In addition, five genes were DEGs but had variable expression across flesh developmental stages.

Transcript reads of *SmuDFR2* (Sm02G02175) were significantly higher in the purple-skinned genotype than in the yellow-skinned cultivar (Figure 3). Reads of less than 10 per stage were present in flesh of both PS and YS (Figures 3A, B). A similar trend was seen with SmuF3'5'H (Sm11G02385), where transcripts were significantly higher in the PS genotype than in the YS cultivar.

10.3389/fpls.2024.1408202

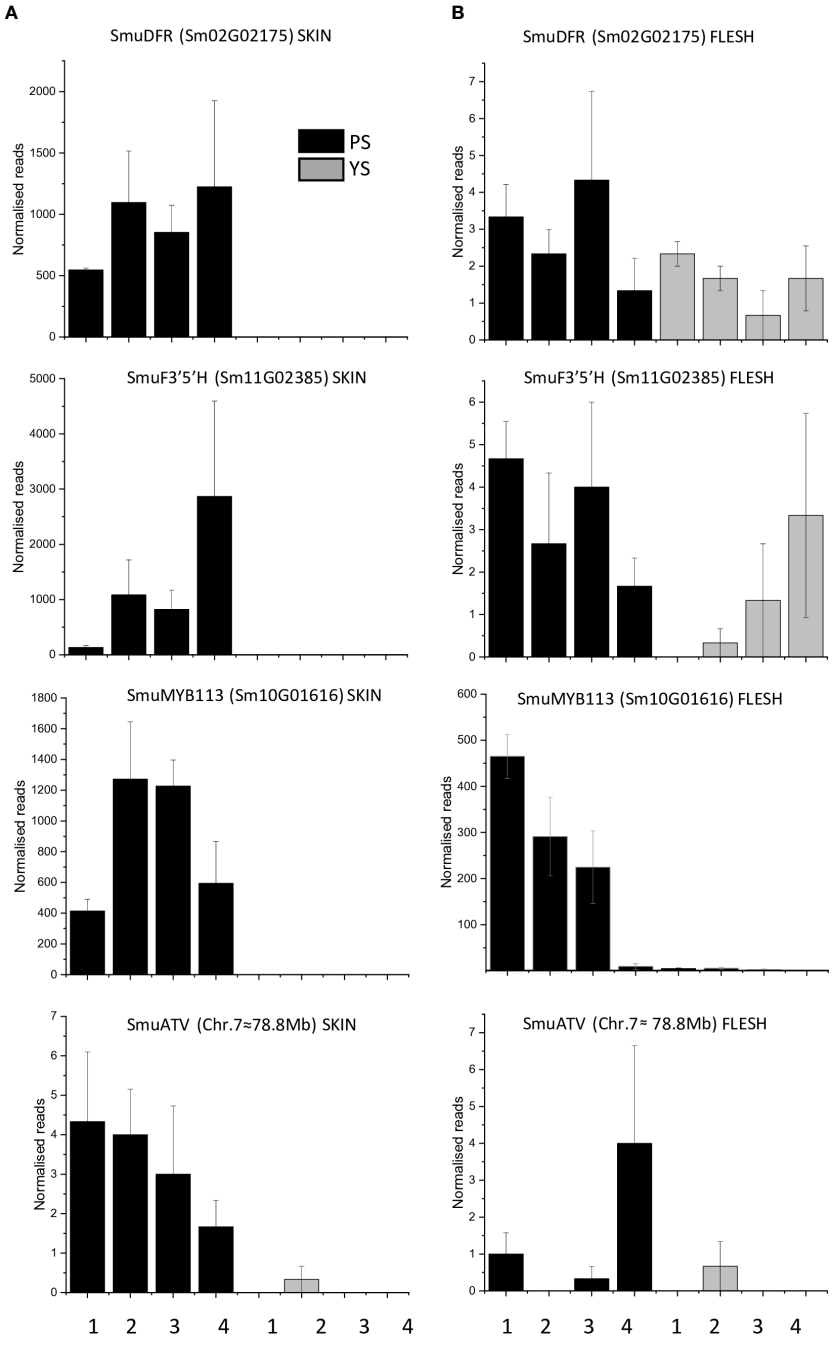


FIGURE 3
Transcript levels in skin **(A)** and flesh **(B)** of anthocyanins biosynthetic genes (SmuDFR and SmuF3'5'H) and regulatory genes (SmuMYB113 and SmuATV) in the skin and flesh of two pepino selections at four developmental stages. Error bars represent standard errors of the means (STE) of three biological replicates.

Reads peaked at stage 4, which correlated with the highest concentrations of anthocyanins (Figure 1). In the flesh, reads for SmuF3'5'H were again less than 5 normalized reads per stage.

An examination of candidate transcriptional regulators of anthocyanin-related enzymes was made. There were four candidate MYB transcription factors in the DEG lists, as well as a bHLH, which is homologous (69% identity) to AtTT8 (Figure 2A) and was previously annotated as SmuTT8 (Song et al., 2022) (Table 1). This *TT8*-like bHLH was highly expressed in PS skin, but not as well expressed in PS flesh. However, the R2R3 MYB *SmMYB113* was consistently up-regulated, in both skin and flesh of the purple-skinned selection. Transcript counts were high in all purple fruit tissues (Figure 3B). In contrast, reads were absent or very low in YS fruit tissues.

Candidate repressors of anthocyanins, such as R3 MYBs and LOB domain TFs, did not explain the phenotypes, as they were not up-regulated either in YS, or in PS flesh tissues (Figure 2). In addition, a homologue of the tomato anthocyanin repressor *SlATV*, which we termed *SmATV*, was identified by blast match in the

pepino genome (chr.7, 78.8Mb). Reads mapping to this gene model (previously un-annotated) were low, but present in both skin and flesh (Figures 3A, B). Furthermore, a flavonol-related subgroup 7 MYB (*SmuMYB12*) was expressed in skin of both selections (Table 1), and therefore increased flux of precursor substrates away from anthocyanins does not appear be the cause of the YS phenotype.

Phylogenetic analysis of *SmuMYB113, SmuATV* and MYB repressors of the anthocyanin pathway

Phylogenetic analysis of these MYBs revealed close similarities between SmuMYB113 and StAN1 of potato, and SmuMYB114, and SlANT1 of tomato (Figure 4). In potato, StAN1 is a key regulator of skin coloration (Payyavula et al., 2013). Although differentially expressed, SmMYB114 showed low expression in pepino (Table 1). Therefore, both expression and sequence similarity suggest SmuMYB113 is the causative controller of pepino skin anthocyanin content. However, this does not explain the flesh phenotype, where high levels of SmMYB113 expression are seen, yet no anthocyanins accumulate. In this phylogenetic tree, *SmuATV* clustered with a group of R3 MYB TFs. Within this group, *SmuATV* was closely related to *PhMYBx* and *SlMYBATV*, both negative regulators of anthocyanin biosynthesis (Kroon, 2004; Albert et al., 2011; Albert et al., 2014b).

qPCR analysis validation of candidate genes

Expression of anthocyanin-related genes was verified using RT-qPCR. Accumulation of transcripts of anthocyanin biosynthetic genes *SmuDFR2* and *SmuF3'5'H* was confirmed in purple skin, but

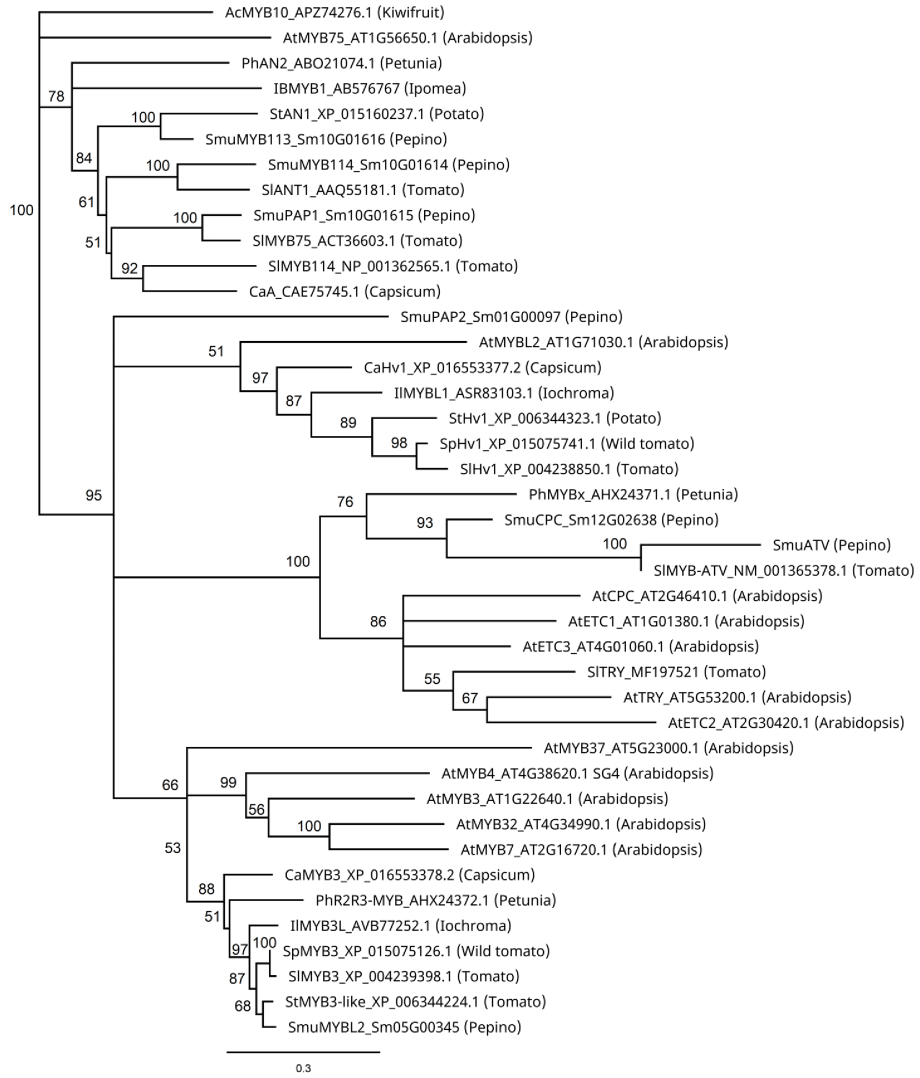


FIGURE 4
Phylogenetic relationship analysis of pepino MYBs and known anthocyanin MYB regulators from other species. Sequences were aligned using predicted protein sequences in Geneious (version Geneious Prime® 2022.0.1) Muscle alignment with eight iterations, and the phylogenetic tree was constructed using the Neighbor-Joining method with 1000 bootstrap replicates.

these were absent from yellow skin and flesh at all studied stages of fruit development (Figure 5). An increase at 34 DAF for *SmuDFR* was followed by a decline at later fruit developmental stages, while transcripts of *SmuF3'5'H* increased at 34 DAF and peaked at 73 DAF.

The expression of *SmuMYB113* and *SmuATV* in flesh and skin was examined using qPCR. For *SmuMYB113*, the purple-skinned stage 2 of development showed the highest value, followed by a decline. For yellow-skinned (YS) pepino, the values were significantly lower in all stages of development, in both skin and flesh. However, in the flesh of the purple-skinned genotype, high levels of expression were seen at three of the four stages. This confirms patterns seen with RNAseq, and suggests a repressive mechanism exists in the pepino flesh. Expression levels of the potential repressor *SmuATV* were low in the purple- and yellow-skinned cultivars. However, PS had higher levels of expression generally than YS.

As *SmuMYB113* appears to be a key regulatory difference between the purple-skinned and yellow-skinned cultivars, further correlations were made between its expression and other putative target genes within the phenylpropanoid and flavonoid pathways. Early steps in the anthocyanin pathway, such as *PAL, C4H, 4CL, CHS*, and *F3'H*, were all well expressed in yellow skin and flesh tissues of both selections (Table 2). This is despite the fact that no anthocyanin accumulated in yellow skin or flesh, or in the flesh of the PS selection. In skin tissues, there was a strong positive correlation between *SmuMYB113*/anthocyanin and expression of *CHI3, F3'5'H, DFR*, and the *GST*, *TT19*. An exception was the previously annotated UFGT (Sm09G02395) (Song et al., 2022), which showed high expression in all fruit tissues. By blasting the Arabidopsis anthocyanidin 3-O-glucosyltransferase (At5g17050) and grape UFGT (Ford et al., 1998), better matching pepino gene models were retrieved, such as Sm10G01884 (Table 2), which was expressed in a skin-specific manner, especially in PS tissue. We therefore hypothesize that the absence of *SmuMYB113* in the yellow-skinned selection of pepino results in a loss of expression of later steps in the anthocyanin pathway. This would also lead to a loss of expression of the bHLH partner, *SmTT8*, which is known to be part of a hierarchy of regulation in solanaceous plants (Montefiori et al., 2015). However, this does not explain the presence of *SmuMYB113* expression in the flesh of the purple-skinned selection, where there is no elevation of expression of these later biosynthetic steps, or indeed of the co-regulator *SmTT8*.

Functional assays of SmuMYB113 and SmuATV in tobacco

To functionally test the R2R3 MYB SmuMYB113 and R3 MYB SmuATV, their full-length coding regions were cloned, and placed under the control of the *35S* promoter. These sequences were

TABLE 1 Normalized reads of candidate transcription factors in pepino selections.

gene model	Description	PS skin	YS skin	PS flesh	YS flesh	Fold change Up in PS vs YS[a]	
		Total normalized reads across all stages				Skin	Flesh
Sm12G00739	SmuMYB11	0	0	0	0		
Sm01G01578	SmuMYB12	95	97	0	0		
Sm06G02794	SmuMYB111	0	0	0	0		
Sm10G01616	SmuMYB113	3511	0	998	12	11.6	6.0
Sm10G01614	SmuMYB114	76	1	22	0		
Sm10G01615	SmuPAP1	0	0	0	0		
Sm01G00097	SmuPAP2	0	0	3	0		
Sm09G01582	SmuTT8/bHLH	1099	93	47	10	2.8	
Sm08G02075	SmuGL3/bHLH	1350	1145	1036	888		
Sm08G00007	SmuEGL3/bHLH	3535	4149	4297	4679		
Sm03G01281	SmuTTG1/WD40	37	40	40	35		
Sm05G00345	SmuMYBL2	33	0	0	0		
Sm12G02638	SmuCPC	489	256	508	345	1.6	
SmATV	SmuATV	13	0	6	1		
Sm02G02436	SmuLBD37	956	729	94	58		
Sm01G03283	SmuLBD38	1823	1795	1596	1713		
Sm11G00468	SmuLBD39	554	307	844	823		

[a]Data presented only if fold change ≥2, p value ≥0.05 and mean count over samples ≥50.
PS, purple-skinned; YS, yellow-skinned.

10.3389/fpls.2024.1408202

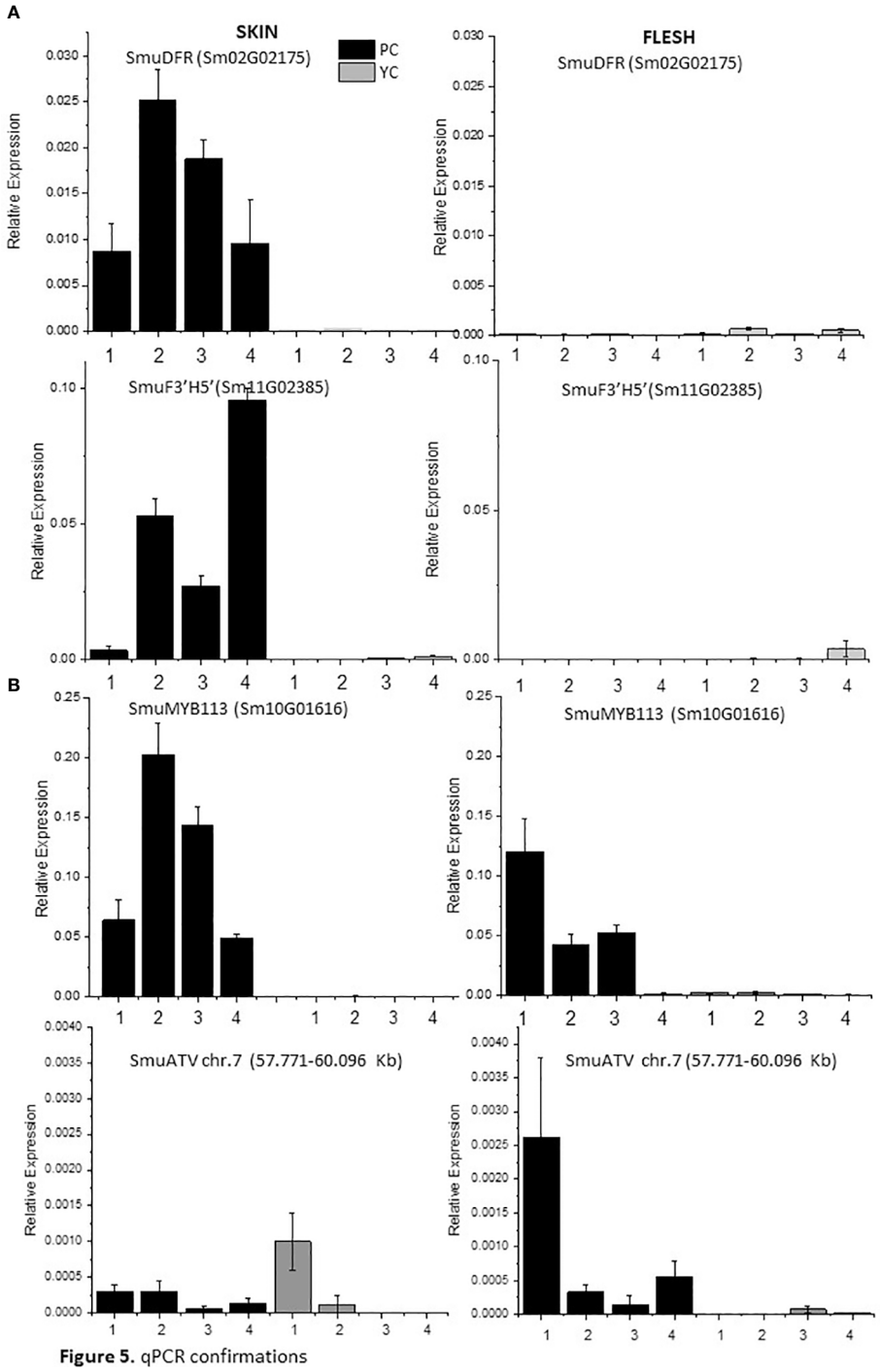


Figure 5. qPCR confirmations

FIGURE 5
Expression analysis of anthocyanin biosynthetic genes SmuDFR and SmuF3′5′H and anthocyanin regulatory genes SmuMYB113 and SmuATV, across four developmental stages of two pepino selections. **(A)** Relative expression levels of biosynthetic genes SmuDFR and SmuF3′5′H and regulatory genes, and **(B)** regulatory genes SmuMYB113 and SmuATV in purple-skinned and yellow-skinned selection skins.

infiltrated using *Agrobacterium* into tobacco leaves (Espley et al., 2007). Infiltration of *SmuMYB113* activated anthocyanin production in tobacco leaves (Figure 6). No anthocyanin production was observed with the negative control GUS construct. The potato *StAN1* (NCBI protein accession number: AKA95392) construct (Liu et al., 2016) was infiltrated as a positive control and resulted in strong activation of anthocyanin accumulation (Figure 6). In contrast, co-infiltration of *StAN1* with *SmuATV* significantly reduced anthocyanin accumulation compared with StAN1 alone. Similarly, when *SmuATV* was co-infiltrated with *SmMYB113*, there was no anthocyanin accumulation, suggesting strong repression of the R2R3 MYB by this R3 MYB. These findings support the hypothesis that *SmuATV* acts as a repressor of anthocyanin accumulation.

10.3389/fpls.2024.1408202

TABLE 2 Normalized reads of candidate biosynthetic-related genes in pepino selections.

gene model	Description	PS skin	YS skin	PS flesh	YS flesh	Fold change Up in PS vs YS[a]	
		Normalized reads across all stages				Skin	Flesh
Sm10G01622	SmuPAL10	9025	4105	11067	2420	1.5	1.1
Sm09G00502	SmuPAL3	11584	5056	7332	7129	1.6	
Sm09G00503	SmuPAL4	8246	4366	1818	2460		
Sm09G00504	SmuPAL5	8475	5022	2159	3176		
Sm09G00505	SmuPAL6	14	10	12	3		
Sm10G02259	SmuC4H	8794	4433	1957	1338		
Sm03G00494	Smu4CL3	2244	1352	725	444		
Sm03G01314	Smu4CL4	673	416	37	32		
Sm06G00927	Smu4CL5	6096	2973	4379	4094		
Sm05G02050	SmuCHI3	328	12	18	3		
SmUnG137G00010	SmuCHI4	1600	1249	1232	1420	1.1	
Sm05G02138	SmuCHS	30706	26715	33015	56022		
Sm02G02082	SmuF3H	7273	235	126	417	4.9	
Sm03G00722	SmuF3'H	874	631	665	329		0.9
Sm11G02385	SmuF3'5'H	4870	1	13	5	11.0	
Sm02G02175	SmuDFR	7846	3	20	3	11.1	
Sm08G01977	SmuANS	2644	1	17	4	10.2	
Sm09G02395	SmuUFGT	514	183	1086	129	1.1	2.9
Sm10G01884	SmuUFGT-like	4011	190	1	0	4.8	
Sm02G01868	SmuTT19	5962	6	5	1	10.2	

[a]Data presented only if fold change ≥0.5, p value ≥0.05 and mean count over samples ≥ 50.
PS, purple-skinned; YS, yellow-skinned.

Functional characterization of SmMYB113 in transformed pepino lines

To characterize the effect of SmMYB113 on pepino anthocyanin biosynthesis, we transformed pepino (YS genotype) with a 35S: *SmMYB113* overexpression construct. Three independent lines were produced that showed intense purple pigmentation in leaves and flowers (Figures 7A, B). Fruit developed normally and had visibly enhanced purple color (Figure 7C). These fruits were tested by qPCR, which showed increased expression of the *SmuMYB113* gene in both fruit skin and flesh (Figure 7D). A correlation between transgene expression and fruit skin and flesh color intensity was apparent, with transgenic pepino Line 1 showing the highest *SmuMYB113* transcript levels and the darkest skin and flesh (Figures 7C, D). In contrast, *SmuATV* transcript was higher in the skin and flesh of a control fruit than in Lines 1 and 3, although increased *SmuATV* was detected in the skin of Line 2. Transcript abundance of *SmuTT8* was also elevated in the transgenic lines, compared with the untransformed YS control. In the flesh of YS wild-type lines, there were very few reads of TT8 in the RNA seq dataset (Table 1). However, when MYB113 was overexpressed as a cDNA transgene, TT8 transcript was induced. Over-expression of *SmuMYB113* therefore changes a YS phenotype into a plant with purple-skinned fruit, although the flesh color was still not as intense as expected from a strong promoter. This is further evidence that a strong repressive background is not fully overcome by over-expression of the R2R3 MYB.

Discussion

Enhancing fruit quality, particularly by providing higher vitamin and antioxidant contents, can provide numerous health benefits. For this reason, accumulation of compounds such as anthocyanins have been studied in many plants. Anthocyanin biosynthesis has been widely characterized in solanaceous species (Albert et al., 2011; Kiferle et al., 2015; Montefiori et al., 2015; Liu et al., 2016; Sun et al., 2020; Yan et al., 2020; Zhou et al., 2022). In tomato, R2R3 MYB genes *SlANT1, SlANT1like* and *SlAN2* have been identified as positive anthocyanin regulators (Kiferle et al., 2015; Colanero et al., 2020). Also in tomato, the R3 MYB gene, *SlATV*, has been shown to be a negative regulator for anthocyanin

10.3389/fpls.2024.1408202

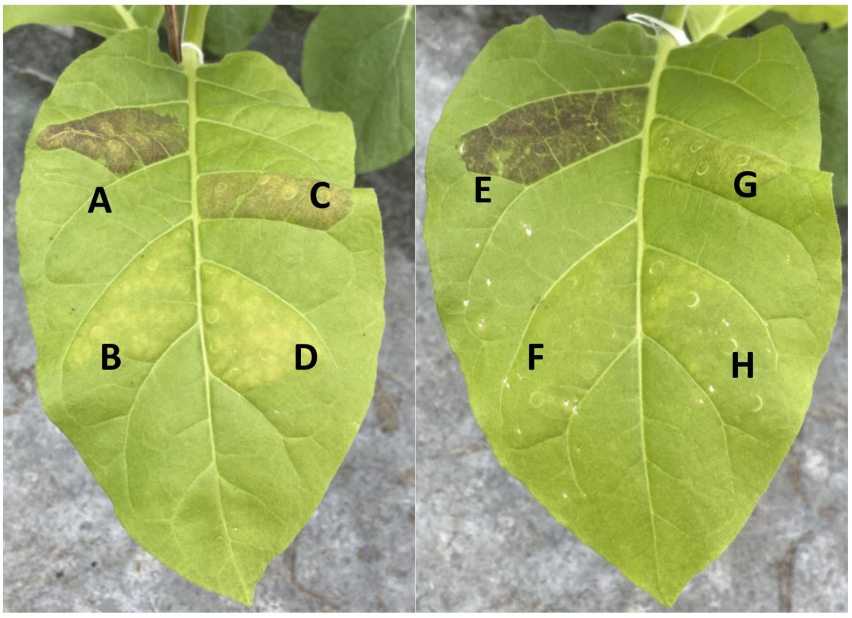


FIGURE 6
SmuMYB113 and SmuATV regulate anthocyanin accumulation in *Nicotiana tabacum* leaves. *N. tabacum* tobacco leaves were infiltrated with *Agrobacterium* strains expressing anthocyanin activators *StAN1* and *SmuMYB113*, or the R3-MYB repressor *SmuATV* is various combinations. *Agrobacterium* carrying a plasmid with GUS-only was included as a negative control and to maintain equivalent proportions of each strain. Anthocyanin accumulation (visibly assessed) 4 d post-infiltration. **(A)** StAN1 + GUS, **(B)** GUS only **(C)** StAN1 + SmuATV **(D)** SmuATV with GUS, **(E)** SmuMYB113 + GUS, **(F)** GUS only **(G)**. SmuMYB113 + SmuATV and **(H)**. SmuATV (+ GUS).

biosynthesis (Yan et al., 2020). In pepino, key anthocyanin-related genes have been reported (Song et al., 2022), as well as several potential transcription factors regulating the anthocyanin pathway, but these are yet to be characterized.

In this current study we characterized two candidate genes: *SmuMYB113*, encoding an R2R3-MYB TF, and *SmuATV*, encoding an R3-MYB TF. Our results suggest that expression of *SmuMYB113* is the key difference between the two evaluated pepino selections, PS and YS. This is based on RNA-seq and qPCR data, as well as tests of gene function in both tobacco and pepino transgenics. A previously un-annotated MYB, *SmuATV*, was shown to be a strong repressor of anthocyanin accumulation, so this could be a candidate for inhibiting accumulation of anthocyanins in the pepino flesh.

Although expression of *MYB113* appears to be a key difference between purple-skinned and yellow-skinned pepinos, there is still expression in flesh of purple-skinned cultivars where no anthocyanins accumulate. In the *35S:SmuMYB113* transformants, which is a yellow-skinned selection, flesh expression of the transgene was high but once again flesh anthocyanins did not accumulate to a high concentration. In the close relative, eggplant (*Solanum melongena*), it has been reported that eggplant *MYB113* significantly increased the anthocyanin and flavonol content in both peel and fruit pulp (Yang et al., 2022).

The RNAseq dataset revealed a large number of differentially expressed transcripts, obtained from comparing the tissues of pepino skin and flesh (PS vs PY) at four stages of development. Genes encoding early anthocyanin biosynthetic enzymes (*PAL, C4H, 4CL, CHI* and *CHS*) were well expressed in fruit tissues even in YS, and in the flesh of PS (Table 2). However, the late stages of the anthocyanin pathway (*F3H, F3'5'H, DFR, ANS, GST*) were low in expression in tissues with low anthocyanin content. The transcript levels of *SmuF3'5'H* and S*muDFR* were validated using qPCR and were significantly up regulated in PS skin, compared with PS flesh, and with YS skin and flesh. These results align with the observed distribution pattern of anthocyanins in both the skin and flesh of the pepino fruit (Figure 1C). The exception to this is the expression of *UFGT* (Sm09G02395), which showed high expression levels in yellow and purple tissues (Table 2). However, re-examination of the pepino genome suggested better candidate genes for *UFGT* (Sm10G01884) that did corelate with anthocyanin contents (Table 2). Recently transcriptomic and metabolomic analysis of ramie (*Boehmeria nivea*) was also used to identify UFGT as a key gene in foliage anthocyanin production (Feng et al., 2021).

Candidate regulatory genes were examined in the RNAseq datasets. The transcript of *SmuMYB113* was upregulated in PS skin and flesh in the four stages of development. Its potential partner, *SmuTT8*, was expressed in PS skin, but not highly expressed in flesh. Strong over-expression of *SmuMYB113* activates more expression of *SmuTT8* in transgenics and triggers some activation of the anthocyanin pathway in flesh. We observed no increase in expression of candidate repressors that could explain the lack of activation of *SmuMYB113* in YS, or the lack of anthocyanins in PS flesh. Cao et al. (2017), suggested that in tomato, *SlATV* inhibits the expression of the genes responsible for anthocyanin biosynthesis, leading to a repression in anthocyanin production. Transcript analysis from the pepino RNA-seq data and RT-qPCR, showed that *SmuATV* was expressed in PS skin and flesh. Although

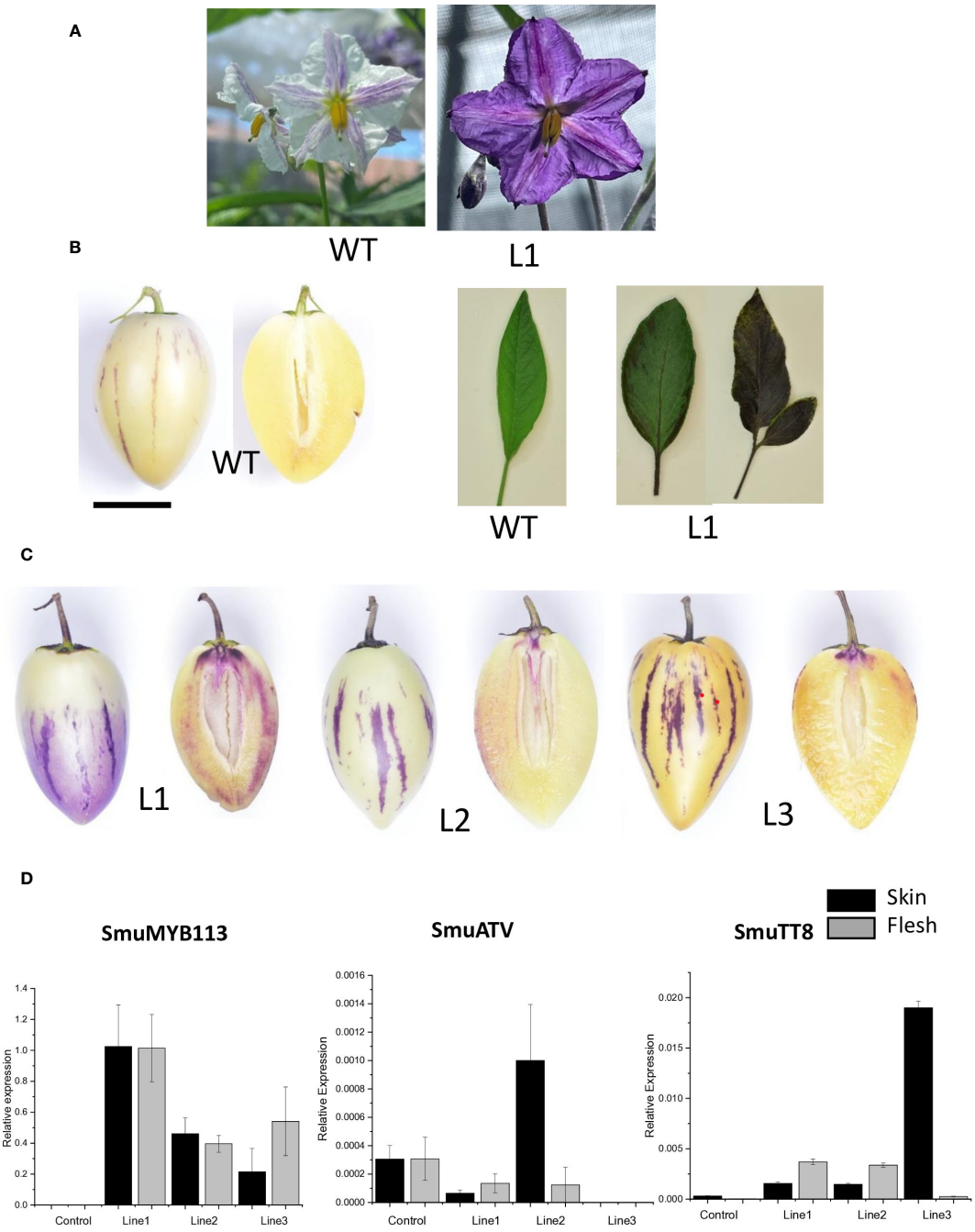


FIGURE 7
Overexpression of SmuMYB113 induced anthocyanin accumulation in flowers, leaves and fruit skin and flesh in pepino plants. **(A)** Wild-type (WT) flower phenotype and line overexpressing SmuMYB113 phenotype (L1). **(B)** Left: WT pepino fruit phenotype. Right: WT leaf phenotype and line over-expressing SmuMYB113 phenotype (L1). **(C)** Transformed lines (L1, L2 and L3) showing anthocyanin accumulation in skin and flesh. **(D)** Transcript levels in skin and flesh of anthocyanin-regulatory genes SmuMYB113 and SmuATV of three independent lines overexpressing SmuMYB113.

SmuATV plays a repressive role in regulating the activity of SmMYB113 from functional assays (Figure 6), its expression did not correlate with loss of anthocyanins in pepino fruit.

Over-expression of *SmMYB113* in the yellow-skinned pepino selection complemented the phenotype, producing a purple foliage, petals and fruit skin. However, flesh anthocyanin concentrations were still low. This further suggests a strong repressive background in the pepino flesh, despite good expression of the SmuMYB113 transgene. In apple, Espley et al. (2007) identified the MYB component, MdMYB10, as the principal factor for strong flesh pigment accumulation in the MBW complex. However, in solanaceous plants this does not appear to occur. Albert et al. (2014a) showed that petunia MYB AN1-like genes are normally regulated by the MBW complex, as part of a feed-forward mechanism, while Montefiori et al. (2015) identified that a hierarchy of bHLHs is needed for the tobacco MYB activators to bind promoters of the anthocyanin pathway. In YS there is no expression of TT8 in the flesh. However, in lines overexpressing

MYB113, some TT8 transcript was induced. Interestingly, skin pigmentation was striped in both purple-fruited cultivars and in transgenics over-expressing *SmuMYB113*. Other levels of regulation, such as degradation of transcripts by RNA processing, or rapid turnover of MYB protein, may well explain why pepino flesh remains relatively un-pigmented or why uneven pigmentation occurs in the skin. If these conundrums can be addressed, considerable improvement in total fruit anthocyanins will be achievable, which is a trait useful in many fruits and vegetables.

Data availability statement

The RNAseq data presented in this study is deposited in the NCBI repository, BioProject PRJNA1122205.

Author contributions

MM-S: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. DH: Data curation, Formal Analysis, Software, Writing – original draft, Writing – review & editing. AS: Data curation, Formal analysis, Investigation, Software, Writing – original draft, Writing – review & editing. CA: Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. EV-G: Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. GC: Methodology, Project administration, Resources, Writing – original draft, Writing – review & editing. EB: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. AA: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, 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 New Zealand Ministry of Business, Innovation and Employment (MBIE) Endeavour fund C11X2101. MM-S is supported by Plant & Food Research and the University of Auckland.

Acknowledgments

We sincerely thank Dr. Richard Espley and Dr. Nick Albert for critically reading the manuscript and providing constructive edits and Geeta Chhiba for media preparation.

Conflict of interest

Authors MM-S and AA are employed by Plant & Food Research and the University of Auckland.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1408202/full#supplementary-material

SUPPLEMENTARY FIGURE 2
Principal component analysis using common variance on the top 500 most variable genes in pepino.

References

Aharoni, A., De Vos, C. H. R., Wein, M., Sun, Z., Greco, R., Kroon, A., et al. (2001). The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco: Strawberry Myb represses flavonoid biosynthesis. *Plant J. Cell Mol. Biol.* 28, 319–332. doi: 10.1046/j.1365-313X.2001.01154.x

Albert, N. W., Davies, K. M., Lewis, D. H., Zhang, H., Montefiori, M., Brendolise, C., et al. (2014a). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell* 26, 962–980. doi: 10.1105/tpc.113.122069

Albert, N. W., Davies, K. M., and Schwinn, K. E. (2014b). Gene regulation networks generate diverse pigmentation patterns in plants. *Plant Signal Behav.* 9, e29526. doi: 10.4161/psb.29526

Albert, N. W., Lewis, D. H., Zhang, H., Schwinn, K. E., Jameson, P. E., and Davies, K. M. (2011). Members of an R2R3-MYB transcription factor family in Petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant J. Cell Mol. Biol.* 65, 771–784. doi: 10.1111/j.1365-313X.2010.04465.x

Allan, A. C., and Espley, R. V. (2018). MYBs drive novel consumer traits in fruits and vegetables. *Trends Plant Sci.* 23, 693–705. doi: 10.1016/j.tplants.2018.06.001

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., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Anderson, G. J., Jansen, R. K., and Kim, Y. (1996). The origin and relationships of the pepino, Solanum muricatum (Solanaceae): DNA restriction fragment evidence. *Econ. Bot.* 50, 369–380. doi: 10.1007/BF02866519

Anderson, G. J., Steinharter, T. P., and Cooper-Driver, G. (1987). Foliar flavonoids and the systematics of Solanum sect. Basarthrum. *Systematic Bot.* 12, 534–540. doi: 10.2307/2418888

Andre, C. M., Oufir, M., Guignard, C., Hoffmann, L., Hausman, J.-F., Evers, D., et al. (2007). Antioxidant profiling of native andean potato tubers (Solanum tuberosum L.) reveals cultivars with high levels of β-carotene, α-tocopherol, chlorogenic acid, and petanin. *J. Agric. Food Chem.* 55, 10839–10849. doi: 10.1021/jf0726583

André, C. M., Schafleitner, R., Legay, S., LefeèVre, I., Alvarado Aliaga, C. A., Nomberto, G., et al. (2009). Gene expression changes related to the production of phenolic compounds in potato tubers grown under drought stress. *Phytochem. (Oxford)* 70, 1107–1116. doi: 10.1016/j.phytochem.2009.07.008

Ban, Y., Honda, C., Hatsuyama, Y., Igarashi, M., Bessho, H., and Moriguchi, T. (2007). Isolation and functional analysis of a MYB transcription factor gene that is a key regulator for the development of red coloration in apple skin. *Plant Cell Physiol.* 48, 958–970. doi: 10.1093/pcp/pcm066

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

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12, 2383–2393. doi: 10.1105/tpc.12.12.2383

Borovsky, Y., Oren-Shamir, M., Ovadia, R., De Jong, W., and Paran, I. (2004). A locus that controls anthocyanin accumulation in pepper encodes a MYB transcription factor homologous to Anthocyanin2 of Petunia. *Theor. Appl. Genet.* 109, 23–29. doi: 10.1007/s00122-004-1625-9

Burge, G. K. (1989). Fruit-set in the pepino (Solanum-muricatum Ait). *Scientia Hortic.* 41, 63–68. doi: 10.1016/0304-4238(89)90050-2

Cao, X., Qiu, Z., Wang, X., Van Giang, T., Liu, X., Wang, J., et al. (2017). A putative R3 MYB repressor is the candidate gene underlying atroviolacium, a locus for anthocyanin pigmentation in tomato fruit. *J. Exp. Bot.* 68, 5745–5758. doi: 10.1093/jxb/erx382

Castellarin, S. D., Pfeiffer, A., Sivilotti, P., Degan, M., Peterlunger, E., and DI Gaspero, G. (2007). Transcriptional regulation of anthocyanin biosynthesis in ripening fruits of grapevine under seasonal water deficit. *Plant Cell Environ.* 30, 1381–1399. doi: 10.1111/j.1365-3040.2007.01716.x

Cavallini, E., Matus, J. T., Finezzo, L., Zenoni, S., Loyola, R., Guzzo, F., et al. (2015). The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine. *Plant Physiol. (Bethesda)* 167, 1448–1470. doi: 10.1104/pp.114.256172

Christie, P. J., Alfenito, M. R., and Walbot, V. (1994). Impact of low-temperature stress on general phenylpropanoid and anthocyanin pathways: Enhancement of transcript abundance and anthocyanin pigmentation in maize seedlings. *Planta* 194, 541–549. doi: 10.1007/BF00714468

Chu, H., Jeong, J. C., Kim, W.-J., Chung, D. M., Jeon, H. K., Ahn, Y. O., et al. (2013). Expression of the sweetpotato R2R3-type IbMYB1a gene induces anthocyanin accumulation in Arabidopsis. *Physiologia plantarum* 148, 189–199. doi: 10.1111/j.1399-3054.2012.01706.x

Colanero, S., Tagliani, A., Perata, P., and Gonzali, S. (2020). Alternative splicing in the anthocyanin fruit gene encoding an R2R3 MYB transcription factor affects anthocyanin biosynthesis in tomato fruits. *Plant Commun.* 1, 100006–100006. doi: 10.1016/j.xplc.2019.100006

Davies, K. M., Albert, N. W., and Schwinn, K. E. (2012). From landing lights to mimicry: the molecular regulation of flower colouration and mechanisms for pigmentation patterning. *Funct. Plant Biol.* 39, 619–638. doi: 10.1071/FP12195

Deluc, L., Bogs, J., Walker, A. R., Ferrier, T., Decendit, A., Merillon, J.-M., et al. (2008). The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. *Plant Physiol. (Bethesda)* 147, 2041–2053. doi: 10.1104/pp.108.118919

Du, H., Liang, Z., Zhao, S., Nan, M.-G., Tran, L.-S. P., Lu, K., et al. (2015). The evolutionary history of R2R3-MYB proteins across 50 eukaryotes: new insights into subfamily classification and expansion. *Sci. Rep.* 5, 11037–11037. doi: 10.1038/srep11037

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

Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-Amma, S., and Allan, A. C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10: A MYB transcription factor controlling apple fruit colour. *Plant J. Cell Mol. Biol.* 49, 414–427. doi: 10.1111/j.1365-313X.2006.02964.x

Feng, X., Gao, G., Yu, C., Zhu, A., Chen, J., Chen, K., et al. (2021). Transcriptome and metabolome analysis reveals anthocyanin biosynthesis pathway associated with ramie (Boehmeria nivea (L.) Gaud.) leaf color formation. *BMC Genomics* 22, 1–684. doi: 10.1186/s12864-021-08007-0

Ford, C. M., Boss, P. K., and Hoj, P. B. (1998). Cloning and characterization of Vitis vinifera UDP-glucose:flavonoid 3-O-glucosyltransferase, a homologue of the enzyme encoded by the maize Bronze-1 locus that may primarily serve to glucosylate anthocyanidins *in vivo*. *J. Biol. Chem.* 273, 9224–9233. doi: 10.1074/jbc.273.15.9224

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J. Cell Mol. Biol.* 53, 814–827. doi: 10.1111/j.1365-313X.2007.03373.x

Grotewold, E. (2006). *The Science of Flavonoids* (New York, NY: Springer-Verlag). doi: 10.1007/978-0-387-28822-2

Hellens, R. P., Allan, A. C., Friel, E. N., Bolitho, K., Grafton, K., Templeton, M. D., et al. (2005). Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* 1, 13–13. doi: 10.1186/1746-4811-1-13

Herraiz, F. J., Vilanova, S., Plazas, M., Gramazio, P., Andújar, I., Rodríguez-Burruezo, A., et al. (2015). Phenological growth stages of pepino (Solanum muricatum) according to the BBCH scale. *Scientia Hortic.* 183, 1–7. doi: 10.1016/j.scienta.2014.12.008

Herraiz, F. J., Villano, D., Plazas, M., Vilanova, S., Ferreres, F., Prohens, J., et al. (2016). Phenolic Profile and Biological Activities of the Pepino (Solanum muricatum) Fruit and Its Wild Relative S. caripense. *Int. J. Mol. Sci.* 17, 394. doi: 10.3390/ijms17030394

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7, 1071. doi: 10.2307/3870058

Horsch, R. B., Fry, J. E., Hoffmann, N. L., Wallroth, M., Eichholtz, D., Rogers, S. G., et al. (1985). A simple and general method for transferring genes into plants. *Science* 227, 1229–1231. doi: 10.1126/science.227.4691.1229

Hsu, C. C., Guo, Y. R., Wang, Z. H., and Yin, M. C. (2011). Protective effects of an aqueous extract from pepino (Solanum muricatum Ait.) in diabetic mice. *J. Sci. Food Agric.* 91, 1517–1522. doi: 10.1002/jsfa.4345

Hsu, J. Y., Lin, H. H., Hsu, C. C., Chen, B. C., and Chen, J. H. (2018). Aqueous extract of pepino (Solanum muriactum Ait) leaves ameliorate lipid accumulation and oxidative stress in alcoholic fatty liver disease. *Nutrients* 10, 931–946. doi: 10.3390/nu10070931

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

Jung, C. S., Griffiths, H. M., De Jong, D. M., Cheng, S., Bodis, M., and De Jong, W. S. (2005). potato P locus codes for flavonoid 3',5'-hydroxylase. [Erratum: 2005 June, v. 111, no. 1, p. 184.]. *Theor. Appl. Genet.* 110, 269–275. doi: 10.1007/s00122-004-1829-z

Jung, C. S., Griffiths, H. M., De Jong, D. M., Cheng, S., Bodis, M., Kim, T. S., et al. (2009). Potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. *Theor. Appl. Genet.* 120, 45–57. doi: 10.1007/s00122-009-1158-3

Kähkönen, M. P., and Heinonen, M. (2003). Antioxidant activity of anthocyanins and their aglycons. *J. Agric. Food Chem.* 51, 628–633. doi: 10.1021/jf025551i

Kiferle, C., Fantini, E., Bassolino, L., Povero, G., Spelt, C., Buti, S., et al. (2015). Tomato R2R3-MYB proteins SlANT1 and SlAN2: Same protein activity, different roles. *PloS One* 10, e0136365. doi: 10.1371/journal.pone.0136365

Kobayashi, S., Goto-Yamamoto, N., and Hirochika, H. (2005). Association of VvmybA1 gene expression with anthocyanin production in grape (Vitis vinifera) skin-color mutants. *J. Japanese Soc. Hortic. Sci.* 74, 196–203. doi: 10.2503/jjshs.74.196

Kroon, A. R. (2004). *Transcription regulation of the anthocyanin pathway in Petunia hybrida. Research and graduation internal, Vrije Universiteit Amsterdam* (Amsterdam, The Netherlands: Vrije Universiteit Amsterdam).

LaFountain, A. M., and Yuan, Y. W. (2021). Repressors of anthocyanin biosynthesis. *New Phytol.* 231, 933–949. doi: 10.1111/nph.17397

Lin-Wang, K. U. I., Micheletti, D., Palmer, J., Volz, R., Lozano, L., Espley, R., et al. (2011). High temperature reduces apple fruit colour via modulation of the anthocyanin regulatory complex. *Plant Cell Environ.* 34, 1176–1190. doi: 10.1111/j.1365-3040.2011.02316.x

Liu, Y., Lin-Wang, K., Espley, R. V., Wang, L., Yang, H., Yu, B., et al. (2016). Functional diversification of the potato R2R3 MYB anthocyanin activators AN1, MYBA1, and MYB113 and their interaction with basic helix-loop-helix cofactors. *J. Exp. Bot.* 67, 2159–2176. doi: 10.1093/jxb/erw014

Liu, Y., Tikunov, Y., Schouten, R. E., Marcelis, L. F. M., Visser, R. G. F., and Bovy, A. (2018). Anthocyanin biosynthesis and degradation mechanisms in solanaceous vegetables: A review. *Front. Chem.* 6. doi: 10.3389/fchem.2018.00052

Liu, X.-F., Yin, X.-R., Allan, A. C., Lin-Wang, K., Shi, Y.-N., Huang, Y.-J., et al. (2013). The role of MrbHLH1 and MrMYB1 in regulating anthocyanin biosynthetic genes in tobacco and Chinese bayberry (Myrica rubra) during anthocyanin biosynthesis. *Plant Cell Tissue Organ Culture (PCTOC)* 115, 285–298. doi: 10.1007/s11240-013-0361-8

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

Mathews, H., Clendennen, S. K., Caldwell, C. G., Liu, X. L., Connors, K., Matheis, N., et al. (2003). Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport. *Plant Cell* 15, 1689–1703. doi: 10.1105/tpc.012963

Matsui, K., Umemura, Y., and Ohme-Takagi, M. (2008). AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. *Plant J.* 55, 954–967. doi: 10.1111/j.1365-313X.2008.03565.x

Miguel, M. G. (2011). Anthocyanins: Antioxidant and/or anti-inflammatory activities. *J. Appl. Pharm. Sci.* 1, 07–15.

Montefiori, M., Brendolise, C., Dare, A. P., Lin-Wang, K., Davies, K. M., Hellens, R. P., et al. (2015). In the Solanaceae, a hierarchy of bHLHs confer distinct target specificity to the anthocyanin regulatory complex. *J. Exp. Bot.* 66, 1427–1436. doi: 10.1093/jxb/eru494

Payyavula, R. S., Navarre, D. A., Kuhl, J., and Pantoja, A. (2013). Developmental effects on phenolic, flavonol, anthocyanin, and carotenoid metabolites and gene expression in potatoes. *J. Agric. Food Chem.* 61, 7357–7365. doi: 10.1021/jf401522k

Paz-Ares, J., Ghosal, D., Wienand, U., Petersont, P. A., and Saedler, H. (1987). The regulatory cl 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/embj.1987.6.issue-12

Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 181, 219–229. doi: 10.1016/j.plantsci.2011.05.009

Prohens, J., Anderson, G. J., Rodriguez-Burruezo, A., and Nuez, F. (2003). Exploiting wild species for the genetic improvement of the pepino (Solanum muricatum). *J. Appl. Bot-Angew Bot.* 77, 21–27. doi: 10.1007/s10681-005-3882-3

Prohens, J., Rodriguez-Burruezo, A., and Nuez, F. (2005). Utilization of genetic resources for the introduction and adaptation of exotic vegetable crops: The case of pepino (Solanum muricatum). *Euphytica* 146, 133–142. doi: 10.1007/s10681-005-3882-3

Quattrocchio, F., Wing, J., Woude, K. V. D., Souer, E., Vetten, N. D., Mol, J., et al. (1999). Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *Plant Cell* 11, 1433–1444. doi: 10.1105/tpc.11.8.1433

Rodríguez-Burruezo, A. P. T., Prohens, J., and Fita, A. (2011). Breeding strategies for improving the performance and fruit quality of the pepino (Solanum muricatum): A model for the enhancement of underutilized exotic fruits. *Food Res. Int.* 44, 1927–1935. doi: 10.1016/j.foodres.2010.12.028

Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., Van Den Hoff, M. J. B., et al. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* 37, e45–e45. doi: 10.1093/nar/gkp045

Saha, S., Singh, J., Paul, A., Sarkar, R., Khan, Z., and Banerjee, K. (2020). Anthocyanin profiling using UV-vis spectroscopy and liquid chromatography mass spectrometry. *J. AOAC Int.* 103, 23–39. doi: 10.5740/jaoacint.19-0201

Sarkinen, T., Gonzales, P., and Knapp, S. (2013). Distribution models and species discovery: the story of a new Solanum species from the Peruvian Andes. *PhytoKeys*. 31, 1–20. doi: 10.3897/phytokeys.31.6312

Sarma, A. D., and Sharma, R. (1999). Anthocyanin-DNA copigmentation complex: mutual protection against oxidative damage. *Phytochemistry* 52, 1313–1318. doi: 10.1016/S0031-9422(99)00427-6

Schwinn, K., Venail, J., Shang, Y., Mackay, S., Alm, V., Butelli, E., et al. (2006). A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus antirrhinum. *Plant Cell* 18, 831–851. doi: 10.1105/tpc.105.039255

Shathish, K., and Guruvayoorappan, C. (2014). Solanum muricatum Ait. inhibits inflammation and cancer by modulating the immune system. *J. Cancer Res. Ther.* 10, 623–630. doi: 10.4103/0973-1482.138198

Silva Souza, M. A., Peres, L. E. P., Freschi, J. R., Purgatto, E., Lajolo, F. M., and Hassimotto, N. M. A. (2020). Changes in flavonoid and carotenoid profiles alter volatile organic compounds in purple and orange cherry tomatoes obtained by allele introgression. *J. Sci. Food Agric.* 100, 1662–1670. doi: 10.1002/jsfa.10180

Song, X., Liu, H., Shen, S., Huang, Z., Yu, T., Liu, Z., et al. (2022). Chromosome-level pepino genome provides insights into genome evolution and anthocyanin biosynthesis in Solanaceae. *Plant J. Cell Mol. Biol.* 110, 1128–1143. doi: 10.1111/tpj.15728

Spooner, D. M., Anderson, G. J., and Jansen, R. K. (1993). Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (Solanaceae). *Am. J. Bot.* 80, 676–688. doi: 10.1002/j.1537-2197.1993.tb15238.x

Sun, C., Deng, L., Du, M., Zhao, J., Chen, Q., Huang, T., et al. (2020). A transcriptional network promotes anthocyanin biosynthesis in tomato flesh. *Mol. Plant* 13, 42–58. doi: 10.1016/j.molp.2019.10.010

Takos, A. M., Jaffé, F. W., Jacob, S. R., Bogs, J., Robinson, S. P., and Walker, A. R. (2006). Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiol. (Bethesda)* 142, 1216–1232. doi: 10.1104/pp.106.088104

Tamagnone, L., Merida, A., Parr, A., Mackay, S., Culianez-Macia, F. A., Roberts, K., et al. (1998). The AmMYB308 and AmMYB330 transcription factors from Antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10, 135–154. doi: 10.1105/tpc.10.2.135

Walker, A. R., Lee, E., Bogs, J., Mcdavid, D. A. J., Thomas, M. R., and Robinson, S. P. (2007). White grapes arose through the mutation of two similar and adjacent regulatory genes. *Plant journal: Cell Mol. Biol.* 49, 772–785. doi: 10.1111/j.1365-313X.2006.02997.x

Wang, H., Sun, S., Zhou, Z., Qiu, Z., and Cui, X. (2020). Rapid analysis of anthocyanin and its structural modifications in fresh tomato fruit. *Food Chem.* 333, 127439. doi: 10.1016/j.foodchem.2020.127439

Wu, S. B., Meyer, R. S., Whitaker, B. D., Litt, A., and Kennelly, E. J. (2013). A new liquid chromatography-mass spectrometry-based strategy to integrate chemistry, morphology, and evolution of eggplant (Solanum) species. *J. Chromatogr A* 1314, 154–172. doi: 10.1016/j.chroma.2013.09.017

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, Y., Wang, G., Cao, F., Zhu, C., Wang, G., and El-Kassaby, Y. A. (2014). Light intensity affects the growth and flavonol biosynthesis of Ginkgo (Ginkgo biloba L.). *New Forests* 45, 765–776. doi: 10.1007/s11056-014-9435-7

Yan, S., Chen, N., Huang, Z., Li, D., Zhi, J., Yu, B., et al. (2020). Anthocyanin Fruit encodes an R2R3-MYB transcription factor, SlAN2-like, activating the transcription of SlMYBATV to fine-tune anthocyanin content in tomato fruit. *New Phytol.* 225, 2048–2063. doi: 10.1111/nph.16272

Yang, G., Li, L., Wei, M., Li, J., and Yang, F. (2022). SmMYB113 is a key transcription factor responsible for compositional variation of anthocyanin and color diversity among eggplant peels. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.843996

Zhang, Y., Cheng, S., De Jong, D., Griffiths, H., Halitschke, R., and De Jong, W. (2009). potato R locus codes for dihydroflavonol 4-reductase. *Theor. Appl. Genet.* 119, 931–937. doi: 10.1007/s00122-009-1100-8

Zhou, X., Liu, S., Yang, Y., Liu, J., and Zhuang, Y. (2022). Integrated metabolome and transcriptome analysis reveals a regulatory network of fruit peel pigmentation in eggplant (Solanum melongena L.). *Int. J. Mol. Sci.* 23, 13475. doi: 10.3390/ijms232113475

Zhu, H. F., Fitzsimmons, K., Khandelwal, A., and Kranz, R. G. (2009). CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in Arabidopsis. *Mol. Plant* 2, 790–802. doi: 10.1093/mp/ssp030

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 02 July 2024
DOI 10.3389/fpls.2024.1397874

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Sarahani Harun,
National University of Malaysia, Malaysia
Ning Tang,
Chongqing University of Arts and Sciences, China

*CORRESPONDENCE
Li Wang
liwang@yzu.edu.cn

RECEIVED 08 March 2024
ACCEPTED 14 June 2024
PUBLISHED 02 July 2024

CITATION
Wang Y, Liu X, Chen S, Wang Q, Jin B and Wang L (2024) Functions, accumulation, and biosynthesis of important secondary metabolites in the fig tree (*Ficus carica*). *Front. Plant Sci.* 15:1397874. doi: 10.3389/fpls.2024.1397874

Functions, accumulation, and biosynthesis of important secondary metabolites in the fig tree (*Ficus carica*)

Yawen Wang, Ximeng Liu, Siyu Chen, Qingjie Wang, Biao Jin and Li Wang*

College of Horticulture and Landscape, Yangzhou University, Yangzhou, China

Ficus carica is an economically important horticultural plant. Due to its abundant secondary metabolites, *F. carica* has gained interest for its applications in medicine and as a nutritional supplement. Both external and internal factors affect the accumulation of secondary metabolites in *F. carica*. The assembly of the *F. carica* genome has facilitated functional analysis of key genes and transcription factors associated with the biosynthesis of secondary metabolites, particularly anthocyanin. In this review, we summarize the various types and functions of secondary metabolites, with a particular focus on flavonoids, coumarins, and terpenes. We also explore the factors influencing their biosynthesis and accumulation, including varieties, tissue, environmental factors (e.g., light), stresses (e.g., high temperature, low temperature, drought, nutrient deficiencies, salinity), hormonal treatments, and developmental factors. Furthermore, we discuss the involvement of structural genes and transcription factors in the biosynthesis of secondary metabolites, specifically anthocyanin and furanocoumarins, knowledge of which will promote the breeding and genetic engineering of novel *F. carica* varieties.

KEYWORDS

***Ficus carica*, secondary metabolite, anthocyanin, coumarin, biosynthesis**

1 Introduction

Ficus carica, commonly known as the fig tree, is one of the oldest cultivated plants and is among the most important commercial fruit trees in the genus *Ficus*. *F. carica* is extensively cultivated in numerous regions worldwide, particularly in nations of the Mediterranean Basin such as Portugal and Turkey (Dueñas et al., 2008). Fresh and dried

F. carica fruits are nutritious sources of secondary metabolites, making them popular ingredients in culinary preparations such as sauces, fruit wines, and dried fruit assortments. Extracts of *F. carica* leaves and fruits, as well as latex, have been used to treat gastrointestinal, respiratory, inflammatory, metabolic, and cardiovascular conditions (Mawa et al., 2013). The accumulation of secondary metabolites in *F. carica* is influenced by internal and external factors, such as light quality, high temperature, drought, salt stress, and stress-associated hormones. Development, storage, and processing techniques also affect the accumulation of secondary metabolites in *F. carica* fruit (Sandhu et al., 2023). The availability of the genome of *F. carica* and advancements in molecular biology and sequencing technologies have yielded insights into the biosynthetic mechanisms of important secondary metabolites in *F. carica* (Usai et al., 2020). A comprehensive understanding of secondary metabolites in *F. carica* is essential for its horticultural cultivation, commercial significance, and potential health benefits. Here, we review the types and functions of secondary metabolites in *F. carica*, as well as the factors influencing their biosynthesis and accumulation. We also summarize the regulatory mechanisms of important secondary metabolites, including anthocyanins and furanocoumarins, in *F. carica*, focusing on the related structural genes and transcription factors.

2 Types and functions of secondary metabolites in *F. carica*

More than 100 bioactive compounds have been identified in *F. carica*. Almost all parts of the *F. carica* tree are abundant in secondary metabolites, particularly flavonoids, terpenes, and coumarins (Hajam and Saleem, 2022) (Figure 1). In this section, we focus on these secondary metabolites and their effects on health.

2.1 Flavonoids

Flavonoids, a class of polyphenolic secondary metabolites, are formed through the combination of two six-carbon phenyl rings connected by an oxygen heterocycle-containing ring. *F. carica* produces flavonoids in various organs, and the latex secreted from the vascular bundles of the petioles and peduncles also contains flavonoids and phenolic acids (Abdel-Aty et al., 2019). Flavonoids encompass several types, including anthocyanins, flavonols, and isoflavones, with significant research focusing on anthocyanins in *F. carica* fruit (Mao et al., 2023). The content and composition of these anthocyanins play a crucial role in determining the fruit's color (Castañeda-Ovando et al., 2009). Anthocyanin pigments such as cyanidin, delphinidin, peonidin, petudinin, and malvidin are found in red, purple, and other dark-colored fruits.

Cyanidin is the predominant anthocyanin found in *F. carica*. During the ripening process of purple peel, cyanidin O-malonylhexoside levels increase dramatically. Other cyanidins, including cyanidin 3-O-glucoside and cyanidin-3, 5-O-diglucoside, also experience a significant increase (Wang et al., 2019). Among five different varieties of *F. carica* with distinct peel colors, cyanidin 3-rutinoside was the predominant anthocyanin, accounting for 48–81% of peel content and 68–79% of pulp content, while cyanidin 3-glucoside constituted 5–18% of peel content and

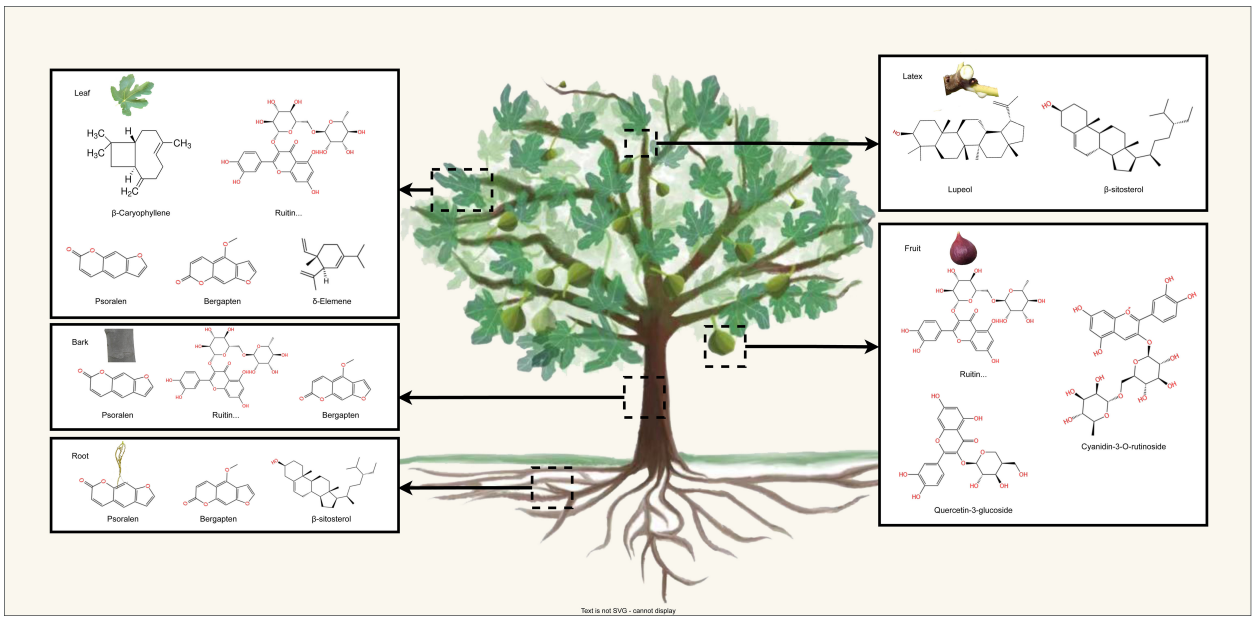


FIGURE 1
The schematic representation illustrates the distribution of main secondary metabolite in the leaves, fruits barks, roots, and latex of *F. carica*.

10–15% of pulp content (Dueñas et al., 2008). Notably, the dark-purple peel of the 'Soltani' variety contains nine identified anthocyanins, primarily cyanidin 3-rutinoside and cyanidin 3,5-diglucoside. The peel of 'Soltani' also revealed the presence of a novel rutinoside and petunidin dimer, yet its function remains unclear (Ammar et al., 2015).

Quercetin, an abundant flavonol in nature, has been extensively studied for its antioxidative and anti-inflammatory properties. *In vitro*, animal, and clinical studies have shown that quercetin can effectively reduce oxidative stress and exhibit anti-inflammatory activity (Ginwala et al., 2019; Jucá et al., 2020). Notably, quercetin has been found to reduce blood glucose levels in diabetic mice (Deepa et al., 2018; Dureshahwar et al., 2019). It is hypothesized that quercetin, along with other compounds, contributes to the observed hypoglycemic effect of the extract. Rutin, a glycoside form of quercetin, is predominantly present in *F. carica* fruit and leaves. Rutin has significant antioxidant activity, anticancer activity, and anti-inflammatory and antibacterial effects (Yang et al., 2008; Hao et al., 2016). It is a popular dietary supplement without any significant side effects. According to the available reports, the highest extraction rate of rutin from *F. carica* pulp was 0.44% (Ammar et al., 2015). In addition to flavonols like quercetin and rutin, *F. carica* also contains isoflavones. Isoflavones, which are 3-phenyl derivatives of benzone rings, are synthesized during the phenylalanine metabolism in plants. *F. carica* fruit has been found to contain 16 types of isoflavones, while the leaf contains 2 types. Interestingly, the isoflavones present in *F. carica* fruit have been shown to possess anti-inflammatory effects by inhibiting the production of nitric oxide (NO) (Liu et al., 2019). Several specific isoflavones, including ficucaricone A, ficucaricone B, and indicanine A, have been reported to exhibit potent anti-inflammatory activities (Magalhães et al., 2006).

2.2 Coumarins

Coumarins account for 51.46% of the volatile oils in *F. carica* leaves (Shahrajabian et al., 2021). Psoralen, bergapten, angelicin, rutaretin, pimpinellin, and seselin are all members of the coumarin family, serving as important secondary metabolites in *F. carica* (Marrelli et al., 2012). Among them, psoralen and bergapten are two of the most prominent coumarins due to their medicinal importance. They have been identified as anticancer agents with the ability to suppress the survival and migration of triple-negative breast cancer MDA-MB-231 cells (Zhang et al., 2018). Psoralens have shown promise in clinical studies, upregulating the level of ERα protein in osteoclasts and inhibiting the growth of breast cancer cells (Zhang et al., 2017). Additionally, psoralens and bergaptens can inhibit acetylcholinesterase (AChE) activity and play a role in the treatment of Alzheimer's disease (Parhiz et al., 2015). Due to their ability to interact with DNA when exposed to ultraviolet A radiation (UV-A), these compounds, especially psoralens, are often used in the treatment of skin diseases such as vitiligo, alopecia areata, psoriasis, and neodermal diseases (Son et al., 2017). Psoralens and bergaptens have also demonstrated significant insecticidal activity properties (Chung et al., 2011; Siyadatpanah et al., 2022).

2.3 Terpenes

F. carica is rich in terpenes. Carotenoids, which belong to the group of tetraterpenoids, are a large group of phytochemicals with powerful antioxidant activities (Kaulmann and Bohn, 2014). *F. carica* contains the carotenoids typically present in human plasma, including β-carotene, lutein, and lycopene (Su et al., 2002). Similar to anthocyanins, carotenoids contribute to the coloration of *F. carica* fruits, giving them an orange or yellow color. The carotenoid contents of fresh and dried fruits differ. B-carotene, zeaxanthin, and lutein are predominant in fresh *F. carica* and β-cryptoxanthin and lutein in dried *F. carica* (Arvaniti et al., 2019). These carotenoids can prevent cardiovascular disease by reducing blood pressure, preventing oxidative stress, and decreasing the secretion of pro-inflammatory factors (Mordente et al., 2011).

Triterpenoids and sesquiterpenes are abundant in the extracts of *F. carica* root, leaf, and latex. Dried and young *F. carica* fruits contain the triterpenoid 9,19-cyclopropyl-24,25-ethylene-5ene-3β-spirosterol, which has anticancer activity (Yin et al., 1997). The main sesquiterpenes in *F. carica* leaf are genanthene, β-caryophyllene, and τ-elemiene (Oliveira et al., 2010c). β-Caryophyllene is used in cosmetics for its spicy aroma (Reinsvold, 2010). Sesquiterpenes are also the most abundant terpenes in *F. carica* latex, accounting for 91% of all identified compounds.

Fresh latex from *F. carica* has been used to treat a variety of types of cancers (Oliveira et al., 2010b; Arvaniti et al., 2019).

3 Accumulation and distribution of secondary metabolites in *F. carica*

3.1 Varieties (cultivars)

The accumulation and distribution of secondary metabolites vary markedly among *F. carica* cultivars (Table 1). In general, dark *F. carica* varieties exhibit higher polyphenol content and total anthocyanin content. Compared with those with light-colored peels, such as green and yellowish-green peels, cultivars with black and purple peels have higher levels of phenolics, flavonoids, and anthocyanins (Çalişkan and Polat, 2011; Ercişli et al., 2012; Hssaini et al., 2020). 'Bursa siyahi' (purple) *F. carica* shows a much higher total proanthocyanidin content than 'Sarilop' (yellow) *F. carica* (yellow) (Kamiloglu and Çapanoğlu, 2015). However, some light-colored *F. carica* varieties exhibit slightly higher proanthocyanidin content (Hssaini et al., 2020). Additionally, varieties of different colors reportedly share some similar anthocyanin profiles, principally cyanidin-3-O-rutinoside (> 90% of total anthocyanins). Moreover, varieties with black peels have high levels of total phenolic compounds, and those with green, greenish-yellow, and brown

TABLE 1 Advantageous secondary metabolites in the main cultivated *F. carica* varieties in fruit peel.

Country	Varieties	References	Advantageous Secondary Metabolites
Israel	Mission (Black)	(Solomon et al., 2006)	Total anthocyanin content
Italy	Mattalona (Black) Dottato	(Del Caro and Piga, 2008) (Russo et al., 2014)	Flavonols (rutin, anthocyanins, hydroxycinnamic acids) Total phenolic content Cyaniding-3-rutinoside
Turkey	Bursa siyahi (purple) Siyah 5	(Kamiloglu and Çapanoğlu, 2015) (Çalişkan and Polat, 2011)	Total flavonoid content Total proanthocyanidin content Total anthocyanin content Flavan-3-ols Total phenolic content Total anthocyanin content
Albania	Kraps Zi (dark)	(Hoxha and Kongoli, 2016)	Total phenolic content Total flavonoid content Total anthocyanin content
Tunisia	Soltani (purple) Kohli	(Ammar et al., 2015) (Harzallah et al., 2016)	Total phenolic content Total anthocynin content Prenylhydroxygenistein Total phenolic content Total tannis content
Iran	Sabz (green) Siyah (dark purple)	(Maghsoudlou et al., 2017)	Total phenolic content Total flavonoid content
Spain	Banane VB1 Cuello de Dama	(Pereira et al., 2017) (Vallejo et al., 2012) (Dueñas et al., 2008)	Total Vitamin C Total phenolic content Total anthocynins Luteolin 6C-hexose-8C-pentose Total anthocynins Cyanidin-3-rutinoside
Algeria	Bakkor Khal Onk Elhamam	(Mahmoudi et al., 2018)	Total phenolic content Total anthocynins Total flavonoid content
Morocco	Fassi Noukali	(Hssaini et al., 2020, 2021)	Total phenolic content Total anthocynins

fruit peels have high levels of chlorogenic acid, vitamin C, and total tannins (Solomon et al., 2006; Pereira et al., 2017).

Among the diverse array of *F. carica* varieties, certain varieties have shown exceptional potential. Notably, in the 'Dottato' and 'Cuello de Dama' varieties from Italy, cyanidin-3-rutinoside, a type of anthocyanin, was identified within the fruit's pulp (Russo et al., 2014). Additionally, the 'Boughandjo', a high-yielding variety, and the 'Bither', an early-ripening variety from Algeria, although not the richest in secondary metabolites, have gained consumer favor for their taste (Mahmoudi et al., 2018). Moreover, cluster analysis of 37 F. *carica* varieties has identified 'Grise de Saint Jean' and 'Grise de Tarascon' as having minimal levels of furanocoumarins (which cause skin inflammation), making them promising candidates for the development of functional foods and pharmaceutical products (Takahashi et al., 2017).

3.2 Tissues

Different tissues of *F. carica* have different secondary metabolite profiles (Table 2). Previous studies have focused more on the secondary metabolites found in the fruits, pulp, peel, leaves, and latex of *F. carica*, with only a limited number of studies exploring the roots and bark.

F. carica fruits serve as a rich reservoir of flavonoids, which notably include quercetin 3-O-rutinoside, quercetin-3-glucosides, polymeric procyanidins, and cyanidin-3-O-rutinoside among their key constituents (Mopuri and Islam, 2016). Quercetin-3-O-rutinoside (rutin) is the predominant flavonoid found in light-colored *F. carica* peel, whereas cyanidin-3-O-rutinoside is the primary flavonoid present in dark-colored *F. carica* peel (Palmeira et al., 2019; Calani et al., 2022). Within fruits, the peel is richer in secondary metabolites than the pulp. For example, the anthocyanin content of the peel of the dark-colored 'Mission' variety was 100-fold that of the pulp (Solomon et al., 2006). Portuguese *F. carica* variety had higher levels of nutritional and phenolic components 158 (Palmeira et al., 2019). Interestingly, an exception is observed in the green *F. carica* varieties 'Bidhi' and 'Kadota', where the concentration of total phenolics in the edible pulp significantly surpasses that found in the peel (Mahmoudi et al., 2018). However, in the *F. carica* processing industry, peels are frequently discarded as waste, highlighting the importance of studying the bioactivity of *F. carica* peels and the need to develop and utilize by-products.

Notably, the fruit pulp is a significantly richer source of prenylated flavonoids compared to the fruit peel (Ammar et al., 2015; Liu et al., 2019). A study analyzed the phenolic composition of leaves, fruit, peel, and pulp, identifying an isomer of prenylhydroxygenistein as the primary compound in the pulps (Ammar et al., 2015). This compound, a type of prenylated flavonoids, is known for its anti-inflammatory and cancer-preventive activities, making it a valuable constituent for the development of functional foods (Chang et al., 2021).

The bioactive compounds present in *F. carica* leaves were extensively summarized in a previous review (Li et al., 2021). Leaves are found to be more abundant in secondary metabolites compared to their fruit counterparts. Overall, *F. carica* leaves are a good source of flavonoids, coumarins, and terpenes. Rutin is the main flavonoid in *F. carica* leaves, and psoralen is the main furanocoumarin compound (Ammar et al., 2015). An experiment measuring coumarin content in leaves, bark, and the woody part of *F. carica* showed that the leaves contain higher levels of psoralen and bergapten (Conforti et al., 2012). Psoralen and bergapten are also major coumarin compounds in the root heartwood (Jaina et al., 2013). Recent research on the phenolic composition of *F. carica* bark identified rutin as the predominant compound (Yahiaoui et al., 2024). In addition, compared with pulp and peel, *F. carica* leaves contain a notably higher content of sesquiterpenes, with β-Caryophyllene and δ-Elemene as the main components, but monoterpenes are rare (Oliveira et al., 2010c).

Latex is rich in coumarins and has a large amount of sterols. A hexane extract of Tunisian common 'Jrani caprifig' latex identified

10.3389/fpls.2024.1397874

TABLE 2 Dominate compounds in different tissues of *F. carica*.

Tissue	Classification	Constituents	Dominant Compounds
Fruit	Flavonoids	Flavonols	Quercetin 3-O-rutinoside; Quercetin-3-glucosides
		Anthocyanins	Polymeric Procyanidins; Cyanidin-3-O-rutinoside
		Isoflavones	Ficucaricone A; Ficucaricone B; Ficucaricone C; Ficucaricone D
Peel	Flavonoids	Flavonols	Quercetin 3-O-rutinoside; Quercetin-3-glucosides
		Anthocyanins	Polymeric Procyanidins; Cyanidin-3-O-rutinoside
Pulp	Flavonoids	Anthocyanins	Cyanidin-3-O-rutinoside
		Isoflavones	Prenylgenistein III; 7methoxy 2'hydroxy genistein (cajanin)
Leave	Flavonoids	Flavonols	Quercetin 3-O-rutinoside
	Coumarins		Psoralen; Bergapten
	Terpenes	Sesquiterpenes	β-Caryophyllene; δ-Elemene
Latex	Sterols		β-sitosterol; Lupeol; 6-O-acyl-beta-D-glucosyl-beta-sitosterols
Woody Part	Coumarins		Psoralen; Bergapten
Bark	Flavonoids	Flavonols	Quercetin 3-O-rutinoside
Root Heartwood	Coumarins		Psoralen; Bergapten
	Sterols		β-sitosterol

36 compounds, 14 of which were coumarins (Lazreg-Aref et al., 2012). A total of 7 phytosterols were identified, among which β-sitosterol and lupeol were the compounds with higher concentrations (Oliveira et al., 2010a). Notably, a mixture of 6-O-acyl-beta-D-glucosyl-beta-sitosterols, which has a significant inhibitory effect on cancer cells, was only isolated from latex (Rubnov et al., 2001).

4 Factors that affect the secondary metabolite biosynthesis and accumulation

4.1 Abiotic stresses

Environmental factors and stressors affect secondary metabolite accumulation in *F. carica*. Light influences the biosynthesis of anthocyanins in plants (Mao et al., 2023). Although light deprivation did not affect the coloration or anthocyanin content of *F. carica* female flower tissues, it significantly inhibited the biosynthesis of anthocyanins in fruit peel (Wang et al., 2019). In the bagged fruit (female flower tissue), 12 flavonoids were detected, of which 11 increased, and apigenin increased by 11 times (Wang et al., 2021). In addition, pulse light treatment using a high-energy xenon flashlight after harvesting resulted in a 20-fold increase in the anthocyanin content of *F. carica* fruit peel. Moreover, pulse light treatment increased the total phenolic content of fruit. Therefore, pulse light markedly promotes the accumulation of anthocyanins and other phenolic compounds in *F. carica* (Rodov et al., 2012).

Increased production of secondary metabolites can ameliorate stress-induced damage in plants (Figure 2). The total phenolic level of *F. carica* is increased by phosphorus and calcium deficiency (Garza-Alonso et al., 2020). Salt stress increased the total phenolic content of *F. carica* by 5.6%, with a particularly pronounced increase observed in the levels of epicatechin (Francini et al., 2021). Additionally, high temperatures increase the phenolic content of the leaves of *F. carica* seedlings (Guo et al., 2017), and drought stress increases the α-tocopherol content (Gholami et al., 2012). In a study of fresh, frozen, and processed fruit, storage of *F. carica* at −18°for 4 months resulted in significant decreases in the total phenolic, total flavonoid, and total anthocyanin contents (Petkova et al., 2019).

4.2 Phytohormone treatments

Phytohormones related to stress resistance, such as abscisic acid (ABA), jasmonic acid, and ethylene, can modulate the biosynthesis of anthocyanins in *F. carica* (Figure 3) (Guo et al., 2017; Lama et al., 2019; Cui et al., 2021). ABA promotes the accumulation of anthocyanins during fruit ripening, thereby initiating color development (Lama et al., 2019). *F. carica* treated with ABA had higher levels of cyanidin 3-O-glucoside and cyanidin 3-O-rucoside and developed a dark-purple color earlier compared with those treated with an ABA inhibitor (Lama et al., 2020). Transcriptomic analysis indicated that ABA upregulated the expression levels of genes related to anthocyanin biosynthesis, whereas an ABA inhibitor reversed this effect, confirming the importance of ABA for the biosynthesis of anthocyanins. Ethylene is a phytohormone associated with plant maturation and is implicated in fruit development and stress responses. Ethylene has bidirectional regulatory effects on female flower and peel coloration in *F. carica*. Treatment of *F. carica* with exogenous ethylene

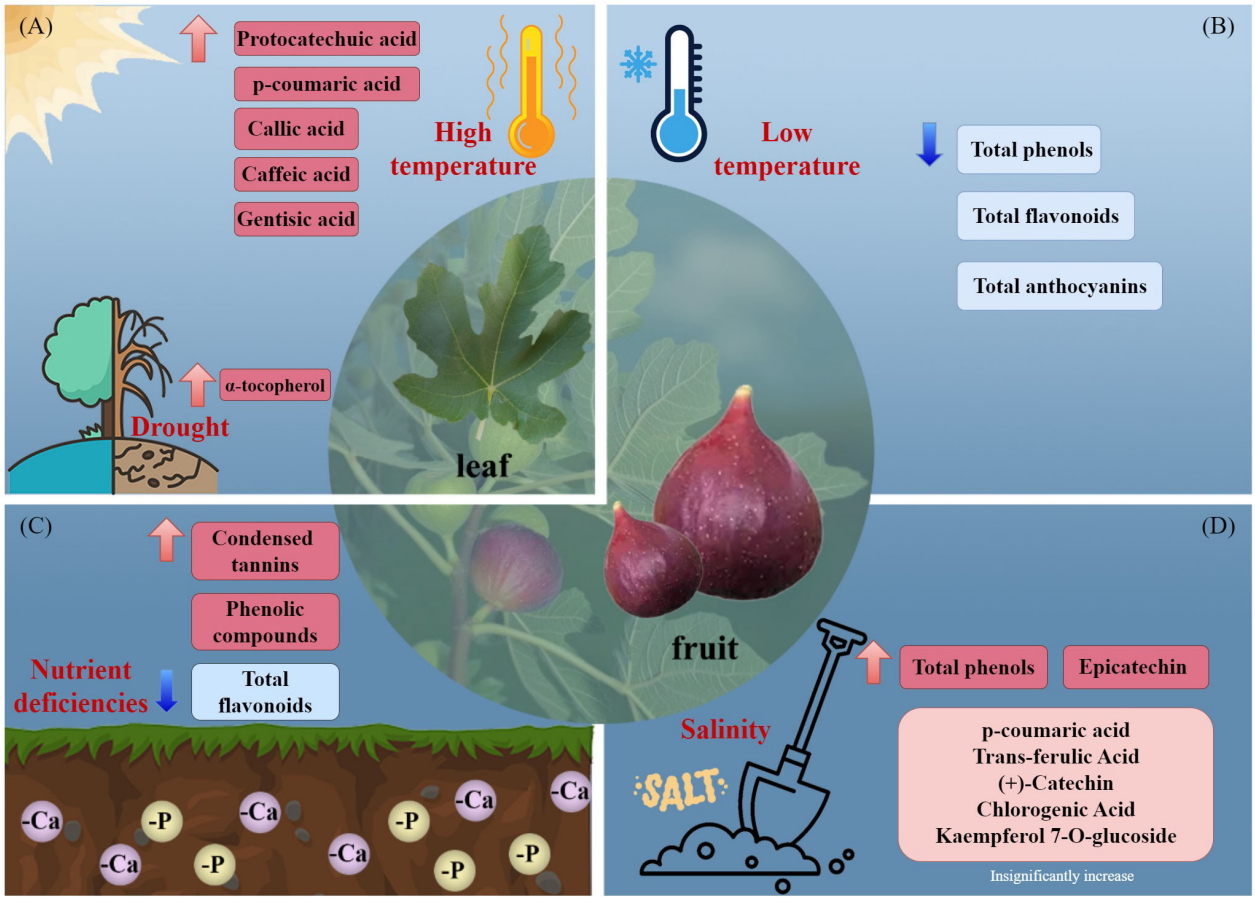


FIGURE 2
Effects of environmental stresses affect the accumulation of secondary metabolites. **(A)** high temperature and drought, **(B)** low temperature, **(C)** nutrient deficiencies, and **(D)** salinity.

upregulated the expression levels of genes related to flavonoid biosynthesis, thereby accelerating the coloration of the receptacle, whereas downregulation of these genes delayed the coloration of female tissues (Cui et al., 2019). In addition, *F. carica* seedlings treated with methyl jasmonate under high-temperature stress had elevated total phenolic contents (Guo et al., 2017).

4.3 Developmental factors

Developmental and ripening stages affect the accumulation of secondary metabolites in *F. carica* fruit (Veberic et al., 2008). Some species of *F. carica* can ripen twice and be harvested twice annually: first from late May until the end of June ('Breba') and subsequently from mid-July to early September (main crop). In the *F. carica* cultivars 'Zuccherina', 'Crna Petrovka', and 'Miljska', fruits from the second ripening had higher levels of phenolic than those from the first ripening (Veberic et al., 2008). Similarly, the main crops of two 'Albanian' cultivars had higher phenolics contents compared with the 'Brebas' from the second ripening (Hoxha and Kongoli, 2016). However, in the cultivars 'VB1', 'Antonio', and 'Santado', the levels of phenolic compounds were higher in fruits from the first ripening than in those from the second ripening, suggesting that secondary metabolite accumulation varies among developmental and ripening stages (Vallejo et al., 2012). In addition, the contents of furanocoumarins and pyranocoumarins are higher in the first ripening in June, whereas those of polyphenolics are highest in the second ripening in September (Marrelli et al., 2012).

The types and levels of secondary metabolites also vary according to the maturity of *F. carica* fruit within a ripening stage. Among fruits at the immature, mature, and fully mature stages, the total phenolic content increased during maturation and peaked at the fully mature stage (Pereira et al., 2017). However, others have reported that the total phenolic content decreases with fruit development (Gündeli et al., 2021; Hoxha et al., 2022). In addition, the polyphenolic contents of dark-colored *F. carica* cultivars change more during ripening compared with light-colored cultivars (Wang et al., 2019). Overall, the accumulation of secondary metabolites in *F. carica* fruit is influenced by the developmental time, stage of maturity, and also depends on the cultivars.

4.4 Post-harvest processing techniques

As a result of its susceptibility to physical damage and post-harvest infection, fresh *F. carica* fruit is typically processed into

10.3389/fpls.2024.1397874

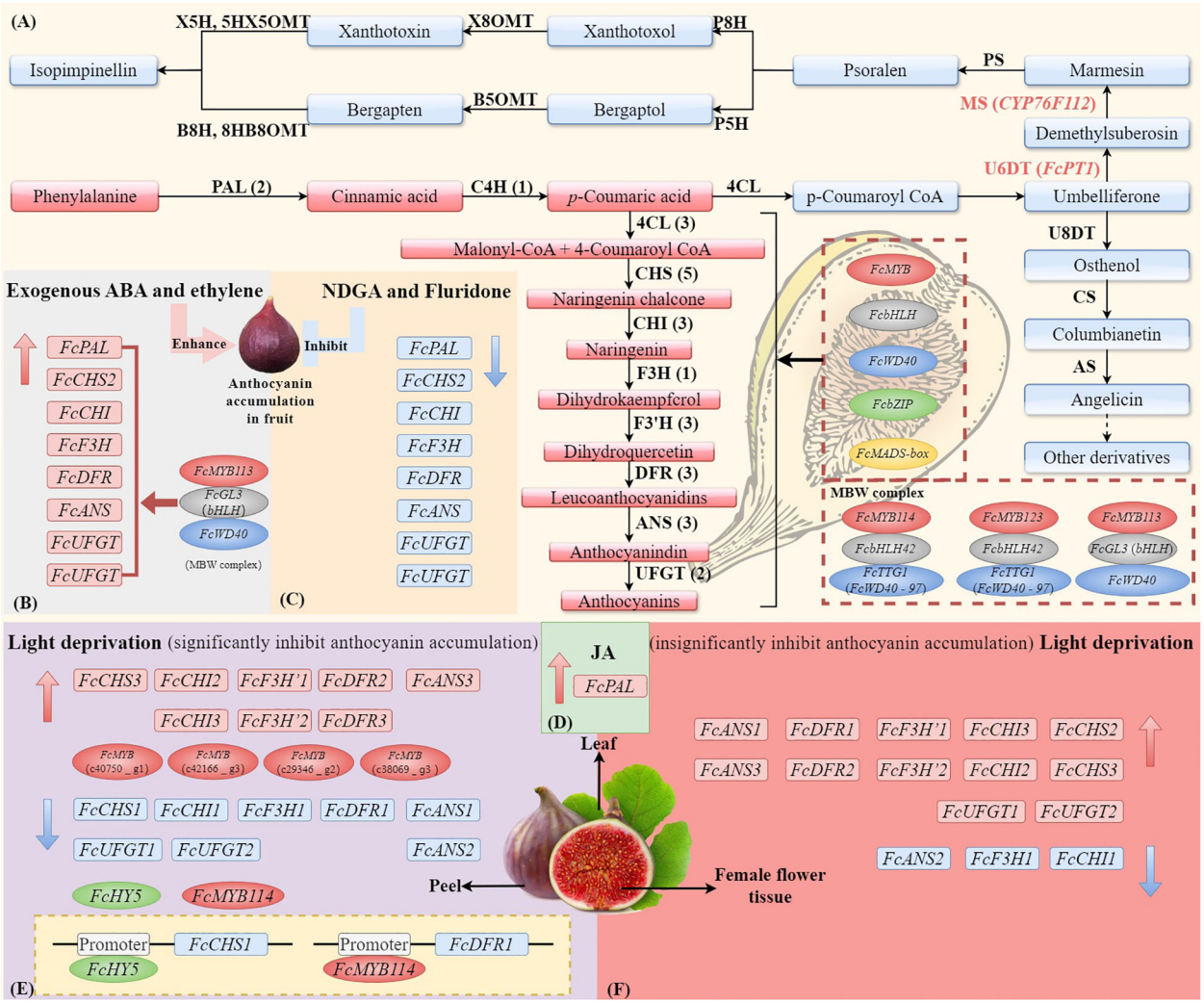


FIGURE 3
The synthesis pathways of anthocyanin and coumarin in *F. carica* and the external factors (ABA, light, and JA) affect the expression of genes and TFs. **(A)** The biosynthesis of anthocyanins in *F. carica* is regulated by MYB, bHLH, WD40, bZIP, and MADS-box transcription factors, as well as the MBW (MYB-bHLH-WD40) complex. Effects of exogenous abscisic acid **(ABA)** or ethylene **(B)**, ABA inhibitors nordihydroguaiaretic acid (NDGA) or fluridone **(C)**, and jasmonic acids (JAs) **(D)** on the expression of anthocyanin biosynthesis genes in fruit. Effects of light deprivation on anthocyanin biosynthesis genes in fig peel **(E)** and female flower tissue **(F)**. The structural genes involved in the biosynthesis of anthocyanins include phenylalanine ammonium lyase (PAL), cinnamate-4-hydroxylase (C4H), 4-coumarate: coenzyme A ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavanone 3'-hydroxylase (F3'H), dihydroflavonol 4-reductase (DFR), anthocyanin synthetase (ANS), and UDP glucose: flavonoid 3-O-glucosyltransferase (UFGT). The simplified biosynthetic pathways of coumarin and furanocoumarin (FC). The core skeleton structures of linear and angular FCs are psoralen and angelicin. The key enzymes involved in the biosynthesis of FCs include umbelliferone 6-dimethylallyltransferase (U6DT), umbelliferone 8-dimethylallyltransferase (U8DT), marmesin synthase (MS), columbianetin synthase (CS), psoralen synthase (PS), angelicin synthase (AS), psoralen 5-hydroxylase (P5H), psoralen 8-hydroxylase (P8H), bergaptol 5-O-methyltransferase (B5OMT), xanthotoxol 8-O-methyltransferase (X8OMT), bergapten 8-hydroxylase (B8H), xanthotoxin 5-hydroxlyase (X5H), 8-hydroxybergapten 8-O-methyltransferase (8HB8OMT), and 5-hydroxyxanthotoxin 5-O-methyltransferase (5HX5OMT).

dried fruits, jams, jellies, and syrups, as examples. These processes affect the secondary metabolite profiles of *F. carica* fruit (Liu et al., 2020). Drying, including sun- and oven-drying, is the most common method for preserving fresh *F. carica* fruit. Sun drying markedly reduces the levels of polyphenols in fruit, and the effect is greater on flavonoids than on phenolic acids. Sun drying decreased the contents of phenolic acids and flavonoids in fruit by 29% and 86%, respectively (Bachir et al., 2016). Similarly, the total phenolic and total anthocyanin contents of a yellow- and a purple-colored fruit variety decreased significantly after drying, but the total flavonoid and proanthocyanidin contents of the yellow fruit variety significantly increased (Kamiloglu and Çapanoğlu, 2015). However, the levels of phenolic compounds, except apigenin-3-O-rutinoside, were high after oven drying, suggesting that oven drying can increase the total phenolic content and the antioxidant activity of fruit (Slatnar et al., 2011). The processing of fruit pulp into jam and honey also reduces the content of secondary metabolites (Tanwar et al., 2014). Although nectar processing, and jam processing reduce the levels of secondary metabolites, considerable quantities, particularly of carotenoids and phenols,

remain after jam processing (Tanwar et al., 2014; Petkova et al., 2019). This suggests that the processing of fruit into jams maintains, at least in part, their contents of phenolics and carotenoids.

5 Enzymes involved in secondary metabolism in *F. carica*

The first preliminary genome sequence of *F. carica* was obtained by sequencing the genome of the Japanese traditional variety 'Horaishi' based on the Illumina platform. The first high-density linkage map was constructed using restriction enzyme site-associated sequencing (Mori et al., 2017). Further sequencing and assembly of the *F. carica* genome, 'Dottato', have been conducted using third-generation sequencing technologies (Usai et al., 2020). This version of the genome assembly of *F. carica* comprises 333 Mbp, 80% of which is anchored to 13 chromosomes (Usai et al., 2020). Later, the newest genome of *F. carica* 'Orphan' was built, with a total length of 366.34 Mb and a contig N50 of 9.78 Mb (Bao et al., 2023). Generally, ten families of structural genes associated with the anthocyanin biosynthesis pathway have been identified: *FcPAL* (two genes), *FcC4H* (one gene), *Fc4CL* (three genes), *FcCHS* (five genes), *FcCHI* (three genes), *FcF3H* (one gene), *FcF3'H* (three genes), *FcDFR* (three genes), *FcANS* (three genes), and *FcUFGT* (two genes) (Li et al., 2020) (Figure 3).

5.1 Chalcone synthase and chalcone isomerase

Chalcone synthase (CHS) is a key enzyme that catalyzes the first committed step in the anthocyanin biosynthesis pathway (Jez and Noel, 2001). As the receptacle tissues of *F. carica* develop and ripen, the contents of flavonoids increase, and the expression level of *FcCHS* is upregulated (Cui et al., 2021). *FcCHS1* encodes 389 amino acid residues, including two phenylalanine residues essential for substrate specificity (Li et al., 2020). In addition, the transcription of *FcCHS1* can be regulated by light-induced signal transduction elements such as HY5 (Wang et al., 2019). Subsequent to the reaction catalyzed by CHS, chalcone isomerase (CHI) catalyzes the conversion of chalcone into naringenin (Ralston et al., 2005). *FcCHI1* encodes a 257 amino-acid type I chalcone isomerase, which includes four conserved active sites (Li et al., 2020). The expression levels of *FcCHS1* and *FcCHI1* increase with the accumulation of anthocyanins and are significantly higher in pigmented tissues (fruit peels and flowers) than non-pigmented tissues. The expression of *FcCHS1* is very low in non-pigmented tissues (Li et al., 2020). *FcCHS1* and *FcCHI1*, together with their homologs associated with anthocyanin accumulation, have similar molecular characteristics and secondary structures, implicating their involvement in *F. carica* fruit coloration.

5.2 Dihydroflavonol-4-reductase

Dihydroflavonol-4-reductase (DFR) is a key enzyme involved in the later steps of anthocyanin biosynthesis. It catalyzes the conversion of dihydroflavonols to colorless anthocyanins. Metabolomic and transcriptomic analyses showed that the expression levels of three *FcDFR* genes were significantly increased in fruit peel. Among them, the expression levels of *FcDFR2* and *FcDFR3* were upregulated by light deprivation, whereas *FcDFR1* was downregulated (Wang et al., 2019). *FcDFR1* encodes a 363-amino-acid protein encompassing a typical NADP-binding region and a substrate-binding region (Li et al., 2020). The secondary structure of *FcDFR1* is similar to *MnDFR1* (*Morus notabilis*), which has been implicated in anthocyanin accumulation, suggesting that *FcDFR1* is involved in the anthocyanin biosynthesis of *F. carica*.

5.3 Anthocyanidin synthase and UDP-glycose flavonoid glycosyltransferase

Anthocyanidin synthase (ANS) catalyzes the penultimate step in the anthocyanin biosynthesis pathway, the conversion of colorless into colored anthocyanins. *FcANS1* encodes a 358-amino-acid protein encompassing a conserved domain 2OG-FeII-Oxy (Cao et al., 2016). *FcANS1* was significantly upregulated in *F. carica* fruit peel compared with flower tissue (Wang et al., 2019). The expression pattern of *FcANS1* and the anthocyanin content are correlated during fruit development. Bagging treatment resulted in downregulation of *FcANS1* expression in fruit peel and female flowers, suggesting that *FcANS1* functions in anthocyanin biosynthesis (Cao et al., 2016). On the contrary, three *FcANS* genes *FcANS1*, *FcANS2* and *FcANS3* in female flower tissues were upregulated after bagging (Wang et al., 2021).

UFGT is the last key enzyme in the anthocyanin synthesis pathway, which can catalyze the glycosylation of unstable anthocyanins to form stable anthocyanins. The open reading frame of *FcUFGT1* is 1374 bp. The expression level of *FcUFGT1* in the fruit peel of *F. carica* at the mature stage was much higher than that in female flower tissue. However, the expression of *FcUFGT1* was significantly downregulated in the fruit peel by light deprivation but significantly upregulated in the female flower tissue (Wang et al., 2019).

5.4 Other enzymes

Glutathione S-transferases (GSTs) are multifunctional enzymes involved in the transport of anthocyanins to vacuoles (Gullner et al., 2018). A total of 53 *GST* genes from five subfamilies are present in the *F. carica* genome. *FcGSTF1* and *FcGSTU5/6/7* may be important in anthocyanin accumulation in *F. carica* peel, but further research is needed (Liu et al., 2023).

Compared to anthocyanins, less research has focused on the regulation of the biosynthesis of other secondary metabolites in *F. carica*. Furanocoumarins (FCs), like anthocyanins, are synthesized via the phenylpropanoid pathway and are structurally classified as angular and linear furanocoumarins (Villard et al., 2019). The first step of the furanocoumarin pathway is catalyzed by the

umbelliferone dimethylallyl transferase (UDT) and is important for the prenylation of umbelliferone (Figure 3). Comparative RNA-seq and enzyme activity analyses of *F. carica* latex organs have shown that *FcPT1* has UDT activity and participates in the biosynthesis of furanocoumarins in *F. carica*, converting umbelliferone into demethylsuberosin (DMS) (Munakata et al., 2020). However, the functions of these enzymes need to be validated by *in vitro* studies. Marmesin synthase is a key enzyme in the biosynthesis of furanocoumarins and is involved in the production of the defense compound psoralen from p-coumaroyl coenzyme A (Villard et al., 2021). The cytochrome P450 gene *CYP76F112* in *F. carica* encodes marmesin synthase, which can convert DMS into marmesin with high affinity. The discovery of *CYP76F112*, the first identified marmesin synthase, expands the potential applications of the furancoumarin pathway (Villard et al., 2021).

6 Transcription factors regulating the biosynthesis of secondary metabolites in *F. carica*

Transcription factors modulate the expression of structural genes associated with the biosynthesis of secondary metabolites. They bind to *cis*-acting regions, thereby inhibiting or promoting gene expression. These factors collaboratively or independently regulate the synthesis of enzymes involved in the biosynthesis of secondary metabolites, such as anthocyanins.

6.1 MYB transcription factors

MYB transcription factors are widely distributed in plants and are involved in a variety of aspects of plant growth and metabolism (Pratyusha and Sarada, 2022). Transcriptomic analysis revealed that light deprivation alters the expression of several *F. carica* genes, such as *MYB114* and three R2R3MYB transcription factors, with *FcMYB114* specifically regulating *FcDFR1* and anthocyanin biosynthesis via promoter interaction (Wang et al., 2019). Virus-induced gene silencing of *FcMYB114* led to significant downregulation of *FcDFR1*, *FcANS1*, and *FcUFGT1*, resulting in a reduction of anthocyanin content in *F. carica* fruit peel (Chen et al., 2023). Overexpression of two R2R3-MYB genes, *FcMYB21* and *FcMYB123*, led to significantly increased anthocyanin content in apple peel and fruit callus tissues (Li et al., 2020). Furthermore, the significant promotion of *MdMYC2* expression by overexpressing *FcMYB21* suggests a potential close relationship between *FcMYB21* and *FcMYC2* in promoting *F. carica* fruit coloring.

6.2 Basic helix-loop-helix transcription factors

The bHLH family is the second largest group of plant transcription factors involved in plant growth, development, and signal transduction. bHLH transcription factors can interact with MYB transcription factors to regulate the biosynthesis of anthocyanins (Hichri et al., 2010). The genome of *F. carica* contains 118 bHLH genes (Song et al., 2021). According to the bHLH family classification in *Arabidopsis*, *FcbHLH* genes can be classified into 25 subgroups. bHLH genes of different subfamilies show different expression patterns in female flower tissues and fruit peel during *F. carica* fruit development. The transient transformation of *FcbHLH42* in tobacco induced a significant increase in anthocyanin content, suggesting that *FcbHLH42* promotes the accumulation of anthocyanins in *F. carica*. The levels of WD40 and other MYB transcription factors were positively correlated with the expression of *FcbHLH42*, providing insight into the regulatory mechanism of flavonoid biosynthesis in *F. carica* (Song et al., 2021).

6.3 WD40

The WD40 transcription factors are structurally stable and important regulators of plant development and physiology, including the biosynthesis of anthocyanins. TTG1 was the first WD40 protein discovered in *Arabidopsis thaliana*, and it forms a trimer with *AtMYB123* and *AtbHLH42* to promote anthocyanin biosynthesis (Baudry et al., 2004). The *F. carica* genome contains a total of 204 WD40 genes. Subcellular localization prediction showed that 109 *FcWD40* proteins were localized to the cytoplasm, 69 to the nucleus, and 26 to other cellular compartments. *FcWD4097* has been identified and named *FcTTG1*. The expression level of *FcTTG1* was significantly highest in the flesh and peel, followed by the stem of *F. carica* (Fan et al., 2022).

MYB transcription factors form the MBW complex by binding to the *FcMYB113*, *FcGL3*, and *FcWD40* proteins of the MYB, bHLH, and WD-repeat families, respectively. This complex directly activates the expression of genes involved in the biosynthesis of anthocyanins (Gonzalez et al., 2008). The MBW complex was identified in pollinated and parthenocarpic *F. carica* fruits. *FcMYB113* was significantly downregulated, whereas *FcGL3* and *FcWD40* were moderately upregulated by exogenous ABA and ethylene. The high expression of *FcMYB113*, *FcGL3*, and *FcWD40* during the later stages of anthocyanin biosynthesis may affect coloration (Lama et al., 2020). In addition, yeast two-hybrid assays and biomolecular fluorescence complementation experiments showed that *FcWD40–97* and *FcbHLH42* regulate the biosynthesis of anthocyanins by forming MBW complexes with *FcMYB114* and *FcMYB123* (Fan et al., 2022).

6.4 Other transcription factors

Other families of transcription factors, such as MADS-box and bZIP, have been identified in *F. carica*. MADS-box proteins are functionally diverse and regulate multiple processes, including floral organ development, fruit development and maturation, and seed coloring (Aerts et al., 2018). The *F. carica* genome has a total of 64 *FcMADS-box* genes, the expression levels of which were shown to be tissue specific (Kmeli et al., 2023). *FcMADS9* has been cloned,

and its regulation of the biosynthesis of anthocyanins in *F. carica* has been validated by heterologous expression in apple fruit and callus tissues (Li et al., 2021).

The bZIP family is one of the largest gene families in plants and is important for biological processes such as secondary metabolism, stress responses, and seed maturation. Y1 screening indicated that *FcHY5* (bZIP) binds to the promoter region of *FcCHS*, implicating the latter in transcription of the structural genes linked to the biosynthesis of anthocyanins (Wang et al., 2019).

7 Concluding remarks and future perspectives

We have conducted a comprehensive review of the functions and biosynthesis of secondary metabolites in the economically significant *F. carica*. A diverse array of secondary metabolites in *F. carica*, such as flavonoids, terpenes, and coumarins, exhibit various beneficial activities. Dark-colored *F. carica* varieties and the fruit peel display elevated levels of secondary metabolites.

To fully exploit the potential application of these metabolites in food and medicine, it is crucial to acquire a more in-depth understanding of their biological activity and pharmacological efficacy. The functions of the flavonoids, terpenoids, and coumarins of *F. carica* require further investigation in cell-culture experiments, animal models, and clinical studies to identify potential health benefits. In addition, while *F. carica* is primarily consumed as fruit, the leaves contain higher levels of coumarin compounds, which are usually discarded as industrial byproducts without utilization; therefore, further research on the use and value of these byproducts is necessary.

The content of secondary metabolites undergoes substantial changes during the development and maturation of *F. carica* fruits, and it is also influenced by the genotype of *F. carica*. Additionally, a range of environmental factors (e.g., light), stresses (e.g., high temperature, low temperature, drought, nutrient deficiencies, salinity), hormone treatments (e.g., ABA, ethylene, NDGA, fluridone, and methyl jasmonate), as well as developmental factors, can affect the accumulation of secondary metabolites in *F. carica*. Post-harvest processing methods are worth further exploration as a low-cost and practical technique to maximize the retention of secondary metabolites in *F. carica*. Some cultivation and management practices can be applied to *F. carica* to achieve better fruit quality, including artificial light supplementation systems, fertilizer rationing, and the application of growth regulators.

Anthocyanins, as secondary metabolites in *F. carica*, play a significant role in determining the fruit's appearance and nutritional value. The availability of the *F. carica* genome has facilitated the investigation of key genes involved in anthocyanin biosynthesis and their regulatory mechanisms. Transcriptomic and bioinformatics analyses have identified several structural genes, including *FcPAL*, *FcCHS*, *FcCHI*, *FcF3H*, *FcDFR*, *FcANS*, *FcUFGT*, and *FcGST*, which regulate anthocyanin synthesis. Furthermore, *CYP76F112* and *FcPT1* have been identified as crucial enzymes in the coumarin compound synthesis pathway. Moreover, MYB, bHLH, WD40, bZIP, and MADS-box transcription factors are implicated in the regulation of the biosynthesis of anthocyanins. Genetic and molecular studies are needed to further identify the functional genes involved in the regulation of secondary metabolite biosynthesis. The application of genetic transformation systems (root hair transformation, virus-induced gene silencing, and gene editing) will facilitate the development of biotechnological techniques for increased contents of bioactive secondary metabolites in *F. carica* varieties. Studies on the biosynthesis mechanisms of anticancer bioactive compounds such as psoralens and saponins in *F. carica* warrant further research.

Author contributions

YW: Writing – original draft, Writing – review & editing, Visualization. XL: Writing – review & editing, Visualization. SC: Writing – review & editing. QW: Writing – review & editing. BJ: Writing – review & editing. LW: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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

Abdel-Aty, A. M., Hamed, M. B., Salama, W. H., Ali, M., Fahmy, A. S., and Mohamed, S. A. (2019). *Ficus carica*, *Ficus* sycomorus and Euphorbia tirucalli latex extracts: Phytochemical screening, antioxidant and cytotoxic properties. *Biocatal. Agric. Biotechnol.* 20, 101199. doi: 10.1016/J.BCAB.2019.101199

Aerts, N., de Bruijn, S., van Mourik, H., Angenent, G. C., and van Dijk, A. D. J. (2018). Comparative analysis of binding patterns of MADS-domain proteins in Arabidopsis thaliana. *BMC Plant Biol.* 18, 131. doi: 10.1186/s12870–018-1348–8

Ammar, S., del Mar Contreras, M., Belguith-Hadrich, O., Belguith-Hadrich, O., Segura-Carretero, A., and Bouaziz, M. (2015). Assessment of the distribution of phenolic compounds and contribution to the antioxidant activity in Tunisian fig leaves, fruits, skins and pulps using mass spectrometry-based analysis. *Food Funct.* 6, 3663–3677. doi: 10.1039/c5fo00837a

Arvaniti, O. S., Samaras, Y., Gatidou, G., Thomaidis, N. S., and Stasinakis, A. S. (2019). Review on fresh and dried figs: Chemical analysis and occurrence of phytochemical compounds, antioxidant capacity and health effects. *Food Res. Int.* 119, 244–267. doi: 10.1016/j.foodres.2019.01.055

Bachir, B. M., Richard, G., Meziant, L., Fauconnier, M. L., and Louaileche, H. (2016). Effects of sun-drying on physicochemical characteristics, phenolic composition and in *vitro* antioxidant activity of dark fig varieties. *J. Food Process. Preserv.* 41, e13164. doi: 10.1111/jfpp.13164

Bao, Y., He, M., Zhang, C., Jiang, S., Zhao, L., Ye, Z., et al. (2023). Advancing understanding of *Ficus carica*: a comprehensive genomic analysis reveals evolutionary patterns and metabolic pathway insights. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1298417

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

Calani, L., Bresciani, L., Rodolfi, M., Del Rio, D., Petruccelli, R., Faraloni, C., et al. (2022). Characterization of the (poly)phenolic fraction of fig peel: comparison among twelve cultivars harvested in Tuscany. *Plants* 11, 3073. doi: 10.3390/plants11223073

Çalişkan, O., and Polat, A. A. (2011). Phytochemical and antioxidant properties of selected fig (*Ficus carica* l.) accessions from the eastern Mediterranean region of Turkey. *Sci. Hortic.* 128, 473–478. doi: 10.1016/j.scienta.2011.02.023

Cao, L., Xu, X., Chen, S., and Ma, H. (2016). Cloning and expression analysis of *Ficus carica* anthocyanidin synthase 1 gene. *Sci. Hortic.* 211, 369–375. doi: 10.1016/j.scienta.2016.09.015

Castañeda-Ovando, A., Pacheco-Hernández, M. D. L., Páez-Hernández, M. E., Rodríguez, J. A., and Galán-Vidal, C. A. (2009). Chemical studies of anthocyanins: a review. *Food Chem.* 113, 859–871. doi: 10.1016/j.foodchem.2008.09.001

Chang, S. K., Jiang, Y., and Yang, B. (2021). An update of prenylated phenolics: Food sources, chemistry and health benefits. *Trends Food Sci. Tech.* 108, 197–213. doi: 10.1016/j.tifs.2020.12.022

Chen, Z. Y., Gao, W. K., Li, M. H., He, D. N., Long, Y. D., Huang, L., et al. (2023). Function of *FcMYB114* gene in fig (*Ficus carica* L.) by virus-induced gene silencing. *Southwest. China J. Agric. Sci.* 36, 623–628. doi: 10.16213/j.cnki.scjas.2023.3.021

Chung, I. M., Kim, S. J., Yeo, M. A., Park, S. W., and Moon, H. I. (2011). Immunotoxicity activity of natural furocoumarins from milky sap of *Ficus carica* L. against *Aedes aEgypti* L. *Immunopharmacol. Immunotoxicol.* 33, 515–518. doi: 10.3109/08923973.2010.543907

Conforti, F., Menichini, G., Zanfini, L., Tundis, R., Statti, G. A., Provenzano, E., et al. (2012). Evaluation of phototoxic potential of aerial components of the fig tree against human melanoma. *Cell Prolif.* 45, 279–285. doi: 10.1111/j.1365–2184.2012.00816.x

Cui, Y., Wang, Z., Chen, S., Vainstein, A., and Ma, H. (2019). Proteome and transcriptome analyses reveal key molecular differences between quality parameters of commercial-ripe and tree-ripe fig (*Ficus carica* L.). *BMC Plant Biol.* 19, 146. doi: 10.1186/s12870–019-1742-x

Cui, Y., Zhai, Y., Flaishman, M., Li, J., Chen, S., Zheng, C., et al. (2021). Ethephon induces coordinated ripening acceleration and divergent coloration responses in fig (*Ficus carica L.*) flowers and receptacles. *Plant Mol. Biol.* 105, 347–364. doi: 10.1007/s11103–020-01092-x

Deepa, P., Sowndhararajan, K., Kim, S., and Park, S. J. (2018). A role of *Ficus* species in the management of diabetes mellitus: a review. *J. Ethnopharmacol.* 215, 210–232. doi: 10.1016/j.jep.2017.12.045

Del Caro, A., and Piga, A. (2008). Polyphenol composition of peel and pulp of two Italian fresh fig fruits cultivars (*Ficus carica* L.). *Eur. Food Res. Technol.* 226, 715–719. doi: 10.1007/s00217–007-0581–4

Dueñas, M., Pérez-Alonso, J. J., Santos-Buelga, C., and Escribano-Bailón, T. (2008). Anthocyanin composition in fig (*Ficus carica* L.). *J. Food. Compost. Anal.* 21, 107–115. doi: 10.1016/j.jfca.2007.09.002

Dureshahwar, K., Mubashir, M., Upaganlwar, A., Sangshetti, J. N., Upasani, C. D., and Une, H. D. (2019). Quantitative assessment of tactile allodynia and protective effects of flavonoids of *Ficus carica* Lam. leaves in diabetic neuropathy. *Pharmacogn. Mag.* 15, 128–134. doi: 10.4103/pm.pm_553_18

Ercişli, S., Tosun, M., Karlidag, H., Dzubur, A., Hadziabulic, S., and Aliman, Y. (2012). Color and antioxidant characteristics of some fresh fig (*Ficus carica* L.) genotypes from northeastern Turkey. *Plant Foods. Hum. Nutr.* 67, 271–276. doi: 10.1007/s11130–012-0292–2

Fan, Z., Zhai, Y., Wang, Y., Zhang, L., Song, M., Flaishman, M. A., et al. (2022). Genome-wide analysis of anthocyanin biosynthesis regulatory WD40 gene *FcTTG1* and related family in *Ficus carica* L. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.948084

Francini, A., Sodini, M., Vicario, G., Raffaelli, A., Gucci, R., Caruso, G., et al. (2021). Cations and phenolic compounds concentrations in fruits of fig plants exposed to moderate levels of salinity. *Antioxid* 10, 1865. doi: 10.3390/antiox10121865

Garza-Alonso, C. A., Niño-Medina, G., Gutiérrez-Díez, A., García-López, J. I., Vázquez-Alvarado, R. E., López-Jiménez, A., et al. (2020). Physicochemical characteristics, minerals, phenolic compounds, and antioxidant capacity in fig tree fruits with macronutrient deficiencies. *Not. Bot. Horti. Agrobo.* 48, 1585–1599. doi: 10.15835/nbha48311867

Gholami, M., Rahemi, M., Kholdebarin, B., and Rastegar, S. (2012). Biochemical responses in leaves of four fig cultivars subjected to water stress and recovery. *Sci. Hortic.* 148, 109–117. doi: 10.1016/j.scienta.2012.09.005

Ginwala, R., Bhavsar, R., Chigbu, D., Jain, P., and Khan, Z. K. (2019). Potential role of flavonoids in treating chronic inflammatory diseases with a special focus on the anti-inflammatory activity of apigenin. *Antioxid* 8, 35. doi: 10.3390/antiox8020035

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J.* 53, 814–827. doi: 10.1111/j.1365–313X.2007.03373.x

Gullner, G., Komives, T., Király, L., and Schröder, P. (2018). Glutathione s-transferase enzymes in plant-pathogen interactions. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01836

Gündeli, M. A., Kafkas, N. E., Güney, M., and Ercişli, S. (2021). Determination of phytochemicals from fresh fruits of fig (*Ficus carica* L.) at different maturity stages. *Acta Sci. Pol. Hortorum. Cultus.* 20, 73–81. doi: 10.24326/asphc.2021.2.8

Guo, Z. B., Han, B. M., and Guo, Q. (2017). Effects of methyl jasmonate on high temperature stress and biological encroachment of fig seedlings. *JiangSu. Agric. Sci.* 45, 138–143. doi: 10.15889/j.issn.1002–1302.2017.20.033

Hajam, T. A., and Saleem, H. (2022). Phytochemistry, biological activities, industrial and traditional uses of fig (*Ficus carica*): A review. *Chem. Biol. Interact.* 368, 110237. doi: 10.1016/j.cbi.2022.110237

Hao, G., Dong, Y., Huo, R., Wen, K., Zhang, Y., and Liang, G. (2016). Rutin inhibits neuroinflammation and provides neuroprotection in an experimental rat model of subarachnoid hemorrhage, possibly through suppressing the RAGE-NF-κB inflammatory signaling pathway. *Neurochem. Res.* 41, 1496–1504. doi: 10.1007/s11064–016-1863–7

Harzallah, A., Bhouri, A. M., Amri, Z., Soltana, H., and Hammami, M. (2016). Phytochemical content and antioxidant activity of different fruit parts juices of three figs (*Ficus carica* L.) varieties grown in Tunisia. *Ind. Crops Prod.* 83, 255–267. doi: 10.1016/j.indcrop.2015.12.043

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

Hoxha, L., and Kongoli, R. (2016). Evaluation of antioxidant potential of Albanian fig varieties "Kraps Zi" and "Kraps Bardhe" cultivated in the region of Tirana. *Int. J. Hydrog. Energy* 16, 70–74.

Hoxha, L., Kongoli, R., and Dervishi, J. (2022). Influence of maturity stage on polyphenolic content and antioxidant activity of fig (*Ficus carica* L.) fruit in native Albanian varieties. *Chem. Proc.* 10, 49. doi: 10.3390/IOCAG2022–12199

Hssaini, L., Charafi, J., Razouk, R., Hernández, F., and Hanine, H. (2020). Assessment of morphological traits and fruit metabolites in eleven fig varieties (*Ficus carica* L.). *Int. J. Fruit Sci.* 20, 8–28. doi: 10.1080/15538362.2019.1701615

Hssaini, L., Elfazazi, K., Razouk, R., Ouaabou, R., Hernandez, F., Hanine, H., et al. (2021). Combined effect of cultivar and peel chromaticity on figs' primary and secondary metabolites: preliminary study using biochemical and ftir fingerprinting coupled to chemometrics. *Biology* 10, 573. doi: 10.3390/biology10070573

Jaina, R., Jain, S., Bhagchandani, T., and Yadav, N. (2013). New furanocoumarins and other chemical constituents from *Ficus carica* root heartwood. *Z. für. Naturforschung. C.* 68, 3–7. doi: 10.1515/znc-2013–1-201

Jez, J., and Noel, J. (2001). Mechanism of chalcone synthase - pK(a) of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. *J. Biol. Chem.* 275, 39640–39646. doi: 10.1074/jbc.M008569200

Jucá, M., Cysne, F. F., de Almeida, J. C., Mesquita, D. D. S., Barriga, J. R. M., Dias, K. C. F., et al. (2020). Flavonoids: biological activities and therapeutic potential. *Nat. Prod. Res.* 34, 692–705. doi: 10.1080/14786419.2018.1493588

Kamiloglu, S., and Çapanoğlu, E. (2015). Polyphenol content in figs (*Ficus carica* L.): effect of sun-drying. *Int. J. Food Prop.* 18, 521–535. doi: 10.1080/10942912.2013.833522

Kaulmann, A., and Bohn, T. (2014). Carotenoids, inflammation, and oxidative stress–implications of cellular signaling pathways and relation to chronic disease prevention. *Nutr. Res.* 34, 907–929. doi: 10.1016/j.nutres.2014.07.010

Kmeli, N., Hamdi, J., and Bouktila, D. (2023). Genome-wide characterization of *Ficus carica* MADS-box transcription factors with a focus on their roles during fruit development. *Hortic. Environ. Biotechnol.* 64, 311–329. doi: 10.1007/s13580–022-00478–8

Lama, K., Harlev, G., Shafran, H., Peer, R., and Flaishman, M. A. (2020). Anthocyanin accumulation is initiated by abscisic acid to enhance fruit color during fig (*Ficus carica* L.) ripening. *J. Plant Physiol.* 251, 153192. doi: 10.1016/j.jplph.2020.153192

Lama, K., Yadav, S., Rosianski, Y., Shaya, F., Lichter, A., Chai, L., et al. (2019). The distinct ripening processes in the reproductive and non-reproductive parts of the fig syconium are driven by ABA. *J. Exp. Bot.* 70, 115–131. doi: 10.1093/jxb/ery333

Lazreg-Aref, H., Mars, M., Fekih, A., Aouni, M., and Said, K. (2012). Chemical composition and antibacterial activity of a hexane extract of Tunisian caprifig latex from the unripe fruit of Ficus carica. *Pharm. Biol.* 50, 407–412. doi: 10.3109/13880209.2011.608192

Li, J., An, Y., and Wang, L. (2020). Transcriptomic analysis of *Ficus carica* peels with a focus on the key genes for anthocyanin biosynthesis. *Int. J. M. Sci.* 21, 1245. doi: 10.3390/ijms21041245

Li, J., Ma, N., An, Y., and Wang, L. (2021). *FcMADS9* of fig regulates anthocyanin biosynthesis. *Sci. Hortic.* 278, 109820. doi: 10.1016/j.scienta.2020.109820

Liu, Y. P., Guo, J. M., Yan, G., Zhang, M. M., Zhang, W. H., Qiang, L., et al. (2019). Anti-inflammatory and antiproliferative prenylated isoflavone derivatives from the fruits of *Ficus carica*. *J. Agric. Food. Chem.* 67, 4817–4823. doi: 10.1021/acs.jafc.9b00865

Liu, H. P., Li, D., Xu, W. C., Fu, Y., Liao, R., Shi, J., et al. (2020). Effect of packaging design with two functional films on the quality of figs stored at ambient temperature. *J. Food Process. Preserv.* 44, e14836. doi: 10.1111/jfpp.14836

Liu, L., Zheng, S., Yang, D., and Zheng, J. (2023). Genome-wide in silico identification of glutathione S-transferase (GST) gene family members in fig (*Ficus carica* L.) and expression characteristics during fruit color development. *PeerJ* 11, e14406. doi: 10.7717/peerj.14406

Magalhães, A. F., Tozzi, A. M. G. A., Magalhães, E. G., and Souza-Neta, L. C. (2006). New prenylated metabolites of Deguelia longeracemosa and evaluation of their antimicrobial potential. *Planta. Med.* 72, 358–363. doi: 10.1055/s-2005–916232

Maghsoudlou, E., Esmaeilzadeh Kenari, R., and Raftani Amiri, Z. (2017). Evaluation of antioxidant activity of fig (*Ficus carica*) pulp and skin extract and its application in enhancing oxidative stability of canola oil. *J. Food Process. Pres.* 41, e13077. doi: 10.1111/jfpp.13077

Mahmoudi, S., Khali, M., Benkhaled, A., Boucetta, I., Dahmani, Y., Attallah, Z., et al. (2018). Fresh figs (*Ficus carica* L.): pomological characteristics, nutritional value, and phytochemical properties. *Eur. J. Hortic. Sci.* 83, 104–113. doi: 10.17660/eJHS.2018/83.2.6

Mao, D., Zhong, L., Zhao, X., and Wang, L. (2023). Function, biosynthesis, and regulation mechanisms of flavonoids in *Ginkgo biloba*. *Fruit Res.* 3, 18. doi: 10.48130/FruRes-2023–0018

Marrelli, M., Menichini, F., Statti, G. A., Bonesi, M., Duez, P., Menichini, F., et al. (2012). Changes in the phenolic and lipophilic composition, in the enzyme inhibition and antiproliferative activity of *Ficus carica* L. cultivar Dottato fruits during maturation. *Food Chem. Toxicol.* 50, 726–733. doi: 10.1016/j.fct.2011.12.025

Mawa, S., Husain, K., and Jantan, I. (2013). *Ficus carica* L. (Moraceae): Phytochemistry, traditional uses and biological activities. *Evid. Based. Complementary. Altern. Med.* 2013, 974256. doi: 10.1155/2013/974256

Mopuri, R., and Islam, M. S. (2016). Antidiabetic and anti-obesity activity of *Ficus carica: in vitro* experimental studies. *Diabetes Metab. J.* 42, 300–300. doi: 10.1016/j.diabet.2016.07.020

Mordente, A., Guantario, B., Meucci, E., Silvestrini, A., Lombardi, E., Martorana, G. E., et al. (2011). Lycopene and cardiovascular diseases: an update. *Curr. Med. Chem.* 18, 1146–1163. doi: 10.2174/092986711795029717

Mori, K., Shirasawa, K., Nogata, H., Hirata, C., Tashiro, K., Habu, T., et al. (2017). Identification of RAN1 orthologue associated with sex determination through whole genome sequencing analysis in fig (*Ficus carica* L.). *Sci. Rep.* 7, 41124. doi: 10.1038/srep41124

Munakata, R., Kitajima, S., Nuttens, A., Tatsumi, K., Takemura, T., Ichino, T., et al. (2020). Convergent evolution of the UbiA prenyltransferase family underlies the independent acquisition of furanocoumarins in plants. *New Phytol.* 225, 2166–2182. doi: 10.1111/nph.16277

Oliveira, A. P., Silva, L. R., Andrade, P. B., Valentão, P., Silva, B. M., Gonçalves, R. F., et al. (2010a). Further insight into the latex metabolite profile of *Ficus carica*. *J. Agr. Food Chem.* 58, 10855–10863. doi: 10.1021/jf1031185

Oliveira, A. P., Silva, L. R., Ferreres, F., Pinho, P. G. D., Valentão, P., and Silva, B. M. (2010b). Chemical assessment and in *vitro* antioxidant capacity of *Ficus carica* latex. *J. Agr. Food Chem.* 58, 3393–3398. doi: 10.1021/jf9039759

Oliveira, A. P., Silva, L. R., Guedes de Pinho, P., Gil-Izquierdo, A., Valentão, P., Silva, B. M., et al. (2010c). Volatile profiling of *Ficus carica* varieties by HS-SPME and GC-IT-MS. *Food Chem.* 123, 548–557. doi: 10.1016/j.foodchem.2010.04.064

Palmeira, L., Pereira, C., Dias, M. I., Abreu, R. M. V., Corrêa, R. C. G., Pires, T. C. S. P., et al. (2019). Nutritional, chemical and bioactive profiles of different parts of a Portuguese common fig (*Ficus carica* L.) variety. *Food Res. Int.* 126, 108572. doi: 10.1016/j.foodres.2019.108572

Parhiz, H., Roohbakhsh, A., Soltani, F., Rezaee, R., and Iranshahi, M. (2015). Antioxidant and anti-inflammatory properties of the citrus flavonoids hesperidin and hesperetin: an updated review of their molecular mechanisms and experimental models. *Phytother. Res.* 29, 323–331. doi: 10.1002/ptr.5256

Pereira, C., López-corrales, M., Serradilla, M. J., Villalobos, M. D., Ruíz-Moyano, S., and Martín, A. (2017). Influence of ripening stage on bioactive compounds and antioxidant activity in nine fig (*Ficus carica* L.) varieties grown in Extremadura, Spain. *J. Food Compost. Anal.* 64, 203–212. doi: 10.1016/j.jfca.2017.09.006

Petkova, N., Ivanov, I., and Denev, P. (2019). Changes in phytochemical compounds and antioxidant potential of fresh, frozen, and processed figs (*Ficus carica* L.). *Int. Food Res. J.* 26, 1881–1888.

Pratyusha, D. S., and Sarada, D. V. L. (2022). MYB transcription factors-master regulators of phenylpropanoid biosynthesis and diverse developmental and stress responses. *Plant Cell Rep.* 41, 2245–2260. doi: 10.1007/s00299–022-02927–1

Ralston, L., Subramanian, S., Matsuno, M., and Yu, O. (2005). Partial reconstruction of flavonoid and isoflavonoid biosynthesis in yeast using soybean type I and type II chalcone isomerases. *Plant Physiol.* 137, 1375–1388. doi: 10.1104/pp.104.054502

Reinsvold, R. E. (2010). *Expression of the Artemisia annua beta-caryophyllene synthase in Synechocystis sp. strain PCC6803 results in cyanobacterial sesquiterpene accumulation*. (master's thesis). University of Northern Colorado, Greeley (CO).

Rodov, V., Vinokur, Y., and Horev, B. (2012). Brief postharvest exposure to pulsed light stimulates coloration and anthocyanin accumulation in fig fruit (*Ficus carica* L.). *Postharvest. Biol. Technol.* 68, 43–46. doi: 10.1016/j.postharvbio.2012.02.001

Rubnov, S., Kashman, Y., Rabinowitz, R., Schlesinger, M., and Mechoulam, R. (2001). Suppressors of cancer cell proliferation from fig (*Ficus carica*) resin: isolation and structure elucidation. *J. Nat. Prod.* 64, 993–996. doi: 10.1021/np000592z

Russo, F., Caporaso, N., Paduano, A., and Sacchi, R. (2014). Phenolic compounds in fresh and dried figs from Cilento (Italy), by considering breba crop and full crop, in comparison to Turkish and Greek dried figs. *J. Food Sci.* 79, C1278–C1284. doi: 10.1111/1750–3841.12505

Sandhu, A. K., Islam, M., Edirisinghe, I., and Burton-Freeman, B. (2023). Phytochemical composition and health benefits of figs (fresh and dried): a review of literature from 2000 to 2022. *Nutrients* 15, 2623. doi: 10.3390/nu15112623

Shahrajabian, M. H., Sun, W., and Cheng, Q. (2021). A review of chemical constituents, traditional and modern pharmacology of fig (*Ficus carica* L.), a super fruit with medical astonishing characteristics. *Pol. J. Agron.* 44, 22–29. doi: 10.26114/PJA.IUNG.452.2021.452.04

Siyadatpanah, A., Mirzaei, F., Hossain, R., Islam, M. T., Fatemi, M., and Norouzi, R. (2022). Anti-parasitic activity of the olea europaea and *Ficus carica* on leishmania major: new insight into the anti-leishmanial agents. *Biologia* 77, 1795–1803. doi: 10.1007/s11756–022-01066-y

Slatnar, A., Klancar, U., Stampar, F., and Veberic, R. (2011). Effect of drying of figs (*Ficus carica* L.) on the contents of sugars, organic acids, and phenolic compounds. *J. Agric. Food Chem.* 59, 11696–11702. doi: 10.1021/jf202707y

Solomon, A., Golubowicz, S., Yablowicz, Z., Grossman, S., Bergman, M., Gottlieb, H. E., et al. (2006). Antioxidant activities and anthocyanin content of fresh fruits of common fig (*Ficus carica* L.). *J. Agric. Food Chem.* 54, 7717–7723. doi: 10.1021/jf060497h

Son, J. H., Jin, H., You, H. S., Shim, W. H., Kim, J. M., Kim, G. W., et al. (2017). Five cases of phytophotodermatitis caused by fig leaves and relevant literature review. *Ann. Dermatol.* 29, 86–90. doi: 10.5021/ad.2017.29.1.86

Song, M., Wang, H., Wang, Z., Huang, ,. H. ,. T., Chen, S. W., and Ma, H. Q. (2021). Genome-wide characterization and analysis of bHLH transcription factors related to anthocyanin biosynthesis in fig. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.730692

Su, Q., Rowley, K. G., Itsiopoulos, C., and O'Dea, K. (2002). Identification and quantitation of major carotenoids in selected components of the Mediterranean diet: green leafy vegetables, figs and olive oil. *Eur. J. Clin. Nutr.* 56, 1149–1154. doi: 10.1038/sj.ejcn.1601472

Takahashi, T., Okiura, A., and Kohno, M. (2017). Phenylpropanoid composition in fig (*Ficus carica* L.) leaves. *J. Nat. Med.* 71, 770–775. doi: 10.1007/s11418–017-1093–6

Tanwar, B., Andallu, B., and Modgil, R. (2014). Influence of processing on physicochemical, nutritional and phytochemical composition of *Ficus carica* L. (fig) products. *Asian J. Dairy. Foods. Res.* 33, 37–43. doi: 10.5958/j.0976–0563.33.1.008

Usai, G., Mascagni, F., Giordani, T., Vangelisti, A., Bosi, E., Zuccolo, A., et al. (2020). Epigenetic patterns within the haplotype phased fig (*Ficus carica* L.) genome. *Plant J.* 102, 600–614. doi: 10.1111/tpj.14635

Vallejo, F., Marín, J. G., and Tomás-Barberán, F. A. (2012). Phenolic compound content of fresh and dried figs (*Ficus carica* L.). *Food Chem.* 130, 485–492. doi: 10.1016/j.foodchem.2011.07.032

Veberic, R., Colaric, M., and Stampar, F. (2008). Phenolic acids and flavonoids of fig fruit (*Ficus carica* L.) in the northern Mediterranean region. *Food Chem.* 106, 153–157. doi: 10.1016/j.foodchem.2007.05.061

Villard, C., Larbat, R., Munakata, R., and Hehn, A. (2019). Defence mechanisms of *Ficus*: pyramiding strategies to cope with pests and pathogens. *Planta* 249, 617–633. doi: 10.1007/s00425–019-03098–2

Villard, C., Munakata, R., Kitajima, S., van Velzen, R., Schranz, M. E., Larbat, R., et al. (2021). A new P450 involved in the furanocoumarin pathway underlies a recent case of convergent evolution. *New Phytol.* 231, 1923–1939. doi: 10.1111/nph.17458

Wang, Z., Song, M., Li, Y., Chen, S., and Ma, H. (2019). Differential color development and response to light deprivation of fig (*Ficus carica* L.) syconia peel and female flower tissues: transcriptome elucidation. *BMC Plant Biol.* 19, 217. doi: 10.1186/s12870–019-1816–9

Wang, Z., Song, M., Wang, Z., Chen, S., and Ma, H. (2021). Metabolome and transcriptome analysis of flavor components and flavonoid biosynthesis in fig female flower tissues (*Ficus carica* L.) after bagging. *BMC Plant Biol.* 21, 396. doi: 10.1186/s12870–021-03169–1

Yahiaoui, S., Kati, D. E., Chaalal, M., Ali, L. M. A., Cheikh, K. E., Depaepe, G., et al. (2024). Antioxidant, antiproliferative, anti-inflammatory, and enzyme inhibition potentials of *Ficus carica* wood bark and related bioactive phenolic metabolites. *Wood Sci. Technol* 58, 1-25. doi: 10.1007/s00226-024-01549-y

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

Yin, W. P., Chen, H. M., Wang, T. X., and Cai, M. S. (1997). 9,19-Cyclopropyl-24,25-epoxyethane-5-ene-3β the chemical structure and anticancer activity of spirosterols. *Chin. J. Med. Chem.* 7, 49–50. doi: 10.14142/j.cnki.cn21–1313/r.1997.01.010

Zhang, Y., Wan, Y., Huo, B., Li, B., Jin, Y., and Hu, X. (2018). Extracts and components of *Ficus carica* leaves suppress survival, cell cycle, and migration of triple-negative breast cancer MDA-MB-231 cells. *Oncotarg. Ther.* 11, 4377–4386. doi: 10.2147/OTT.S17160

Zhang, W. J., Xie, B. P., Li, W. J., Li, J. P., Gan, G. X., and Zhang, J. (2017). Anti-inflammatory and antiproliferative prenylated isoflavone derivatives from the fruits of *Ficus carica*. *J. Agric. Food. Chem.* 67, 4817–4823. doi: 10.1021/acs.jafc.9b00865

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 02 August 2024
DOI 10.3389/fpls.2024.1427359

OPEN ACCESS

EDITED BY
Xiaomeng Liu,
Wuhan Polytechnic University, China

REVIEWED BY
Wei Li,
Chinese Academy of Agricultural Sciences, China
Gen Li,
Xi'an Botanical Garden of Shaanxi Province, China

*CORRESPONDENCE
Zhenming Yu
yuzhenming@zcmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 03 May 2024
ACCEPTED 17 July 2024
PUBLISHED 02 August 2024

CITATION
Xu Z, Zhang G, Chen J, Ying Y, Yao L, Li X, Teixeira da Silva JA and Yu Z (2024) Role of *Rubus chingii* BBX gene family in anthocyanin accumulation during fruit ripening. *Front. Plant Sci.* 15:1427359. doi: 10.3389/fpls.2024.1427359

Role of *Rubus chingii* BBX gene family in anthocyanin accumulation during fruit ripening

Zhangting Xu[1†], Guihua Zhang[2†], Junyu Chen[1], Yuxin Ying[3], Lingtiao Yao[1], Xiaoxian Li[1], Jaime A. Teixeira da Silva[4] and Zhenming Yu[1,5*]

[1]School of Pharmaceutical Sciences, Academy of Chinese Medical Sciences, Zhejiang Chinese Medical University, Hangzhou, China, [2]Zhejiang Academy of Forestry, Hangzhou, China, [3]College of Food and Health, Zhejiang A & F University, Hangzhou, China, [4]Independent Researcher, Miki, Kagawa, Japan, [5]Songyang Institute of Zhejiang Chinese Medical University, Lishui, China

The B-box (BBX) family, which is a class of zinc finger transcription factors, exhibits special roles in plant growth and development as well as in plants' ability to cope with various stresses. Even though *Rubus chingii* is an important traditional medicinally edible plant in east Asia, there are no comprehensive studies of BBX members in *R. chingii*. In this study, 32 RcBBX members were identified, and these were divided into five groups. A collinearity analysis showed that gene duplication events were common, and when combined with a motif analysis of the *RcBBX* genes, it was concluded that group V genes might have undergone deletion of gene fragments or mutations. Analysis of *cis*-acting elements revealed that each *RcBBX* gene contained hormone-, light-, and stress-related elements. Expression patterns of the 32 *RcBBX* genes during fruit ripening revealed that highest expression occurred at the small green fruit stage. Of note, the expression of several *RcBBX* genes increased rapidly as fruit developed. These findings, combined with the expression profiles of anthocyanin biosynthetic genes during fruit ripening, allowed us to identify the nuclear-targeted RcBBX26, which positively promoted anthocyanin production in *R. chingii*. The collective findings of this study shed light on the function of *RcBBX* genes in different tissues, developmental stages, and in response to two abiotic stresses.

KEYWORDS

***Rubus chingii*, BBX, anthocyanidin, expression analysis, fruit ripening**

1 Introduction

Transcription factors (TFs) are a class of proteins that are able to bind to specific genes, allowing for gene expression to be regulated (Spitz and Furlong, 2012). They can act as communication hubs by networking with other interacting proteins, synergizing with them, and resulting in gene transcription (Stortz et al., 2024). Among a wealth of plant TFs, the zinc finger protein is a specialized TF that folds upon itself to form a finger structure, and is composed of Zn (II) and different amounts of histidine (His) and cysteine (Cys). Different zinc finger proteins play varying roles in plants (Laity et al., 2001). For instance, zinc finger proteins are involved in mRNA recognition, DNA packing, transcription activation, and protein-protein interactions, and can thus impact plant growth and development (Noman et al., 2019), while playing an integral role in plants' responses to stresses (Cao et al., 2023).

The B-box domain protein, or BBX, which is one type of zinc finger protein, has attracted widespread attention. In recent years, BBX families have been identified and functionally characterized in the genomes of several plants, such as *Dioscorea opposita* (Chang et al., 2023), *Vaccinium corymbosum* (Liu et al., 2023), *Dendrobium officinale* (Cao et al., 2019), *Glycine max* (Shan et al., 2022), and *Vitis vinifera* (Wei et al., 2020). The BBX family is well-studied in the model plant *Arabidopsis thaliana* at physiological, molecular, and biochemical levels (Khanna et al., 2009). All BBX proteins contain one or two highly conserved B-box domains, B-box1 ($C\text{-}X_2\text{-}C\text{-}X_{7\text{-}8}\text{-}C\text{-}X_2\text{-}D\text{-}X\text{-}A\text{-}X\text{-}L\text{-}C\text{-}X_2\text{-}C\text{-}D\text{-}X_3$-HB) and B-box2 ($C\text{-}X_2\text{-}C\text{-}X_3\text{-}P\text{-}X_4\text{-}C\text{-}X_2\text{-}D\text{-}X_3\text{-}L\text{-}C\text{-}X_2\text{-}C\text{-}D\text{-}X_3$-H), with about 40 amino acids at the N-terminus, and some of them may possess a conserved CCT (CONSTANS, CO-like and TOC1) domain with 42-43 amino acids at the C-terminus (Cao et al., 2023). Based on their structures and functions in *A. thaliana*, the BBX genes have been divided into five subfamilies (groups I, II, III, IV, and V) that are closely associated with a number of B-box and CCT conserved domains (Gangappa and Botto, 2014). Groups I and II contain two B-box domains (B-box1 and B-box2) and one CCT domain, group III contains a B-box1 domain and a CCT domain, group IV contains two B-box domains but no CCT domain, whereas group V only contains a B-box1 domain (Khanna et al., 2009). Hence, addressing the diversity of the BBX family may be conducive to understand its roles.

Increasing evidence has demonstrated that BBX is involved in various developmental processes and in response to environmental stresses (Cao et al., 2023). In Chinese yam (*D. opposita*), overexpression of *DoBBX2* and *DoBBX8* accelerated tuber formation under short-day conditions (8-h photoperiod; light intensity: 38 μmol m^{-2} s^{-1}) while overexpression of *DoBBX8* alone promoted tuber formation in the dark (Chang et al., 2023). In tomato (*Solanum lycopersicum*), SlBBX17 interacted with SlHY5 and stimulated transcription of the *SlHY5* gene, positively enhanced CBF-dependent cold tolerance, whereas silencing of *SlBBX17* promoted susceptibility to low-temperature stress (Song et al., 2023). The expression of 68.4% of *D. officinale BBX* genes increased to varying degrees after treatment with methyl jasmonate (MeJA), especially *DoBBX17*, which was strongly up-regulated (by 65-fold) after 24 h (Cao et al., 2019). In apple (*Malus domestica*), MdBBX20, induced by UV-B and low temperatures, interacted with MdHY5 to activate *MdMYB1* expression, binding directly to the promoters of *MdDFR* and *MdANS* genes, and promoting the accumulation of anthocyanins (Fang et al., 2019). In pear (*Pyrus pyrifolia*), the expression of *PpBBX16* was significantly enhanced under constant white light (light intensity: 13.19 μmol m^{-2} s^{-1}) induction for 10 d, and strongly interacted with PpHY5 to induce the promoter activity of *PpMYB10*, while integrating the well-characterized COP1-HY5-MYB10 regulatory complex, which is involved in the determination of red coloration and anthocyanin biosynthesis (Bai et al., 2019a). Despite these studies in other plants, no information regarding the BBX family is available for the medicinally edible plant, *Rubus chingii* Hu (Rosaceae).

R. chingii, commonly known as hanging hooks in English, acquired its name due to the shape of ripe fruits, which resemble an upside-down hanging bowl on branches. *R. chingii* is used in the Chinese pharmacopoeia, and is widely cultivated in Zhejiang and Jiangxi provinces of China. Not only is it consumed as a fresh fruit when ripe, its unripe dried fruit serves as a medicinal herb (Wang et al., 2022). Ripening *R. chingii* fruit undergoes a change in color from green to red, and this is dependent on relevant TFs that can regulate the biosynthesis and accumulation of anthocyanins (Li et al., 2021). For instance, BBX can directly or indirectly regulate anthocyanin biosynthetic genes at transcriptional and post-transcriptional levels, thereby adjusting the levels of anthocyanin production (Wang et al., 2023). In *A. thaliana* seedlings, overexpression of *BBX21*, *BBX22* and *BBX23* accumulated dramatically more anthocyanin than control seedlings when induced by light, while *BBX24*, *BBX25* and *BBX32* negatively regulated the accumulation of anthocyanins (Job et al., 2018; Bai et al., 2019b; Podolec et al., 2022). Furthermore, *R. chingii* fruit has an abundance of flavonoids, terpenoids, alkaloids, and phenolics (Wang et al., 2021). Collectively, these exhibit a wealth of pharmacological properties, including antibacterial, antioxidant, anti-tumor, and anti-inflammatory (Chen, 2023). The accumulation of these secondary metabolites is potentially related to induction by phytohormones, and may be transcriptionally regulated by the BBX family (Gangappa and Botto, 2014; Song et al., 2020). The expression profiles of the BBX family genes in *R. chingii* are still unknown.

The objective of the present study was to identify the BBX family from *R. chingii* at the genome-wide level. The physicochemical properties, chromosome location, gene structure, phylogenetic relationships, analysis of *cis*-acting elements (CAEs), conserved motifs, gene replication, and protein-protein interactions were systematically evaluated. Furthermore, the expression levels of BBX members in different organs (roots, stems, leaves, flowers, and fruits), at different developmental stages, and following exposure to abscisic acid (ABA), were comparatively evaluated. Consequently, these findings will allow candidate genes to be identified while

providing a theoretical foundation for understanding the role of BBX family genes during the ripening of *R. chingii* fruit.

2 Materials and methods

2.1 Plant materials and hormonal treatment

The erect and medium-sized shrub *R. chingii* Hu, which was identified as such by Professor Xiaoxia Shen (Zhejiang Chinese Medical University), was cultivated in a greenhouse using sandy and clay soil (4:1, v/v) and irrigated every fortnight with 1000-times diluted Hyponex 20-20-20 fertilizer (Hyponex Co., Tokyo, Japan) under natural conditions at the medicinal herb garden (*E* 119.96°, *N* 30.05°) of Zhejiang Chinese Medical University, in Zhejiang Province, China. To detect the organ-specific expression of *RcBBX* genes, roots, stems, leaves, flowers, and unripe fruits of one-year-old *R. chingii* plants were collected in May 2023, quick-frozen in liquid nitrogen, and stored in a -80°; refrigerator (Thermo Scientific, Waltham, MA, USA). To investigate the dynamic expression of *RcBBX* genes during fruit ripening, small green (5-6 mm in diameter), big green (11-13 mm in diameter), yellow and red fruits were harvested. Their surfaces were wiped clean with a dry cloth, then stored in a -80°; refrigerator. To understand the effect of ABA on the transcript levels of *RcBBX* genes, ten leaves of one-year-old *R. chingii* plants from top to bottom were sampled after treatment with 100 μM ABA (Sigma-Aldrich, St. Louis, MO, USA) 10 mL for each plant over 6 h, then stored in a -80°; refrigerator. Triplicate samples were prepared.

2.2 Identification of *R. chingii* BBX family members at the genome-wide level

The chromosome-scale genome of *R. chingii* was retrieved from the Rosaceae Genome Database (https://www.rosaceae.org) with accession number tfGDR1051. A BLASTP search (score value ≥ 100, e-value ≤ 1e-10) was employed to identify all possible *R. chingii* BBX members by utilizing the 32 published *A. thaliana* BBX proteins (Khanna et al., 2009) as queries. Each candidate BBX was subsequently verified with the Conserved Domain Database (CDD, www.ncbi.nlm.nih.gov/cdd) and Simple Modular Architecture Research Tool (SMART, https://smart.embl-heidelberg.de/). Redundant sequences or incomplete sequences, excluding the conserved BBX domain (PF00643) which was retrieved from Pfam (https://pfam-legacy.xfam.org/), were excluded. Consequently, all 32 RcBBX members with complete BBX domains were identified (Supplementary Table 1). In addition, physical and chemical properties of all RcBBX proteins, including their molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index, and grand average of hydrophobicity, were investigated with the ExPASy server (https://www.expasy.org). The subcellular localization of all RcBBX members was predicted using the WoLF server (https://wolfpsort.hgc.jp/) and the Plant-mPLoc predictor (Chou and Shen, 2010).

2.3 Phylogenetic analysis of RcBBX members and multiple sequence alignment

The *A. thaliana* BBX sequences were identified at the *Arabidopsis* Information Resource (https://www.arabidopsis.org/). The full-length amino acid sequences of the BBX proteins from *A. thaliana* and *R. chingii* were subjected to multiple sequence alignment using ClustalX 2.1 (www.clustal.org/). A phylogenetic tree was established using Molecular Evolutionary Genetics Analysis (MEGA) software (Tamura et al., 2021) applying a neighbor-joining (NJ) algorithm (Saitou and Nei, 1987) with 1000 bootstrap replicates, and visualized using the iTOL online tool (https://itol.embl.de).

2.4 *Cis*-acting elements, gene structure, and conserved motifs of RcBBX members

The 2000-bp long upstream sequences of the 32 *RcBBX* genes were mined from the *R. chingii* genome (Wang et al., 2021), sorted in numerical order and inputted into the PlantCARE platform (Lescot et al., 2002) to predict the potential CAEs. According to the annotation information of the *R. chingii* reference genome (Wang et al., 2021), the structure of *RcBBX* genes was revisualized with the TBtools program (Chen et al., 2020). To better understand the structure of these RcBBX members, the Multiple Expectation Maximization for Motif Elicitation website (MEME, https://meme-suite.org/) was employed to determine their conserved motifs (motif width: 6-50; max number: 10), and visualized with TBtools software (Chen et al., 2020).

2.5 Chromosome distribution, duplications and synteny analysis of RcBBX members

Based on the relative positions of the 52 characterized *RcBBX* genes on the seven *R. chingii* chromosomes, the TBtools program (Chen et al., 2020) was applied to map them on the chromosomes. To explore gene replication events among *RcBBX* genes, collinearity analysis of RcBBX members was performed using the multiple collinear scanning toolkit MCScanX (Wang et al., 2012) with default parameters. The Dual Synteny Plotter in TBtools tool (Chen et al., 2020) was used to carry out a syntenic analysis of three BBX gene families in *R. chingii*, *A. thaliana*, and rice (*O. sativa*).

2.6 Identification of enzyme-encoding genes related to anthocyanidin biosynthesis

The anthocyanidin biosynthetic pathway in *R. chingii* was putatively implicated based on an assessment of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (Kanehisa et al., 2023) and previously published reports regarding flavonoid biosynthesis (Li et al., 2021; Lei et al., 2023). Finally,

information about the transcriptome (NCBI accession no. PRJNA671545) of *R. chingii* fruits during different developmental stages (small green, big green, yellow, and red fruits) were retrieved to appreciate the enzyme-encoding genes involved in the biosynthesis of flavonoids. The TBLASTN server (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was employed to align homologous genes against the enzyme-encoding genes. Redundant and incomplete sequences without functional domains were removed. Finally, enzyme-encoding genes were selected from the *R. chingii* transcriptome database when their FPKM values exceeded 5.0.

2.7 GO and KEGG enrichment analysis of RcBBX sequences

All RcBBX sequences were uploaded into the eggNOG-mapper website (https://eggnog-mapper.embl.de/). The functionally annotated results were downloaded, selected with a *P*-value less than 0.05, and the GO and KEGG enrichment information was visualized on the Chiplot server (https://www.chiplot.online/).

2.8 Expression profiles of *RcBBX* genes in different organs, at different growth stages, and following treatment with ABA

Total RNA was isolated from the lyophilized powder of different organs (roots, stems, leaves, flowers, and fruits), different developmental fruits (small green, big green, yellow and red fruits), and ABA-treated *R. chingii* leaves, using the Quick RNA Isolation Kit (Huayueyang Biotechnology Co., Beijing, China) according to the company's instruction manual (Yu et al., 2021a). Thereafter, the quality and integrity of RNA were estimated using agarose gel electrophoresis (Bio-Rad Laboratories, Hercules, CA, USA) and a NanoDrop 2000 spectrophotometer (Thermo Scientific). The acquired 1 μg of RNA was reverse-transcribed to first-strand cDNA using the PrimeScript Reagent Kit with gDNA Eraser (Takara, Dalian, China). Real-time quantitative polymerase chain reaction (qRT-PCR) was implemented on an ABI 7500 (Applied Biosystems, Foster City, CA, USA) at the following conditions: 95°C for 30 s, followed by 35 cycles of 95°C for 2 s, and 60°C for 30 s. A total of 10 μL detection solution included 5 μL of 2× iTaq™ Universal SYBR® Green Supermix (Bio-Rad Laboratories), 50 ng of obtained cDNA, 500 nM of forward/reverse primer, and the remaining volume with deionized water. Relative transcript abundance was assessed based on the $2^{-\Delta\Delta Ct}$ protocol (Livak and Schmittgen, 2001), and the elongation factor-1 alpha gene (*EF-1α*) was employed as the house-keeping gene (Wang et al., 2021). All primers that were used are indicated in Supplementary Table 1.

2.9 Molecular cloning, sequence analysis and subcellular localization of RcBBX26

The sequences encoded by *RcBBX26* were propagated and purified from *R. chingii* fruits using the ApexHF HS DNA Polymerase FS (Accurate Biotechnology (Hunan) Co., Ltd., Changsha, China). The secondary structure of RcBBX26 was visualized at the SOPM webserver (https://npsa-pbil.ibcp.fr). The subcellular localization of RcBBX26 was assessed as described previously (Yu et al., 2021a), NLS-mCherry was used as a nuclear localization marker, and YFP fluorescence was observed and images were captured under a TCS SP8 STED microscope (Leica Camera AG, Solms, Germany).

2.10 Transient overexpression of *RcBBX26* in *R. chingii* leaves

To induce gene overexpression, a 765-bp coding sequence of *RcBBX26* without a stop codon (TGA) was introduced into the pCAMBIA1301 vector (CAMBIA, Canberra, Australia) at the *EcoR*I and *BamH*I sites using the In-Fusion solution (Takara). After verification by sequencing (Zhejiang SUNYA Co., Hangzhou, China), the pCAMBIA1301-RcBBX26 recombinant was introduced into *Agrobacterium tumefaciens* GV3101 (pSoup-p19; Weidi Biotechnology Co., Shanghai, China) using a previously published freeze-thaw protocol (Yu et al., 2021b). Infiltration solution, containing 10 mM 2-morpholinoethanesulfonic acid (MES; Sigma-Aldrich), 10 mM $MgCl_2$ (Sigma-Aldrich), and 20 mM acetosyringone (Sigma-Aldrich) (pH = 5.6) was injected into the third leaves from the terminus of one-year-old *R. chingii* plants. Total RNA was isolated from MOCK and overexpressing *RcBBX26* leaves as mentioned above. Semi-quantitative RT-PCR was performed following thermocycling as initially published (Yu et al., 2021b). Eventually, positive leaves overexpressing *RcBBX26* were verified by semi-quantitative RT-PCR and qRT-PCR as indicated above.

2.11 Determination of total anthocyanidin content during fruit ripening

To investigate total anthocyanidin content during fruit ripening, four developmental fruits (small green, big green, yellow and red fruits) of one-year-old greenhouse *R. chingii* were harvested in May 2023. Anthocyanin content in different *R. chingii* fruits was determined by a pH differential protocol (Yu et al., 2018). Dried powder (1 g) derived from a vertical-type pulverizer (Xinda Machinery Co., Jiangyin, China) was mixed with 2 mL methanol (Sigma-Aldrich), mixed with 1.0% formic acid (Sigma-Aldrich), and extracted by ultrasonication with a 40 Hz ultrasonic homogenizer (Ningbo Scientz Biotechnology Co., Ningbo, China) in the mixture of ice and water for 15 min. The supernatant was obtained after centrifugation (4000 ×*g*) at 25°. Anthocyanidin content was spectroscopically evaluated at 510 and 700 nm in separate acidic buffers at pH 1.0 and 4.5, respectively. The formula utilized to quantify anthocyanin was: $A = (A_{510nm} - A_{700nm})_{pH1.0} - (A_{510nm} - A_{700nm})_{pH4.5}$. Cyanidin 3-*O*-rutinose (CAS no. 18719-76-1; Sigma-Aldrich) was used as the reference standard to establish a standard curve ($A = 2.3016C + 0.021$, R^2=0.9991; C indicates the concentration of cyanidin 3-*O*-rutinose). Total anthocyanidin content of different *R. chingii* fruits was expressed as mg of cyanidin 3-*O*-rutinose equivalents per gram of dry weight (DW).

2.12 Statistical analysis

All utilized data are presented as the mean ± standard deviation (SD) of no less than three independent replicates. Statistical analysis was executed using SPSS Statistics version 22.0 (IBM Corp., Armonk, NY, USA). In graphs, asterisks above columns indicate statistical differences in expression abundance between CK and ABA, as assessed by a student's *t*-test at $P < 0.01$. Furthermore, different lowercase letters above columns indicate significant differences of anthocyanin content among different fruit ripening stages (small green, big green, yellow and red fruits), as assessed by Duncan's multiple range test at $P < 0.01$. Significant differences between control and treated group were evaluated by student's *t*-test ($P < 0.01$). Heatmaps of the expression abundance of all *RcBBX* genes at different tissues, or at different ripening stages, were charted using TBtools software (Chen et al., 2020), and color scales indicate the log2-transformation of average expression levels, with high expression noted by red/orange and low expression indicated by blue/cyan. A correlation of expression abundance between all *RcBBX* genes and anthocyanidin biosynthetic genes was implemented using Pearson's correlation coefficient (*r*) at $P < 0.05$.

3 Results

3.1 Identification and characterization of RcBBX protein family members

To mine BBX members on the seven chromosomes in the *R. chingii* genome, local Hidden Markov Model searches with the BBX core domain (PF00643) and a BLAST with all 32 A. *thaliana* BBX proteins as queries were carried out. The obtained BBX proteins were confirmed using the Simple Modular Architecture Research Tool and an NCBI CD-Search. Thereafter, redundant or erroneous sequences without PF00643 were removed. A total of 32 BBX members were finally identified from the *R. chingii* genome and sequentially named RcBBX1-RcBBX32 according to their chromosomal positions (Table 1; Supplementary Table 2).

The RcBBX proteins ranged from 173 aa (RcBBX25) to 866 aa (RcBBX16) in length, and their molecular weight ranged from 18.89 to 95.32 kDa. The pI ranged from 4.06 (RcBBX14) to 9.45 (RcBBX19), similar to the pIs of BBX proteins from soybean (Shan et al., 2022) and grapevine (Wei et al., 2020). The number of unstable proteins was high (29/32), suggesting that these RcBBX proteins, with an instability index > 40, exhibited low overall stability. The grand average of hydropathicity of RcBBX proteins was less than 0, suggesting that they are hydrophilic. Most of the RcBBX proteins (23/32) were located in the nucleus, although some of them were predicted to be located in the cytosol or chloroplast, indicating that these proteins might function as TFs.

3.2 Phylogenetic analysis of RcBBX protein family members

To explore the classification and evolution of these RcBBX family members, the 32 identified RcBBX proteins, together with 32 A. *thaliana* AtBBX proteins, 26 *Cucumis melo* CmBBX proteins, 26 C. *sativus* CsBBX proteins, 30 *Solanum lycopersicum* SlBBX proteins, and 31 *Zea mays* ZmBBX proteins (Supplementary Table 3), were sequentially aligned to assess their phylogenetic relationships (Figure 1). According to the phylogenetic analysis, and confirming

TABLE 1 Physicochemical properties of RcBBX proteins in *R. chingii*.

Name	Length (aa)	Molecular weight (kD)	pI	Instability index	Aliphatic index	Grand average of hydropathicity	Subcellular localization
RcBBX1	687	75.16	5.32	53.25	60.36	-0.669	Nucleus
RcBBX2	686	75.04	5.32	54.25	62.16	-0.648	Nucleus
RcBBX3	186	20.63	6.49	53.27	70.27	-0.559	Cytosol
RcBBX4	763	82.15	8.61	67.41	59.34	-0.617	Nucleus
RcBBX5	255	29.10	8.68	59.00	72.98	-0.552	Nucleus
RcBBX6	809	88.52	6.61	58.20	58.92	-0.662	Nucleus
RcBBX7	385	42.58	5.62	54.36	57.06	-0.679	Cytosol
RcBBX8	339	37.73	5.82	43.34	70.47	-0,463	Nucleus
RcBBX9	302	32.91	5.83	51.42	69.47	-0.344	Nucleus
RcBBX10	215	24.07	5.98	44.61	70.28	-0,470	Nucleus
RcBBX11	265	30.03	7.10	66.47	73.51	-0.450	Nucleus
RcBBX12	285	30.29	4.72	52.52	64.04	-0.235	Nucleus
RcBBX13	248	27.78	4.40	58.52	58.19	-0.955	Nucleus
RcBBX14	198	21.88	4.06	70.49	38.38	-1.221	Nucleus

(Continued)

10.3389/fpls.2024.1427359

TABLE 1 Continued

Name	Length (aa)	Molecular weight (kD)	pI	Instability index	Aliphatic index	Grand average of hydropathicity	Subcellular localization
RcBBX15	223	25.39	8.69	75.07	68.61	-0.638	Nucleus
RcBBX16	866	95.32	7.57	37.66	89.36	-0.344	Nucleus
RcBBX17	239	26.38	4.75	47.99	75.98	-0.307	Nucleus
RcBBX18	239	26.32	4.76	48.93	76.82	-0.286	Nucleus
RcBBX19	197	23.29	9.45	56.63	70.25	-0.650	Chloroplast
RcBBX20	348	38.16	6.70	47.09	71.47	-0.331	Chloroplast
RcBBX21	432	47.96	5.41	54.68	58.08	-0.749	Chloroplast
RcBBX22	501	56.05	5.92	62.28	65.19	-0.577	Nucleus
RcBBX23	610	68.01	6.10	47.35	69.77	-0.610	Nucleus
RcBBX24	417	45.18	5.43	55.83	58.99	-0.509	Nucleus
RcBBX25	173	18.89	7.04	37.34	83.99	-0.054	Cytosol
RcBBX26	254	27.40	8.61	61.05	74.17	-0.158	Chloroplast
RcBBX27	211	23.85	8.72	62.78	65.12	-0.609	Nucleus
RcBBX28	251	28.14	8.55	54.10	82.03	-0.301	Nucleus
RcBBX29	400	43.98	5.52	36.80	59.80	-0.605	Nucleus
RcBBX30	265	28.99	4.57	57.31	56.04	-0.985	Nucleus
RcBBX31	434	47.38	5.10	58.27	67.19	-0.506	Chloroplast
RcBBX32	243	27.06	8.85	61.86	82.26	-0.190	Chloroplast

a prior classification of BBX members into five subfamilies (Gangappa and Botto, 2014), these BBX members were divided into five groups (groups I, II, III, IV, and V), including 3, 6, 2, 7, and 14 RcBBX members (Supplementary Table 4), respectively. Among them, group II was the smallest subgroup, including 20 BBX proteins, whereas group IV had the most BBX proteins (52).

3.3 Chromosomal location and gene duplication analysis of all *RcBBX* genes

The genomic location of the 32 *RcBBX* genes revealed that they were unevenly mapped across all seven chromosomes (Figure 2). Eight *RcBBX* genes were distributed on chromosome 6, exhibiting the highest number of *RcBBX* members. In addition, seven genes were distributed on chromosome 4, while chromosomes 1, 2, and 3 included five *RcBBX* genes each. Chromosomes 5 and 7 contained only one gene each, *RcBBX23* and *RcBBX32*, respectively.

Duplication analysis demonstrated that six gene pairs (*RcBBX1-RcBBX2*, *RcBBX9-RcBBX10*, *RcBBX13-RcBBX14*, *RcBBX17-RcBBX18*, *RcBBX21-RcBBX23*, *RcBBX14-RcBBX29*) were defined as tandemly-duplicated segments at the chromosome level (Figure 2). Furthermore, four pairs of gene duplications occurred within the same chromosomes (1, 2, 3, and 4) while another two (*RcBBX14-RcBBX29*, *RcBBX21-RcBBX23*) occurred on two different chromosomes (3 and 6, 4 and 5). In addition, collinearity analysis of *BBX* genes among *R. chingii*, *A. thaliana*, and *O. sativa* was sequentially executed (Figure 3). Comparative syntenic maps of the *R. chingii* genome with the genomes of two model plants *A. thaliana* and *O. sativa*, indicated 28 homologous gene pairs between *R. chingii* and *A. thaliana* on 7 chromosomes, and seven homologous gene pairs between *R. chingii* and *O. sativa* on 4 chromosomes.

3.4 Conserved domains, motifs and gene structure analysis of RcBBX protein family members

Apart from RcBBX28 and RcBBX31, nine RcBBX proteins in groups I and II included two B-box structural domains and one CCT structural domain (Figure 4A). Two RcBBX proteins in group III included one B-box and one CCT domain. Most RcBBX proteins (5/7) in group IV contained two B-box structural domains. All RcBBX proteins in group V included only one B-box structural domain (Figure 4B). Therefore, RcBBX proteins within the same subgroup exhibited similar domain compositions.

Furthermore, 12 motifs were observed and named as motifs 1-12 (Figure 4C). Except for 10 proteins in group V, motif 1 was widely distributed in all other RcBBX members. Motif 7 was distributed in groups I, II, III, and IV. Of note, the distribution of motifs in group V was divided into three types: type 1 (4 members), which only included motif 1; type 2 (4 members), which mainly included motifs 2, 3, 5, and 10; type 3 (6 members), which included motifs 4, 8, and 9. Overall, the number and types of motifs within

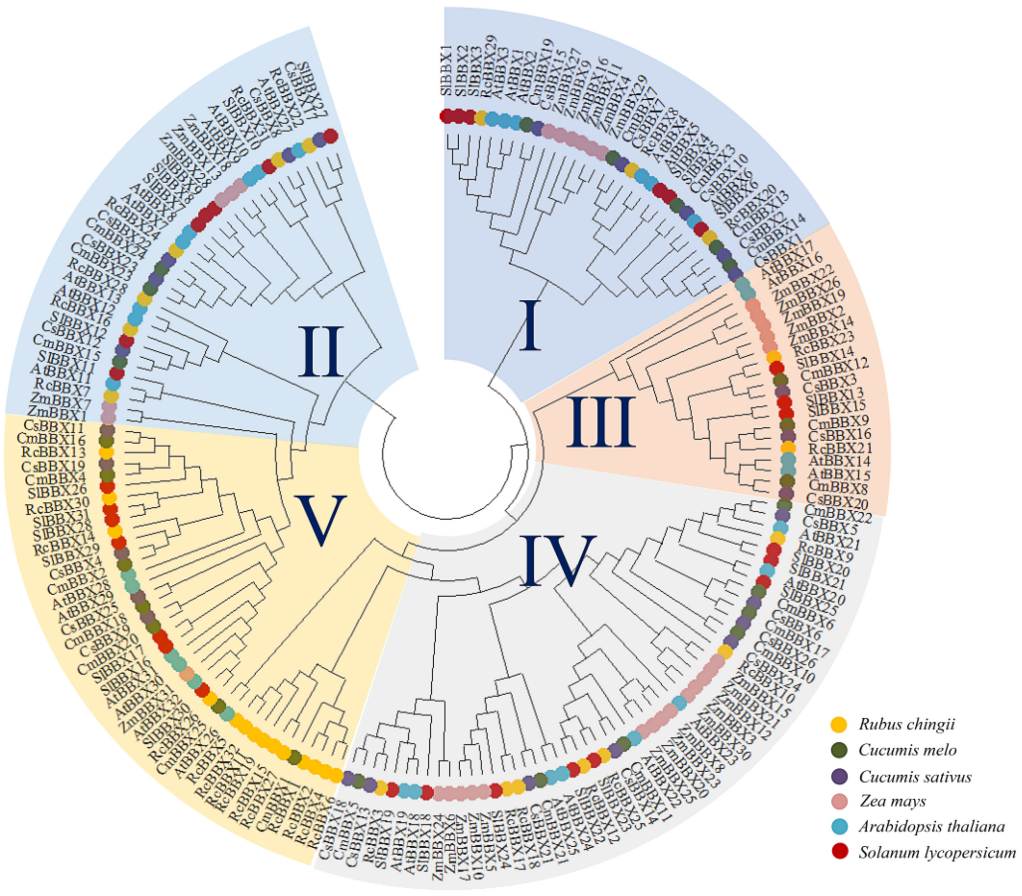


FIGURE 1
Phylogenetic tree of 177 BBX protein family members in *A. thaliana*, *C. melo*, *C. sativus*, *R. chingii*, *S. lycopersicum*, and *Z. mays*. The phylogenetic tree was constructed with MEGA software using a maximum likelihood approach with 1000 bootstrap replications.

the same group were relatively similar, although there were considerable differences among different groups.

The CDS/intron patterns of *RcBBX* genes were also analyzed (Figure 4D). The number of exons in groups I, II, III, and IV was 1-6 (except for 19 exons in *RcBBX16*). Most *RcBBX* genes harbored 1-6 exons in group V whereas five genes (*RcBBX1*, *RcBBX2*, *RcBBX4*, *RcBBX6*, *RcBBX16*) possessed 13-15 exons. Intriguingly, the number of exons was relatively conserved within the same group, for example, group IV genes contained 2-5 exons.

3.5 *Cis*-acting element analysis of the promoter region of *RcBBX* genes

The 2000-bp upstream promoter of the 32 *RcBBX* genes was utilized to analyze the CAEs, which were divided into three categories: those associated with light, hormone, and stress (Figure 5). A total of 836 CAEs were identified, and 405 of them were associated with light responsiveness. For instance, the G-box, a common CAE in plants exposed to external light stimulation, was the most widely distributed CAE, accounting for 33.83% of all light-responsive CAEs. *RcBBX13* contained the most G-boxes (19), followed by *RcBBX26* (12) and *RcBBX27* (10). The I-box was found in 3.7% of all light-responsive CAEs, and together with the G-box, formed a key component with conserved modular array 5 (CMA5) for the regulation of photosensitive pigments. A total of 278 hormone-responsive CAEs were discovered. ABRE, an ABA-responsive element, was the most widely distributed and accounted for 43.17% of all hormone-responsive CAEs. Most *RcBBX* genes (23/32) contained an ABRE, indicating that they were potentially induced by ABA. Similar to light responsiveness, *RcBBX13* contained the most ABREs (16), followed by *RcBBX26* (10) and *RcBBX27* (9). CGTCA and TGACG were MeJA-responsive elements, accounting for 29.5% of all hormone-responsive CAEs. In addition, there were 153 CAEs involved in stress response. ARE and GC motifs, two anaerobic inducible CAEs, were the most common, accounting for 52.94% of all stress-responsive elements. Multiple CAEs were associated with stress responsiveness, such as MBS and LTR, which were involved in drought and low temperature responsiveness (Yu et al., 2021b; Nian et al., 2022), respectively. Hence, it is noteworthy that the promoter regions of each *RcBBX* gene exhibited a different number and composition of CAEs, and might play multifarious roles in photomorphogenesis and stress resistance in *R. chingii*.

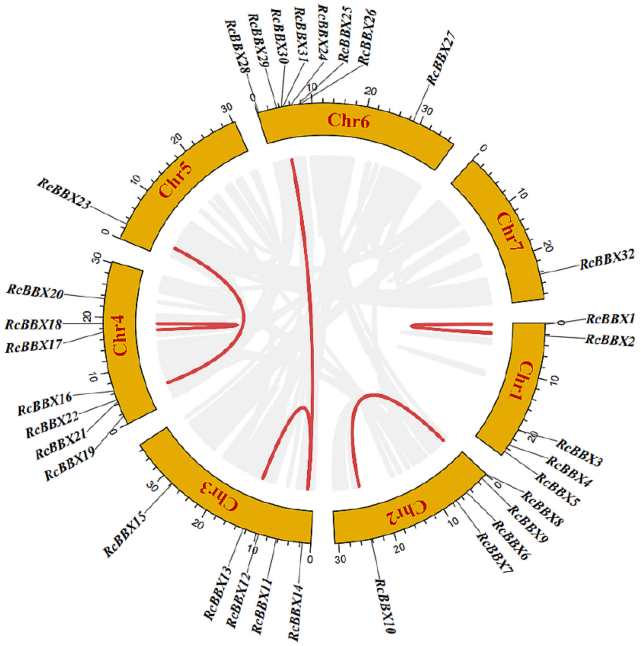


FIGURE 2
Chromosomal location and collinearity analysis of the *RcBBX* genes.

3.6 Protein-protein interaction network of RcBBX proteins

To depict the protein-protein interaction of all RcBBX proteins, the STRING database was employed. Results demonstrate that 24 RcBBX proteins were predicted as being homologous to *A. thaliana* BBX proteins whereas seven other proteins (RcBBX2, RcBBX4, RcBBX6, RcBBX11, RcBBX19, RcBBX27, and RcBBX32) exhibited relatively low connectivity (Figure 6). Among the 17 interacting proteins, the amount of connectivity ranged from 1-12, and the most highly connected protein was RcBBX26, which interacted with 16 other RcBBX proteins. In addition, 2 CONSTANS-like (COL), RcCOL6 and RcCOL7, also interacted with RcBBX proteins.

3.7 GO and KEGG enrichment annotation of RcBBX proteins

As shown in Figure 7A, the 32 RcBBX members had 74 annotated GO terms, and they were categorized into three types, including 66 molecular functions (MF), five cellular components

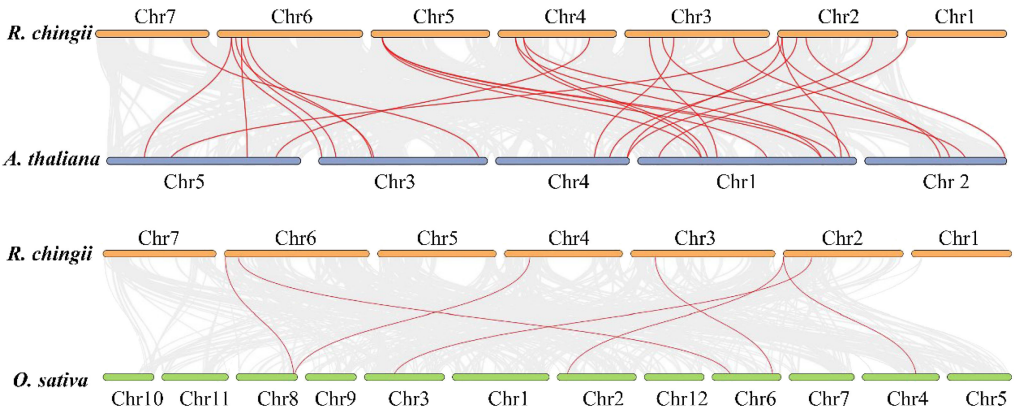


FIGURE 3
Collinearity analysis of *BBX* genes among *A. thaliana*, *O. sativa*, and *R. chingii*.

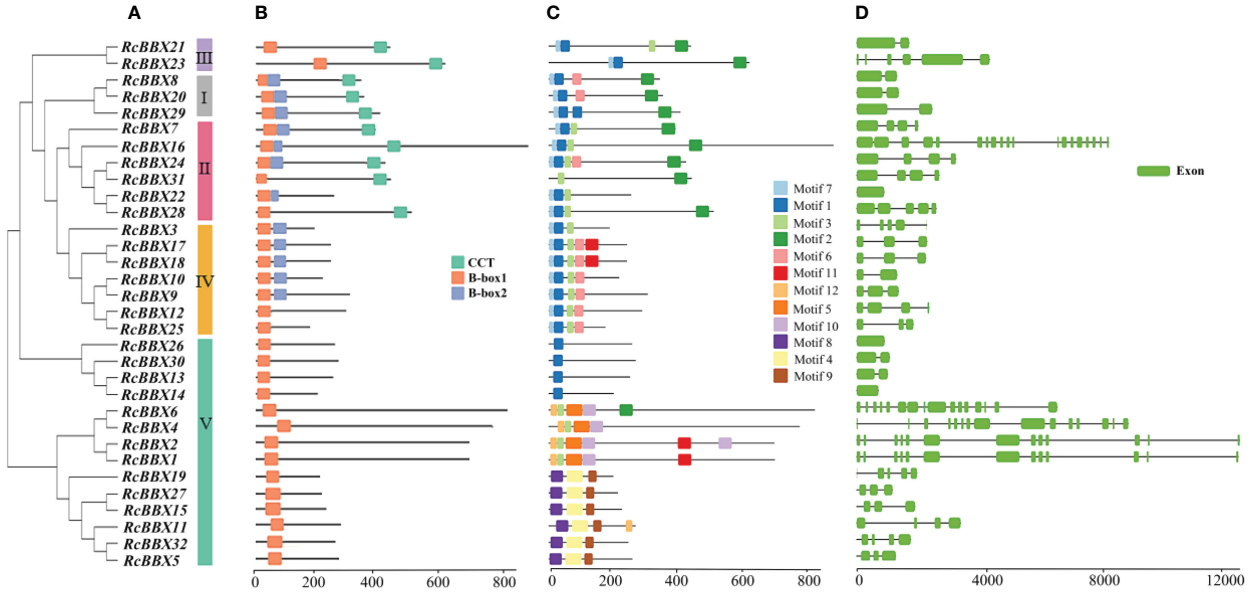


FIGURE 4
Phylogenetic tree, structural domains, conserved motifs, and gene structure of the 32 *RcBBX* members. **(A)** Phylogenetic classification of the 32 RcBBX proteins. **(B)** Structural domains of the 32 RcBBX proteins. **(C)** Conserved motifs of the 32 RcBBX proteins. **(D)** Gene structure of the 32 *RcBBX* genes.

(CC), and one biological process (BP). In the MF group, RcBBX proteins were mainly associated with the activity of a transcription regulator or a DNA-binding TF, which can activate or repress the expression levels of downstream structural genes, playing an important regulatory role in plant growth, development and defense against stresses. In the CC group, the RcBBX proteins were enriched in the nucleus, inferring that they might be involved in nuclear transcription. In the BP group, RcBBX proteins showed significant enrichment in response to light and abiotic stress stimuli, indicating they were linked to photo- and bio-stimulation.

KEGG annotation revealed that RcBBX proteins were frequently linked to protein families associated with the processing of genetic information (Figure 7B). Moreover, RcBBX proteins were associated with environmental adaptations, concurrent with the BP group in GO enrichment.

3.8 Tissue-specific expression levels of 32 *RcBBX* genes

To appreciate the expression patterns of *RcBBX* genes in various tissues, expression levels in flowers, fruits, leaves, roots, and stems were determined (Figure 8). *RcBBX* genes exhibited differential expression in these five tissues, divided into four groups (I, II, III, and IV). There were 12 *RcBBX* genes in group I, with the highest expression in stems. Group II contained five *RcBBX* genes, which displayed the highest expression in leaves. Group III also consisted of five *RcBBX* genes, with the highest expression in roots. The remaining 10 genes were clustered into group IV, and most of them (9/10) displayed highest expression in fruits, except for *RcBBX9*, which was highly expressed in flowers. These results suggest that *RcBBX* genes might play different critical roles during the growth and development of *R. chingii*.

3.9 Expression levels of *RcBBX* genes following exposure to ABA

Since ABRE, an ABA-responsive CAE, existed widely in *RcBBX* genes (Figure 5), the response of *RcBBX* genes to exogenously applied 100 μM ABA was evaluated (Figure 9). A total of 23 *RcBBX* genes were upregulated, five genes were downregulated, while four genes showed no significant change in expression. Among them, *RcBBX26* expression increased the most (39.82-fold), followed by *RcBBX10* and *RcBBX28* expression, suggesting that ABA might be one of the most influential factors affecting the growth and development of *R. chingii* and its ability to accumulate secondary metabolites.

3.10 Expression levels of 32 *RcBBX* genes during the accumulation of anthocyanin

R. chingii fruits pass through four developmental stages: small green fruits (SG), big green fruits (BG), yellow fruits (YE), and red fruits (RE). As shown in Figure 10, the differential anthocyanin accumulation and expression levels of *RcBBX* genes was observed among SG, BG, YE, and RE, and was divided into three groups (I, II, and III). Most *RcBBX* genes (23/32) were highly expressed in SG or BG, and three genes were highly expressed in BG. The *RcBBX* genes displayed different expression levels at all four stages of fruit development. Two genes (*RcBBX13* and *RcBBX31*) had the highest

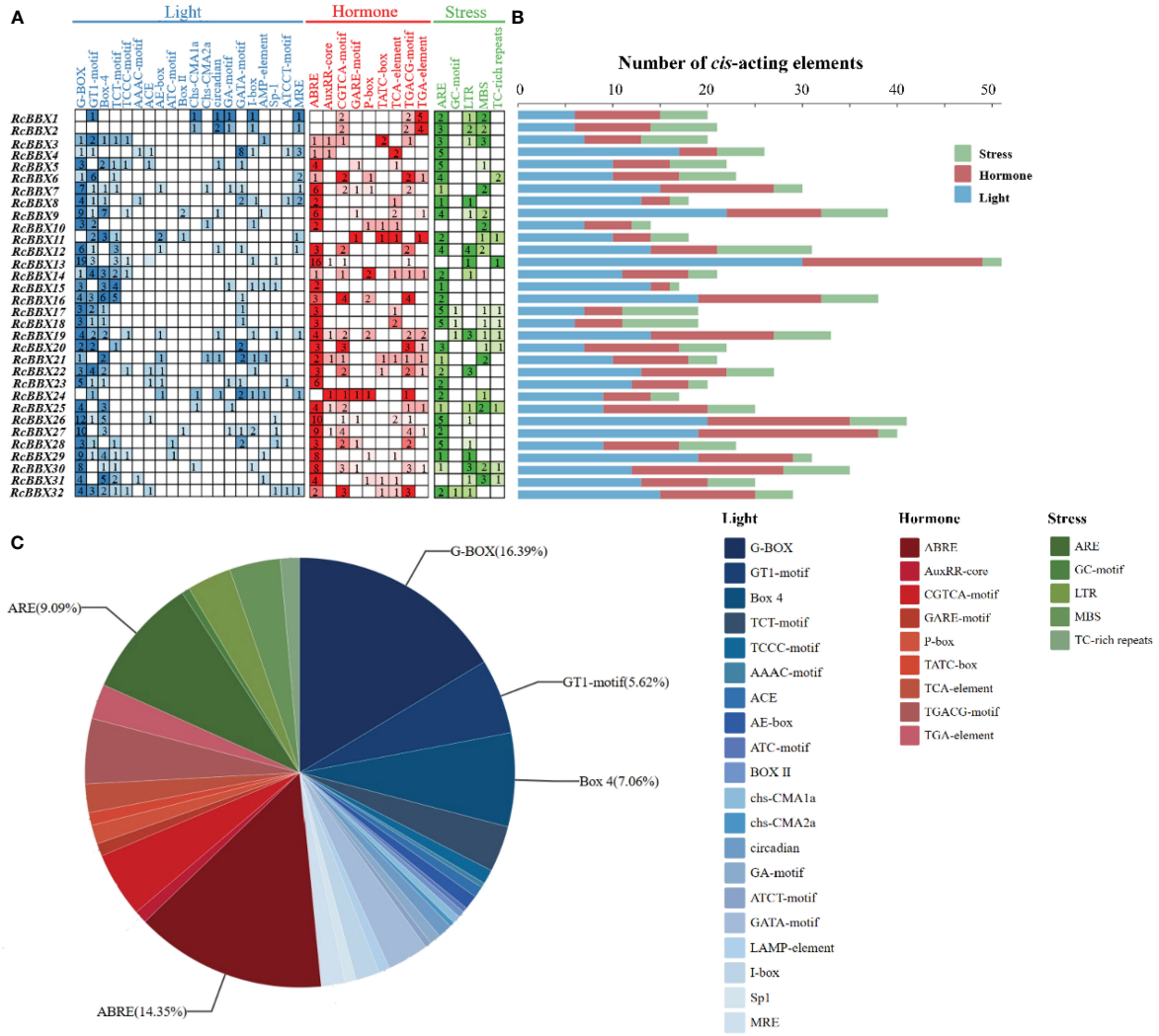


FIGURE 5
Analysis of *cis*-acting elements in the *RcBBX* genes. **(A)** All *cis*-acting elements were classified into three categories, namely those associated with light, hormone, and stress, and their number is presented as a heatmap. **(B)** Number of *cis*-acting elements in each *RcBBX* gene. **(C)** Histogram of the number of *cis*-elements in each category.

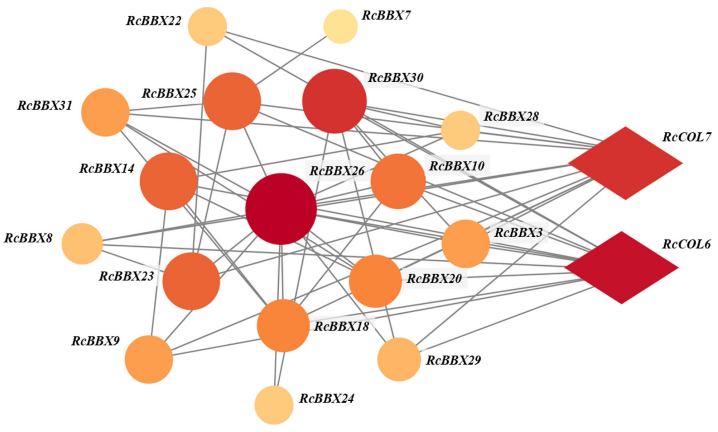


FIGURE 6
Interactive networks of RcBBX proteins assessed with STRING software.

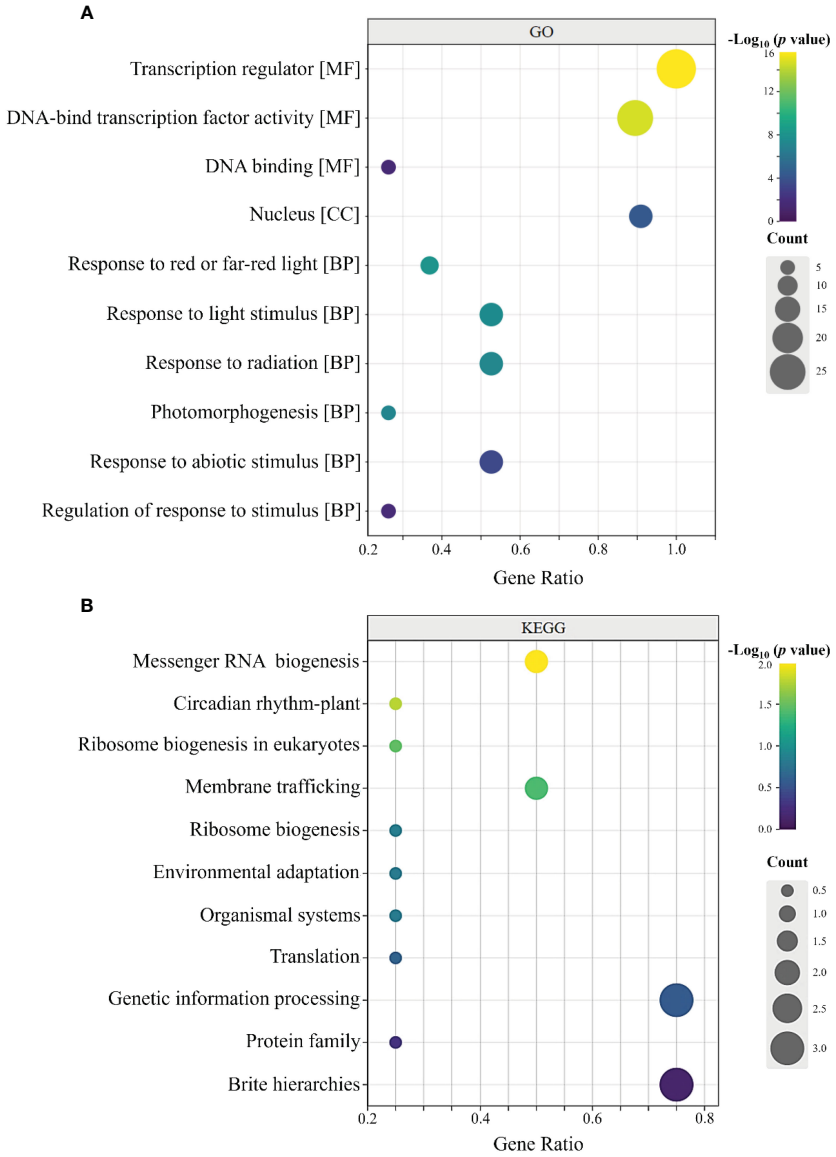


FIGURE 7
GO and KEGG annotation of *RcBBX* proteins. **(A)** GO enrichment analysis of 32 *RcBBX* proteins. All annotated GO terms including biological processes (BP), cellular components (CC) and molecular functions (MF). **(B)** KEGG enrichment analysis of 32 *RcBBX* proteins.

expression in RE. Highest expression of seven genes (*RcBBX7*, *RcBBX9*, *RcBBX14*, *RcBBX15*, *RcBBX25*, *RcBBX26*, *RcBBX27*) was observed in YE. The pH differential approach was used to determine total anthocyanin content in SG, BG, YE, and RE, with YE having the highest content, 6.83 μg g^{-1}, 1.68-fold more than SG (4.05 μg g^{-1}), and 2.42-fold more than RE (2.82 μg g^{-1}) (Figure 10B). This result was in agreement with the results of another report, which indicated that total anthocyanin content first increased, then decreased during the ripening of fruits (Lei et al., 2023).

Furthermore, putative anthocyanin biosynthetic enzymes, including phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate CoA ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3′-hydroxylase (F3′H), flavonoid 3′5′-hydroxylase (F3′5′H), flavonol synthase (FLS), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), and UDP-glucose flavonoid 3-*O*-glucosyl transferase (UFGT), were identified in *R. chingii* (Figure 11; Supplementary Table 5) based on the reported genome or transcriptome data (Li et al., 2021; Wang et al., 2021). The expression of 44 enzyme-coding genes at all four stages (SG, BG, YE, and RE) was detected (Figure 11). Among them, 14 genes (*Rc4CL4*, *Rc4CL5*, *Rc4CL6*, *Rc4CL7*, *Rc4CL8*, *Rc4CL12*, *RcCHI1*, *RcF3H1*, *RcFLS1*, *RcF3'H3*, *RcANS*, *RcUFGT8*, *RcUFGT9*, and *RcUFGT11*) were highly expressed in YE, which was positively related with the seven *RcBBX* genes that were also highly expressed in YE. Combined with the tissue-specific expression of *RcBBX* genes in fruits (Figure 8), *RcBBX26* was potentially related to anthocyanin accumulation in *R. chingii* fruits.

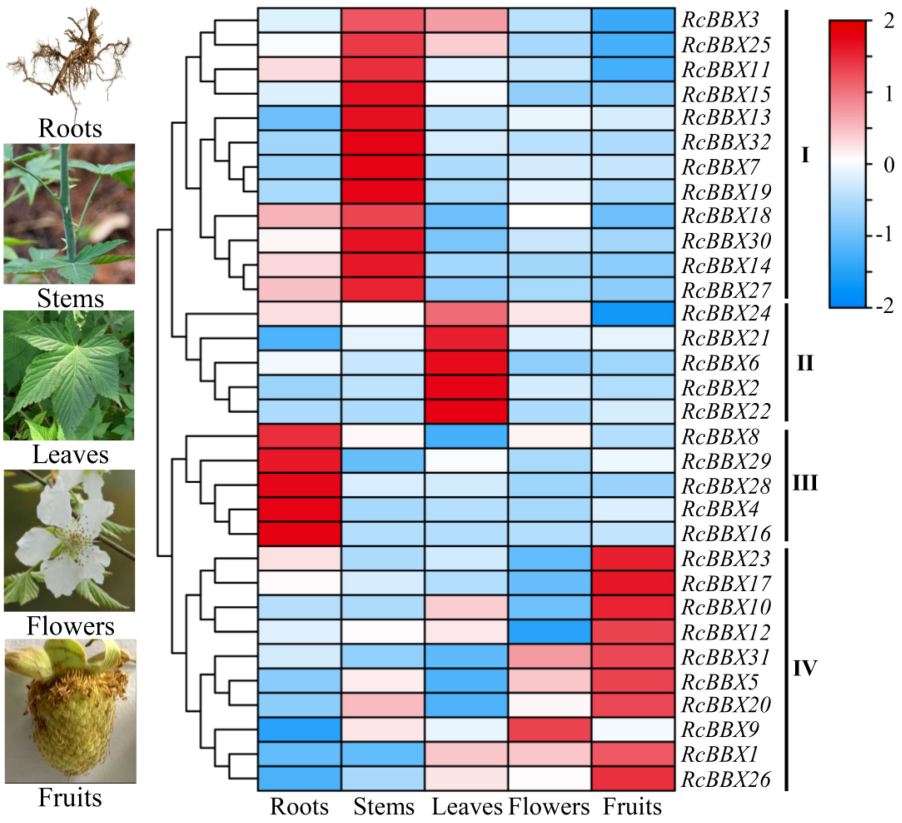


FIGURE 8
Expression profiles of 32 *RcBBX* genes in flowers, fruits, leaves, roots, and stems of *R. chingii*.

3.11 Co-expression network of *RcBBX* genes with anthocyanin biosynthetic genes

Expression levels of *RcBBX* genes and multiple anthocyanin biosynthetic genes at four developmental stages were matched to construct co-expression networks (Figure 12A). A consistent trend was observed between *RcBBX26* and seven anthocyanin biosynthetic genes (Pearson's $r > 0.7$, $P < 0.05$), *Rc4CL4*, *Rc4CL5*, *Rc4CL6*, *Rc4CL12*, *RcUFGT8*, *RcUFGT9*, and *RcUFGT11* (Figure 12B), with a correlation coefficient of 0.55 to 0.92 (Supplementary Table-6), indicating that *Rc4CL* and *RcUFGT* might be the potential target genes of *RcBBX26*.

3.12 Cloning, subcellular localization and functional analysis of RcBBX26

RcBBX26 was amplified, sequenced and submitted to NCBI under accession no. PP723082. RcBBX26 harbored a coding sequence of 765 bp, encoding 254 amino acids with a molecular weight of 27.40 kDa (Figure 13A). The two-dimensional structure showed that RcBBX26 consisted of 50.00% random coils, 43.31% alpha helices, 5.12% extended strands and 1.57% beta turns (Figure 13B). Besides, RcBBX26 was a nuclear-localized protein (Figure 13C), supporting the conclusion that RcBBX26 functions as a TF.

R. chingii is a medicinal xylophyte whose genetic transformation is difficult to achieve (Wang et al., 2021). Herein, efficient transient overexpression was utilized to confirm the function of RcBBX26, resulting in positive transgenic lines through semi-qRT-PCR (Supplementary Figure 1) and qRT-PCR (Figures 13D, E). Overexpression of RcBBX26 accelerated anthocyanin production in *R. chingii* leaves, representing a 3.57-fold increase compared with the untreated MOCK (Figure 13F). Besides, nine anthocyanin biosynthetic genes (*Rc4CL4*, *Rc4CL5*, *Rc4CL6*, *Rc4CL7*, *Rc4CL12*, *RcF3'H3*, *RcANS*, *RcUFGT8*, *RcUFGT9*, and *RcUFGT11*) were upregulated by 1.43- to 13.80-fold, with *RcUFGT11* showing the largest increase (Supplementary Figure 2).

4 Discussion

The BBX family, which belongs to the zinc finger TF superfamily, is considered to play an important role in plant growth and development. The BBX family has not only been widely studied in model plants *A. thaliana*, *O. sativa*, and *S. lycopersicum*, but also in economic plants *Fagopyrum tataricum* (Zhao et al., 2021), *Lycium barbarum* (Yin et al., 2022), *Phyllostachys edulis* (Ma et al., 2021), and *Castanea mollissima* (Yu et al., 2024). However, no information is available on the BBX family in *R. chingii*, which is an edible and medicinal dual-purpose herb.

10.3389/fpls.2024.1427359

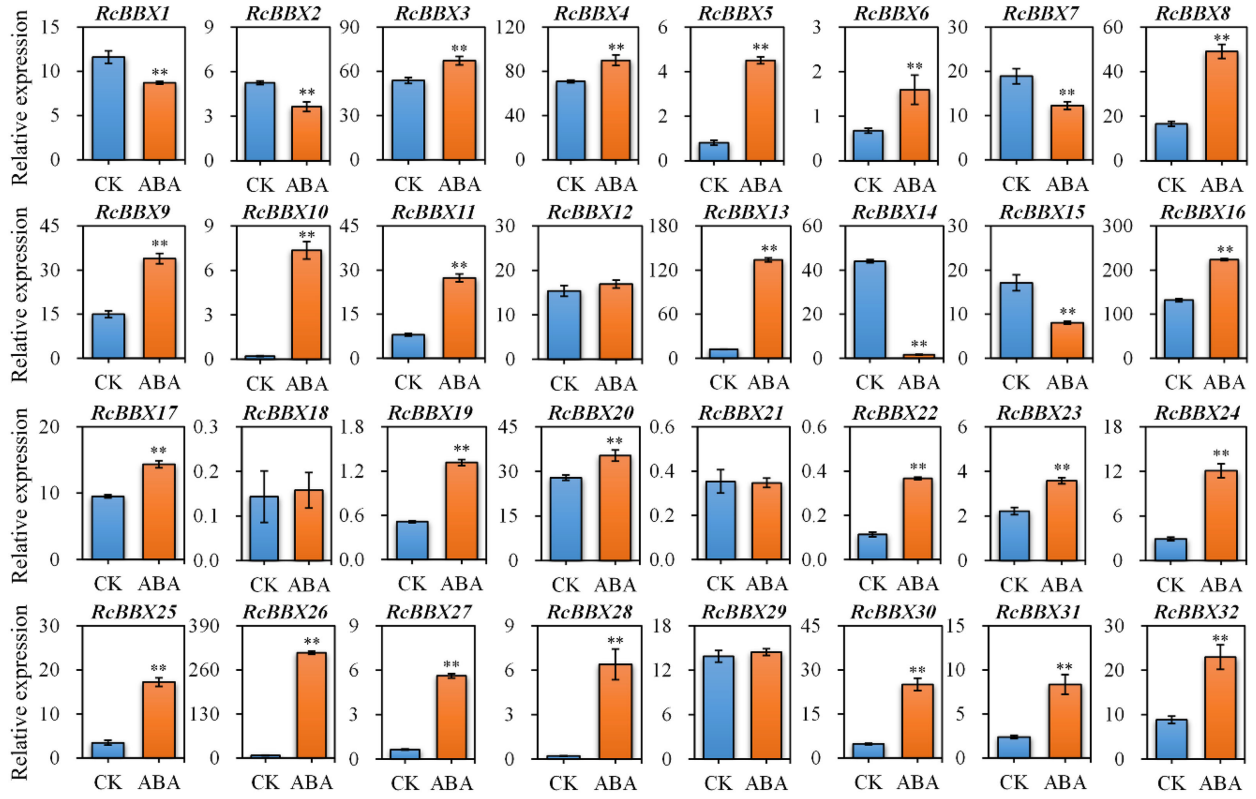


FIGURE 9
Expression profiles of 32 *RcBBX* genes after treatment with ABA in *R. chingii*. Each bar refers to the mean ± SD (standard deviation) of three independent replicates. Double asterisks indicate the significance between CK and ABA based on the student's *t*-test at $P < 0.01$. CK, untreated group. ABA, treatment of 100 μM abscisic acid.

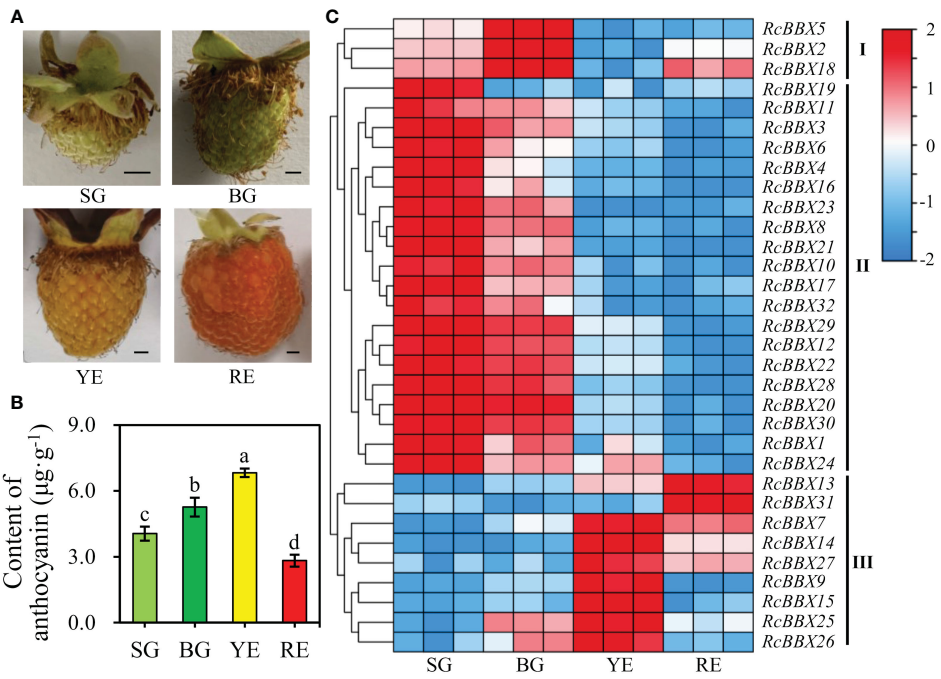


FIGURE 10
Expression profiles of 32 *RcBBX* genes in *R. chingii* during fruit ripening. **(A)** Phenotype of *R. chingii* fruits. **(B)** Anthocyanin content at four different stages of *R. chingii* fruits. Different lowercase letters indicate significant differences among different fruit ripening stages based on Duncan's multiple range test at $P < 0.01$. **(C)** Expression of *RcBBX* genes at four stages of *R. chingii* fruits: SG, small green fruits (5-6 mm in diameter); BG, big green fruits (11-13 mm in diameter); YE, yellow fruits (14-15 mm in diameter); RE, red fruits (18-20 mm in diameter). Bar = 2 mm.

10.3389/fpls.2024.1427359

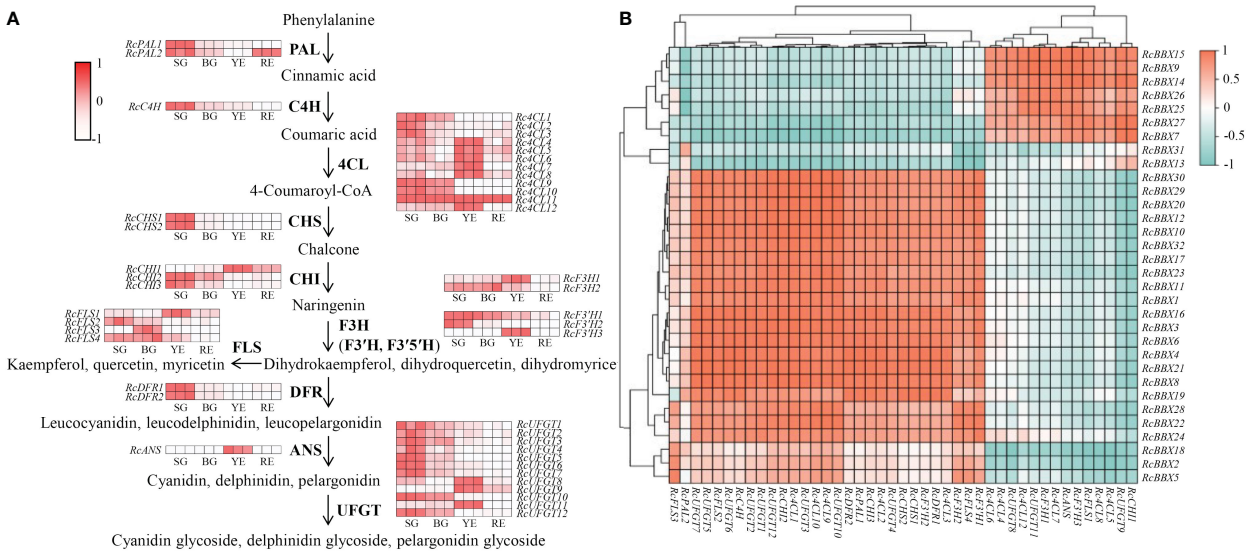


FIGURE 11
Expression profiles of enzyme-encoding genes involved in anthocyanidin biosynthesis during the fruit ripening stage, and their correlation with 32 *RcBBX* genes. **(A)** Expression profiles of enzyme-encoding genes associated with anthocyanidin biosynthesis during different fruit ripening stages (SG, small green fruits; BG, big green fruits; YE, yellow fruits; and RE, red fruits). PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid 3′5′-hydroxylase; FLS, flavonol synthase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; UFGT, UDP-glucose flavonoid 3-*O*-glucosyl transferase. **(B)** Correlation analysis of the expression between *RcBBX* genes and anthocyanidin biosynthetic genes.

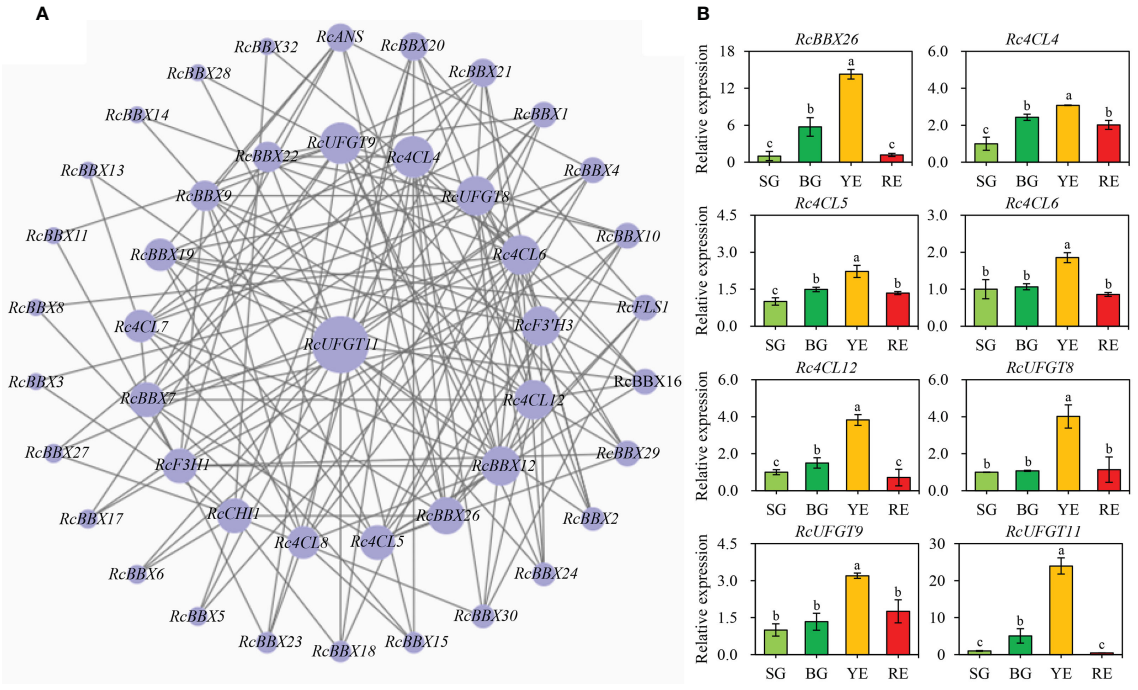


FIGURE 12
Co-expression network between *RcBBX* genes with anthocyanin biosynthetic genes. **(A)** Co-expression map using Cytoscape. **(B)** qRT-PCR analysis of RcBBX26 and its potential target genes (Pearson correlation coefficient $r > 0.7$, $P < 0.05$) during fruit ripening. The four stages of *R. chingii* fruits are: SG, small green fruits; BG, big green fruits; YE, yellow fruits; RE, red fruits. Different lowercase letters indicate significant differences during fruit ripening based on Duncan's multiple range test at $P < 0.01$.

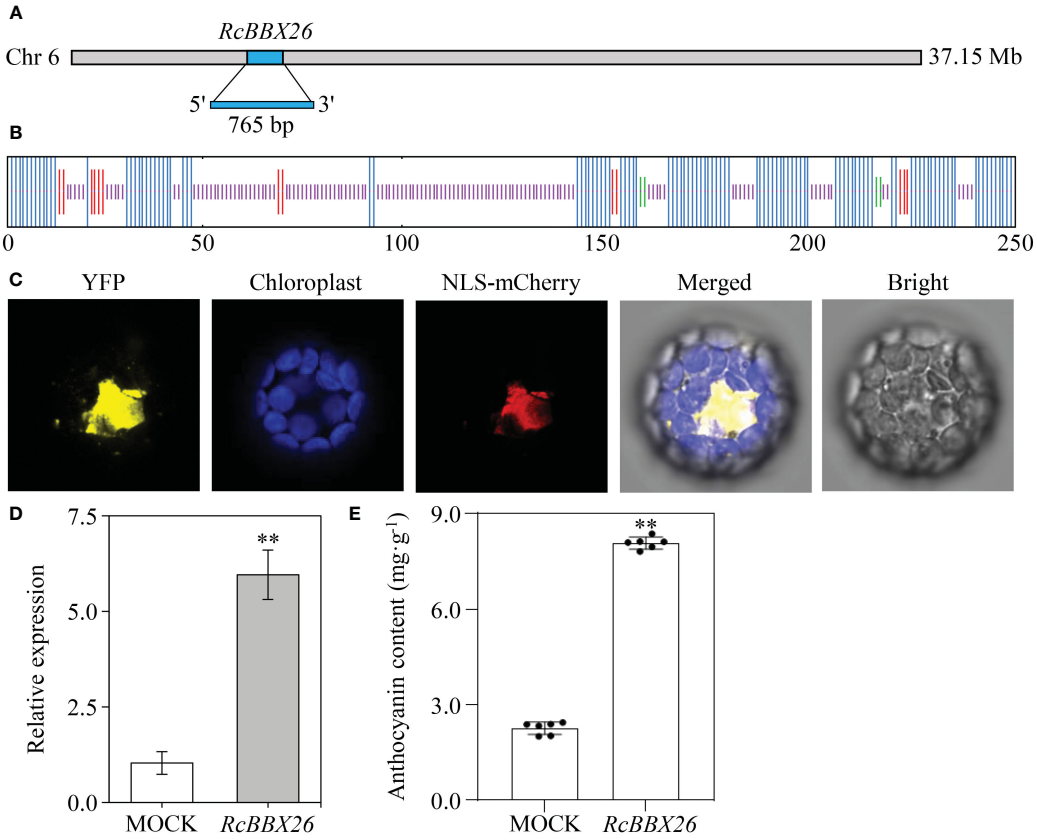


FIGURE 13
Functional analysis of *RcBBX26* in *R. chingii* leaves. **(A)** *RcBBX26* in the *R. chingii* chromosome. **(B)** Secondary structure of RcBBX26. Alpha helixes, extended strands, beta turns, and random coils are represented in blue, red, cyan, and purple, respectively. **(C)** Subcellular localization of RcBBX26. **(D)** qRT-PCR analysis of *RcBBX26* in transgenic *R. chingii* leaves. **(E)** Content of anthocyanin in transgenic leaves. Data denote the mean ± SD (standard deviation) of six independent replicates. Double asterisks indicate significance between MOCK and *RcBBX26* based on a student's *t*-test at $P < 0.01$.

In this study, a total of 32 genes were identified and named *RcBBX1-32* on the basis of their chromosomal positions (Table 1; Supplementary Table 2). This number is the same as the amount of *BBX* genes in *A. thaliana* (32) (Khanna et al., 2009) and *S. miltiorrhiza* (32) (Li et al., 2023), but more than in *C. mollissima* (18) (Yu et al., 2024), *Ananas comosus* (19) (Ouyang et al., 2022), *F. tataricum* (28) (Zhao et al., 2021), *L. barbarum* (29) (Yin et al., 2022) and *O. sativa* (30) (Huang et al., 2012), suggesting the diversification of BBX members among different plant species.

Three conserved domains (B-box1, B-box2 and CCT) of the 32 RcBBX proteins displayed high similarity, inferring that the RcBBX sequences have been strongly conserved throughout evolution. A phylogenetic tree categorized the 32 *RcBBX* family genes into five groups (Figure 1), and the RcBBX members within the same group harbored a strictly consistent combination of conservative domains, which was in line with the *BBX* genes of *A. thaliana* (Khanna et al., 2009) and *O. sativa* (Huang et al., 2012). The number of different groups varied among species. For instance, 13, 4, 8 and 7 A. *thaliana BBX* genes were clustered into groups I/II, III, IV and V, respectively, compared to 9, 2, 7, and 14 in *R. chingii* (Supplementary Table 4), and 10, 3, 7, and 16 in *S. miltiorrhiza* (Li et al., 2023). These findings indicate that the BBX family might share a common ancestor among different plant species, although their evolutionary changes occurred independently. Furthermore, the number of group V *RcBBX* genes, only harboring a single B-box 1, was significantly higher than in any of the other groups. The loss of B-box 2 and CCT domains in group V may be due to a highly frequent gene expansion (Gangappa and Botto, 2014), while critically functional genes might also be hidden in this group.

Six duplication sets covering 32 *RcBBX* genes were phylogenetically distributed in a discrete group with similar domains (Figure 4B), motif compositions (Figure 4C) and gene structure (Figure 4D), suggesting that each pair of duplicated genes probably underwent the closest evolutionary processes and shared similar roles in *R. chingii*. Furthermore, *RcBBX* genes exhibited 4-fold more homologous gene pairs with *A. thaliana* than with *O. sativa* (Figure 3), possibly indicating that different gene duplication events occurred during the evolution of monocotyledonous and dicotyledonous plants. The differences between groups may be associated with widespread diversity of the BBX family (Shan et al., 2022).

Different CAEs present in the promoter region play an essential role in functional diversity (Wittkopp and Kalay, 2011). A total of 836 CAEs were identified, and 48.44% of them were involved in light responsiveness (Figure 5). Notably,

the G-box, which was mainly distributed in the promoter region of *RcBBX* genes (28/32), is an important binding site for BBX regulators, including HY5, PIF3, and PIF8 (Job and Datta, 2021). In *A. thaliana*, HY5 binds to the G-box at the promoter region of BBX11, positively activates its expression, and ultimately affects light-mediated growth and development (Zhao et al., 2020). Through GO and KEGG enrichment annotation, most *RcBBX* members were enriched in response to light or abiotic stress stimuli (Figure 7). ABRE, a well-studied CAEs associated with ABA-induced expression, was widely distributed in the promoter region of *RcBBX* genes (28/32), indicating that they might be of great importance to cope with various environmental stresses. It is widely known that ABA can promote the product of specialized metabolites in multiple medicinal plants, such as tanshinone and salvianolic acid in *S. miltiorrhiza*, artemisinin in *Artemisia annua*, and ginsenoside in *Panax ginseng* (Zheng et al., 2023). Herein, most *RcBBX* genes were up or down regulated (87.50%) after treatment with ABA (Figure 9), suggesting that they might be involved in ABA-induced anthocyanin biosynthesis.

R. chingii is employed as both food and medicine, and the fruit is the primary tissue. Tissue-specific expression analyses showed that nine *RcBBX* genes (*RcBBX1*, *RcBBX5*, *RcBBX10*, *RcBBX12*, *RcBBX17*, *RcBBX20*, *RcBBX23*, *RcBBX26*, and *RcBBX31*) were dominantly expressed in fruit (Figure 8), suggesting that they might be related to the accumulation of anthocyanin. Total anthocyanin content gradually increased as ripening proceeded, attained a maximum at the YE stage, then decreased (Figure 10B). Besides, seven *RcBBX* genes (*RcBBX7*, *RcBBX9*, *RcBBX14*, *RcBBX15*, *RcBBX25*, *RcBBX26*, and *RcBBX27*) initially increased, then decreased (Figure 10C), a trend that was consistent with the dynamic accumulation of anthocyanin. *RcBBX26*, which showed fruit-specific expression, is a candidate gene to explain anthocyanin accumulation in *R. chingii*. Overexpression of *PpBBX16* in *P. pyrifolia* callus promoted red coloration, and resulted in the activated expression of key TF PpMYB10 and structural genes, for instance *PpCHS*, *PpCHI*, *PpUFGT*, and *PpDFR* (Bai et al., 2019a). Additionally, PpBBX16 interacted with PpHY5, thereby stimulating the expression of *PpMYB10*, while PpBBX21 interacted with PpHY5, but suppressed anthocyanin biosynthesis (Bai et al., 2019b). During fruit ripening, seven enzyme-coding genes (*Rc4CL4*, *Rc4CL5*, *Rc4CL6*, *Rc4CL12*, *RcUFGT8*, *RcUFGT9*, and *RcUFGT11*) related to anthocyanin biosynthesis (Figure 12) also exhibited almost the same trend as *RcBBX26* expression and anthocyanin accumulation. The nuclear-located RcBBX26 was conductive to anthocyanin production in transgenic *R. chingii* leaves (Figure 13). These results suggest that RcBBX26 is a positive regulator that potentially activates the expression of anthocyanin biosynthetic genes (Supplementary Table-2) and thus promotes the accumulation of anthocyanin in *R. chingii* fruits.

5 Conclusion

In this study, a total of 32 *BBX* genes were identified from the high-quality genome of *R. chingii*. The complete series of *RcBBX* genes was analyzed, including a phylogenetic analysis, an assessment of their structure and motifs, prediction of chromosome location, and analysis of CAEs in the gene promoter region. Expression profiles of the 32 *RcBBX* genes in different tissues, at different developmental stages, and following treatment with ABA were diverse. A combination of the co-expression of *RcBBX* genes and functional overexpression unveiled the role of *RcBBX26*, which was closely involved in anthocyanin biosynthesis in *R. chingii* fruits. This study lays a foundation for further studies of these *RcBBX* genes and contributes to the ability of breeders to genetically improve the quality of *R. chingii* varieties.

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

ZX: Data curation, Formal analysis, Investigation, Writing – original draft. GZ: Data curation, Formal analysis, Funding acquisition, Writing – original draft, Writing – review & editing. JC: Data curation, Formal analysis, Investigation, Writing – original draft. YY: Data curation, Formal analysis, Investigation, Writing – original draft. LY: Data curation, Formal analysis, Investigation, Writing – original draft. XL: Data curation, Formal analysis, Investigation, Writing – original draft. JT: Investigation, Supervision, Writing – original draft, Writing – review & editing. ZY: Conceptualization, Funding acquisition, Project administration, Supervision, Visualization, Writing – original draft, 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 jointly supported by Key Scientific and Technological Grant of Zhejiang for Breeding New Agricultural Varieties (2021C02074-1), National Natural Science Foundation of China (32000257), Zhejiang Provincial Natural Science Foundation of China (LY23H280003), Zhejiang Provincial Scientific Research Institute Special Project (2024F1068-2), Guangzhou Basic and Applied Basic Research Program (202201010122), and Research Project of Zhejiang Chinese Medical University (2022JKZKTS15).

Acknowledgments

We appreciate the considerable experimental support from the Public Platform of Pharmaceutical Research Center, Academy of Chinese Medical Science, Zhejiang Chinese Medical University.

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

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1427359/full#supplementary-material

References

Bai, S., Tao, R., Tang, Y., Yin, L., Ma, Y., Ni, J., et al. (2019a). BBX16, a B-box protein, positively regulates light-induced anthocyanin accumulation by activating MYB10 in red pear. *Plant Biotechnol. J.* 17, 1985–1997. doi: 10.1111/pbi.13114

Bai, S., Tao, R., Yin, L., Ni, J., Yang, Q., Yan, X., et al. (2019b). Two B-box proteins, PpBBX18 and PpBBX21, antagonistically regulate anthocyanin biosynthesis via competitive association with *Pyrus pyrifolia* ELONGATED HYPOCOTYL 5 in the peel of pear fruit. *Plant J.* 100, 1208–1223. doi: 10.1111/tpj.14510

Cao, J., Yuan, J., Zhang, Y., Chen, C., Zhang, B., Shi, X., et al. (2023). Multi-layered roles of BBX proteins in plant growth and development. *Stress Biol.* 3, 1. doi: 10.1007/s44154-022-00080-z

Cao, Y., Meng, D., Han, Y., Chen, T., Jiao, C., Chen, Y., et al. (2019). Comparative analysis of B-box genes and their expression pattern analysis under various treatments in *Dendrobium officinale*. *BMC Plant Biol.* 19, 245. doi: 10.1186/s12870-019-1851-6

Chang, Y., Sun, H., Liu, S., He, Y., Zhao, S., Wang, J., et al. (2023). Identification of BBX gene family and its function in the regulation of microtuber formation in yam. *BMC Genomics* 24, 354. doi: 10.1186/s12864-023-09406-1

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, J. S. (2023). Essential role of medicine and food homology in health and wellness. *Chin. Herb. Med.* 15, 347–348. doi: 10.1016/j.chmed.2023.05.001

Chou, K. C., and Shen, H. B. (2010). Plant-mPLoc: A top-down strategy to augment the power for predicting plant protein subcellular localization. *PloS One* 5, e11335. doi: 10.1371/journal.pone.0011335

Fang, H., Dong, Y., Yue, X., Hu, J., Jiang, S., Xu, H., et al. (2019). The B-box zinc finger protein MdBBX20 integrates anthocyanin accumulation in response to ultraviolet radiation and low temperature. *Plant Cell Environ.* 42, 2090–2104. doi: 10.1111/pce.13552

Gangappa, S. N., and Botto, J. F. (2014). The BBX family of plant transcription factors. *Trends Plant Sci.* 19, 460–470. doi: 10.1016/j.tplants.2014.01.010

Huang, J., Zhao, X., Weng, X., Wang, L., and Xie, W. (2012). The rice B-box zinc finger gene family: genomic identification, characterization, expression profiling and diurnal analysis. *PloS One* 7, e48242. doi: 10.1371/journal.pone.0048242

Job, N., and Datta, S. (2021). PIF3/HY5 module regulates BBX11 to suppress protochlorophyllide levels in dark and promote photomorphogenesis in light. *New Phytol.* 230, 190–204. doi: 10.1111/nph.17149

Job, N., Yadukrishnan, P., Bursch, K., Datta, S., and Johansson, H. (2018). Two B-box proteins regulate photomorphogenesis by oppositely modulating HY5 through their diverse C-terminal domains. *Plant Physiol.* 176, 2963–2976. doi: 10.1104/pp.17.00856

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 51, 587–592. doi: 10.1093/nar/gkac963

Khanna, R., Kronmiller, B., Maszle, D. R., Coupland, G., Holm, M., Mizuno, T., et al. (2009). The *Arabidopsis* B-box zinc finger family. *Plant Cell.* 21, 3416–3420. doi: 10.1105/tpc.109.069088

Laity, J. H., Lee, B. M., and Wright, P. E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* 11, 39–46. doi: 10.1016/S0959-440X(00)00167-6

Lei, T., Huang, J., Ruan, H., Qian, W., Fang, Z., Gu, C., et al. (2023). Competition between FLS and DFR regulates the distribution of flavonols and proanthocyanidins in *Rubus chingii* Hu. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1134993

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in *silico* analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, X., Jiang, J., Chen, Z., and Jackson, A. (2021). Transcriptomic, proteomic and metabolomic analysis of flavonoid biosynthesis during fruit maturation in *Rubus chingii* Hu. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.706667

Li, Y., Tong, Y., Ye, J., Zhang, C., Li, B., Hu, S., et al. (2023). Genome-wide characterization of B-box gene family in *Salvia miltiorrhiza*. *Int. J. Mol. Sci.* 242146. doi: 10.3390/ijms24032146

Liu, X., Sun, W., Ma, B., Song, Y., Guo, Q., Zhou, L., et al. (2023). Genome-wide analysis of blueberry B-box family genes and identification of members activated by abiotic stress. *BMC Genomics* 24, 584. doi: 10.1186/s12864-023-09704-8

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

Ma, R., Chen, J., Huang, B., Huang, Z., and Zhang, Z. (2021). The BBX gene family in Moso bamboo (*Phyllostachys edulis*): Identification, characterization and expression profiles. *BMC Genomics* 22, 533. doi: 10.1186/s12864-021-07821-w

Nian, L., Zhang, X., Liu, X., Li, X., Liu, X., Yang, Y., et al. (2022). Characterization of B-box family genes and their expression profiles under abiotic stresses in the *Melilotus albus*. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.990929

Noman, A., Aqeel, M., Khalid, N., Islam, W., Sanaullah, T., Anwar, M., et al. (2019). Zinc finger protein transcription factors: Integrated line of action for plant antimicrobial activity. *Microb. Pathog.* 132, 141–149. doi: 10.1016/j.micpath.2019.04.042

Ouyang, Y., Pan, X., Wei, Y., Wang, J., Xu, X., He, Y., et al. (2022). Genome-wide identification and characterization of the BBX gene family in pineapple reveals that candidate genes are involved in floral induction and flowering. *Genomics* 114, 110397. doi: 10.1016/j.ygeno.2022.110397

Podolec, R., Wagnon, T. B., Leonardelli, M., Johansson, H., and Ulm, R. (2022). *Arabidopsis* B-box transcription factors BBX20-22 promote UVR8 photoreceptor-mediated UV-B responses. *Plant J.* 111, 422–439. doi: 10.1111/tpj.15806

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454

Shan, B., Bao, G., Shi, T., Zhai, L., Bian, S., and Li, X. (2022). Genome-wide identification of BBX gene family and their expression patterns under salt stress in soybean. *BMC Genomics* 23, 820. doi: 10.1186/s12864-022-09068-5

Song, J., Lin, R., Tang, M., Wang, L., Fan, P., Xia, X., et al. (2023). SlMPK1- and SlMPK2-mediated SlBBX17 phosphorylation positively regulates CBF-dependent cold tolerance in tomato. *New Phytol.* 239, 1887–1902. doi: 10.1111/nph.19072

Song, Z., Bian, Y., Liu, J., Sun, Y., and Xu, D. (2020). B-box proteins: Pivotal players in light-mediated development in plants. *J. Integr. Plant Biol.* 62, 1293–1309. doi: 10.1111/jipb.12935

Spitz, F., and Furlong, E. E. (2012). Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.* 13, 613626. doi: 10.1038/nrg3207

Stortz, M., Presman, D. M., and Levi, V. (2024). Transcriptional condensates: A blessing or a curse for gene regulation. *Commun. Biol.* 7, 187. doi: 10.1038/s42003-024-05892-5

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, L., Lei, T., Han, G., Yue, J., Zhang, X., Yang, Q., et al. (2021). The chromosome-scale reference genome of *Rubus chingii* Hu provides insight into the biosynthetic pathway of hydrolyzable tannins. *Plant J.* 107, 1466–1477. doi: 10.1111/tpj.15394

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, Y., Xiao, Y., Sun, Y., Zhang, X., Du, B., Turupu, M., et al. (2023). Two B-box proteins, PavBBX6/9, positively regulate light-induced anthocyanin accumulation in sweet cherry. *Plant Physiol.* 192, 2030–2048. doi: 10.1093/plphys/kiad137

Wang, Y., Zhao, Y., Liu, X., Li, J., Zhang, J., and Liu, D. (2022). Chemical constituents and pharmacological activities of medicinal plants from *Rosa* genus. *Chin. Herb. Med.* 14, 187–209. doi: 10.1016/j.chmed.2022.01.005

Wei, H., Wang, P., Chen, J., Li, C., Wang, Y., Yuan, Y., et al. (2020). Genome-wide identification and analysis of B-box gene family in grapevine reveal its potential functions in berry development. *BMC Plant Biol.* 20, 72. doi: 10.1186/s12870-020-2239-3

Wittkopp, P. J., and Kalay, G. (2011). *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13, 59–69. doi: 10.1038/nrg3095

Yin, Y., Shi, H., Mi, J., Qin, X., Zhao, J., Zhang, D., et al. (2022). Genome-wide identification and analysis of the BBX gene family and its role in carotenoid biosynthesis in wolfberry (*Lycium barbarum* L.). *Int. J. Mol. Sci.* 23, 8440. doi: 10.3390/ijms23158440

Yu, L., Wang, D., Huang, R., Guo, C., and Zhang, J. (2024). Genome-wide identification, characterization and expression profile analysis of BBX gene family in Chinese chestnut (*Castanea mollissima*). *Plant Biotechnol. Rep.* 18, 129–142. doi: 10.1007/s11816-023-00845-6

Yu, Z., Liao, Y., Teixeira da Silva, J. A., Yang, Z., and Duan, J. (2018). Differential accumulation of anthocyanins in *Dendrobium officinale* stems with red and green peels. *Int. J. Mol. Sci.* 19, 2857. doi: 10.3390/ijms19102857

Yu, Z., Zhang, G., Teixeira da Silva, J. A., Li, M., Zhao, C., He, C., et al. (2021a). Genome-wide identification and analysis of DNA methyltransferase and demethylase gene families in *Dendrobium officinale* reveal their potential functions in polysaccharide accumulation. *BMC Plant Biol.* 21, 21. doi: 10.1186/s12870-020-02811-8

Yu, Z., Zhang, G., Teixeira da Silva, J. A., Zhao, C., and Duan, J. (2021b). The methyl jasmonate-responsive transcription factor DobHLH4 promotes *DoTPS10*, which is involved in linalool biosynthesis in *Dendrobium officinale* during floral development. *Plant Sci.* 309, 110952. doi: 10.1016/j.plantsci.2021.110952

Zhao, J., Li, H., Huang, J., Shi, T., Meng, Z., Chen, Q., et al. (2021). Genome-wide analysis of BBX gene family in Tartary buckwheat (*Fagopyrum tataricum*). *PeerJ* 9, e11939. doi: 10.7717/peerj.11939

Zhao, X., Heng, Y., Wang, X., Deng, X. W., and Xu, D. (2020). A positive feedback loop of BBX11-BBX21-HY5 promotes photomorphogenic development in *Arabidopsis*. *Plant Commun.* 1, 100045. doi: 10.1016/j.xplc.2020.100045

Zheng, H., Fu, X., Shao, J., Tang, Y., Yu, M., Li, L., et al. (2023). Transcriptional regulatory network of high-value active ingredients in medicinal plants. *Trends Plant Sci.* 28, 429–446. doi: 10.1016/j.tplants.2022.12.007

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 27 August 2024
DOI 10.3389/fpls.2024.1451298

OPEN ACCESS

EDITED BY
Moonhyuk Kwon,
Gyeongsang National University,
Republic of Korea

REVIEWED BY
Vonny Salim,
Louisiana State University in Shreveport,
United States
David Liscombe,
Vineland Research and Innovation
Centre, Canada

*CORRESPONDENCE
Ghislain Deslongchamps
ghislain@unb.ca
Yang Qu
yang.qu@unb.ca

RECEIVED 19 June 2024
ACCEPTED 01 August 2024
PUBLISHED 27 August 2024

CITATION
Farzana M, Richardson MB, Deschênes DAR, Mai Z, Njoku DI, Deslongchamps G and Qu Y (2024) Parallel evolution of methyltransferases leads to vobasine biosynthesis in *Tabernaemontana elegans* and *Catharanthus roseus*.
Front. Plant Sci. 15:1451298.
doi: 10.3389/fpls.2024.1451298

Parallel evolution of methyltransferases leads to vobasine biosynthesis in *Tabernaemontana elegans* and *Catharanthus roseus*

Maisha Farzana, Matthew Bailey Richardson, Daniel André Ramey Deschênes, Zhan Mai, Destiny Ichechi Njoku, Ghislain Deslongchamps* and Yang Qu*

Department of Chemistry, University of New Brunswick, Fredericton, NB, Canada

Monoterpenoid indole alkaloids (MIA) are one of the largest and most complex alkaloid class in nature, boasting many clinically significant drugs such as anticancer vinblastine and antiarrhythmic ajmaline. Many MIAs undergo nitrogen *N*-methylation, altering their reactivity and affinity to the biological targets through a straightforward reaction. Remarkably, all known MIA *N*-methyltransferases (NMT) originate from the neofunctionalization of ancestral γ-tocopherol *C*-methyltransferases (γTMTs), a phenomenon seemingly unique to the Apocynaceae family. In this study, we unveil and characterize a new γTMT-like enzyme from the plant *Tabernaemontana elegans* (toad tree): perivine *N*β-methyltransferase (TePeNMT). TePeNMT and other homologs form a distinct clade in our phylogenetic study, setting them apart from other γTMTs and γTMT-like NMTs discovered to date. Enzyme kinetic experiments and enzyme homology modeling studies reveal the significant differences in enzyme active sites between TePeNMT and CrPeNMT, a previously characterized perivine *N*β-methyltransferase from *Catharanthus roseus* (Madagascar periwinkle). Collectively, our findings suggest that parallel evolution of ancestral γTMTs may be responsible for the occurrence of perivine *N*-methylation in *T. elegans* and *C. roseus*.

KEYWORDS

monoterpenoid indole alkaloid, methyltransferase, *Tabernaemontana elegans*, natural product biosynthesis, parallel evolution, vobasine

Introduction

Natural product methylation plays a key role in diversifying structures, altering compound polarity, membrane permeability and stability, and modifying their electronic properties for target interaction (Morris and Facchini, 2019). A prominent example lies in the frequent *O*-methylation of phenylpropanoids and flavonoids within plants, integral to cell wall biosynthesis and specialized metabolism, with profound implications for human health (Lam et al., 2007). Similarly, nitrogen *N*-methylation, alongside *O*-methylation, significantly contribute to the diversity of plant alkaloids, such as benzylisoquinoline alkaloids (BIA) and monoterpenoid indole alkaloids (MIA). Notable instances include (*S*)-coclaurine *N*-methylation and benzylisoquinoline *O*-methylations in the biosynthesis of morphine (BIA), as well as tabersonine 16-*O*-methylation and 16-methoxy-2,3-dihydrotabersonine *N*-methylation in vinblastine (MIA) production (Levac et al., 2008; Liscombe et al., 2010). Promoting van der Waals interactions and alterations in molecular electronic properties, these methylations are crucial for downstream biosynthetic enzyme recognition, underscoring their indispensability in morphine and vinblastine biosynthesis. The acquisition and retention of these methylations likely conferred adaptive advantages during the evolutionary processes of host plants opium poppy and *Catharanthus roseus* (Madagascar's periwinkle).

Phylogenetic investigations propose that all identified MIA *N*-methyltransferases (NMT) likely originated from the neofunctionalization of ancestral γ-tocopherol (vitamin E) *C*-methyltransferases (γTMT) (Liscombe et al., 2010) (Figure 1). The first characterized MIA NMT is the 16-methoxy-2,3-dihydrotabersonine *N*-methyltransferase (CrDhtNMT) from *C.*

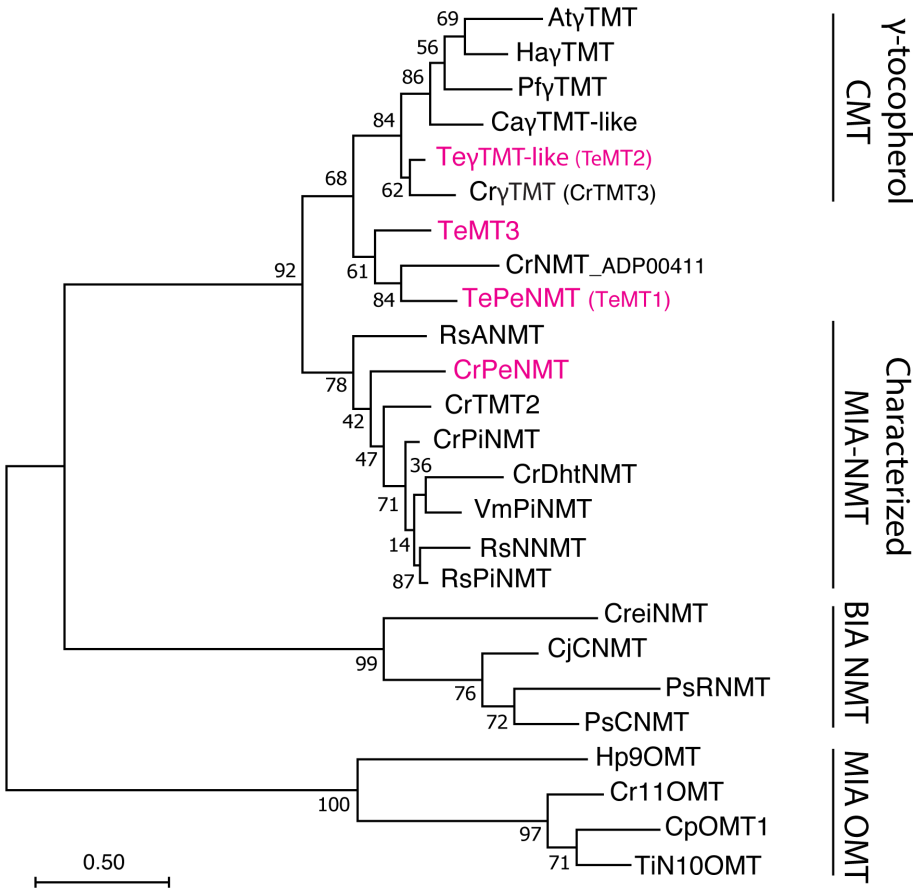


FIGURE 1
Phylogeny of γ-tocopherol *C*-methyltransferase (CMT)-like *N*-methyltransferases (NMT) in Apocynaceae family. The *Tabernaemontana elegans* perivine *N*β-NMT (TePeNMT, TeMT1) and homologs from *Catharanthus roseus* and *Vinca minor* form a new clade, distinguishing them from the *bona fide* γ-TMTs and other characterized monoterpenoid indole alkaloid (MIA) NMTs to date. The enzymes labeled in red were investigated in this study. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches (500 bootstrap replicates). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (scale bar). Evolutionary analyses were conducted in MEGA11. OMT, *O*-methyltransferase; BIA, benzylisoquinoline alkaloids; PiNMT, picrinine NMT; DhtNMT, 16-methoxyl-2,3-dihydrotabersonine NMT; NNMT, norajmaline NMT; ANMT; ajmaline *N*β-NMT; CNMT, coclaurine NMT; RNMT, reticuline NMT; At, *Arabidopsis thaliana*; Ca, *Coffea arabica*; Cp, *Cinchona pubescens*; Cr, Catharanthus roseus; Cre, Chlamydomonas reinhardtii; Cj, *Coptis japonica*; Ha, *Helianthus annuus*; Hp, Hamelia patens; Pf, Perilla frutescens; Ps, *Papaver somniferum*; Rs, *Rauvolfia serpentina*; Te, *Tabernaemontana elegans*; Ti, *Tabernanthe iboga*; Vm, *Vinca minor*. The alignment for constructing the phylogenetic tree is included in Supplementary Data 1.

10.3389/fpls.2024.1451298

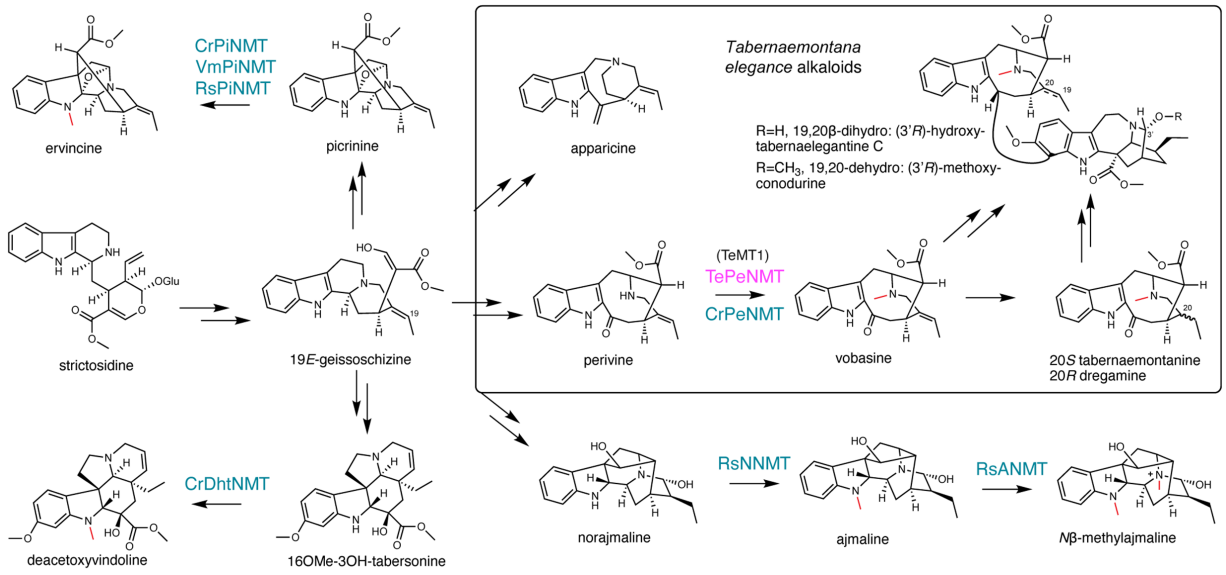


FIGURE 2
γ-tocopherol *C*-methyltransferase-like *N*-methyltransferases (NMT) catalyze diverse reactions, contributing to monoterpenoid indole alkaloid (MIA) diversity. The methyl substitutions on MIAs are labeled in red. The previously characterized NMTs are labeled in blue. The *Tabernaemontana elegans* perivine *N*β-NMT characterized in this study is labeled in magenta.

roseus (Deluca et al., 1987; Luca and Cutler, 1987; Liscombe et al., 2010). Unlike the plastid-located γTMTs, CrDhtNMT lacks a chloroplast transit peptide, indicative of its novel role in cytosolic MIA biosynthesis. While CrDhtNMT does not catalyze γ-tocopherol methylation, it retains the ability to bind γ-tocopherol, as evidenced by its inhibition by γ-tocopherol (Liscombe et al., 2010). Homologous γTMT-like NMTs have also been identified in other MIA-producing species within the Apocynaceae family. RsNNMT, RsANMT, RsPiNMT, VmPiNMT, and CrPeNMT are responsible for the *N*-methylation of norajmaline, ajmaline, picrinine, and perivine respectively, in *Rauwolfia serpentina* (Indian snakeroot), *Vinca minor*, and *C. roseus* (Cázares-Flores et al., 2016; Levac et al., 2016; Koudounas et al., 2022; Levac et al., 2022) (Figure 2). Although previous studies associated CrDhtNMT activity with the chloroplast through sucrose gradient centrifugation (Luca and Cutler, 1987), recent findings utilizing fluorescent protein tagging suggest its localization in peroxisomes, while other MIA NMT members exhibit either cytosolic or vacuolar association (Koudounas et al., 2022).

The *N*β-methylation of perivine (i.e., methylation of the non-indole nitrogen), resulting in the formation of vobasine, is frequently observed in *Tabernaemontana* genus and in other plants in the Apocynaceae family. Vobasine exhibits moderate antifungal and anticancer properties, characteristics shared by its two 19,20-reduced derivatives, tabernaemontanine, and dregamine (Mansoor et al., 2009; Singh et al., 2011; Paterna et al., 2017; Ferreira and Paterna, 2019) (Figure 2). Notably, vobasine and vobasine-type MIAs are often detected as dimerized bisindole alkaloids, either with the same type or a different MIA type, in *Tabernaemontana* spp., demonstrating heightened anticancer activities (Paterna et al., 2016a; Paterna et al., 2016b; Paterna et al., 2017; Ferreira and Paterna, 2019). For example, the iboga-vobasinyl bisindole alkaloid (3'*R*)-hydroxytabernaelegantine C (Figure 2) has been shown to induce potent apoptosis in colon and liver cancer cells (Paterna et al., 2016b). Although *N*β-methylperivine (vobasine) and *N*β-formylperivine (periformyline) have been identified in *Catharanthus* spp (Maloney et al., 1965; Mukhopadhyay and Cordell, 1981)., our research indicated that non-methylated perivine is the primary accumulating form for vobasine type MIAs in *C. roseus* (Eng et al., 2022). In *Tabernaemontana* spp., *N*-methylated vobasine and its derivatives are frequently observed (Beek et al., 1985; Dagnino et al., 1994; Singh et al., 2011).

In this study, we discover and characterize a novel enzyme TePeNMT responsible for perivine *N*β-methylation in the plant *Tabernaemontana elegans* (toad tree). Surprisingly, despite both enzymes catalyzing the same reaction, TePeNMT exhibited a low amino acid identity of 50% when compared to CrPeNMT. Phylogenetic analysis revealed that TePeNMT, along with several other γTMT-like enzymes in the Apocynaceae family, constitutes a distinct subgroup that shares closer evolutionary ties with genuine γTMTs. Homology modeling of CrPeNMT and TePeNMT revealed drastic differences in active site conformation and perivine substrate docking positions. Our findings strongly suggest the possibility of parallel evolution being responsible for perivine *N*-methylation in *C. roseus* and *T. elegans*.

Results

Vobasine and apparicine are the two most abundant MIAs in *T. elegans* leaf

To study the biosynthesis of vobasine and vobasinyl bisindole alkaloids, we investigated the MIA profiles of the leaf, leaf latex,

10.3389/fpls.2024.1451298

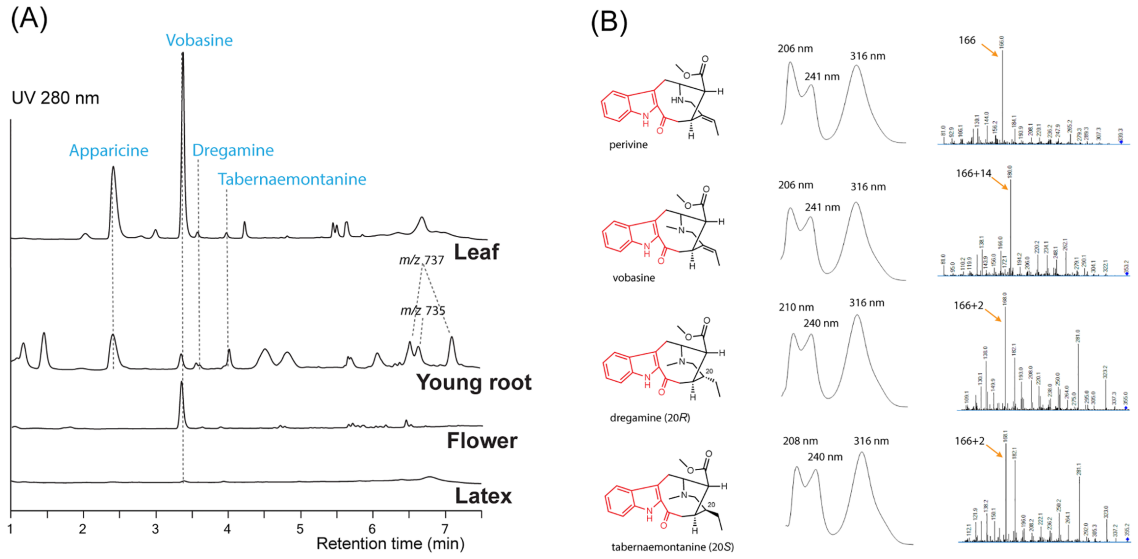


FIGURE 3
Representative LC-MS chromatograms **(A)** for *Tabernaemontana elegans* tissues (leaf, young root, flower, and latex) and alkaloid UV absorption and MS/MS profiles **(B)**. Each chromatogram (UV 280 nm) was normalized by tissue weight and drawn to scale. Apparicine and vobasine were the two most abundant alkaloids in leaf tissues, while peaks tentatively identified as bisindole alkaloids (*m/z* 735 and 737) were detected in young root tissues. The chromophores are labeled in red.

flower, and young root tissues of *T. elegans* growing in our greenhouse using liquid chromatography tandem mass spectrometry (LC-MS/MS). Two major MIAs with $[M+H]^+$ *m/z* 265 and 353 accumulated in leaf and flowers (Figure 3A). In comparison, the MIA profile in *T. elegans* young roots exhibited greater diversity, encompassing the *m/z* 265 and *m/z* 353 MIAs. Notably, leaf latex contained only minimal amounts of MIAs (Figure 3A). This finding suggests that MIAs in *T. elegans* are not sequestered within the leaf latex, in contrast to *C. roseus* and *Rauvolfia tetraphylla*, where laticifers and idioblasts (a special type of parenchyma cells) serve as major sinks for MIAs (Facchini and Luca, 2008; Guedes et al., 2023; Stander et al., 2023; Uzaki et al., 2024).

We purified the *m/z* 265 and *m/z* 353 MIAs from *T. elegans* leaf by thin layer chromatography, and identified them as apparicine and vobasine, respectively, through LC-MS/MS and 1D/2D nuclear magnetic resonance (NMR) analyses. (Supplementary Figures 1–12; Supplementary Tables 1, 2). The identification was further corroborated by comparing the NMR chemical shifts to literature values (Kutney et al., 1980; Pereira et al., 2008; Bennasar et al., 2009). Both MIAs showed UV absorption maxima shifted to longer wavelengths (302 nm for apparicine, and 316 nm for vobasine) compared to the typical indole UV absorption maxima at 280-290 nm (Figure 3B; Supplementary Figure 1), consistent with their extended indole conjugation (Figure 2). Additionally, we purified and identified two diastereomers (*m/z* 355), tabernaemontanine and dregamine (Figures 2, 3; Supplementary Figures 13–22, Supplementary Tables 1, 2), derived from vobasine 19,20-reduction. The disappearance of 19,20-alkene signals in both ^{1}H and ^{13}C NMR spectra was consistent with their 19,20-reduction, and Nuclear Overhauser Effect Spectroscopy (NOESY) cross peaks clearly differentiated the two diastereomers. Similar to vobasine and perivine, both tabernaemontanine and dregamine showed UV absorption profiles and MS/MS fingerprints indicative of the unique indole-C3-ketone chromophore (Supplementary Figure 1). There are several bisindole alkaloids found only in young roots, which we tentatively identified as 3'-methoxytabernaelegantine A and C (*m/z* 737), and 3'-methoxyconodurine (*m/z* 735) based on *m/z* values (Figures 2, 3A). Structural elucidation of these bis-MIAs awaits additional NMR studies.

γTMT-like methyltransferases identified in *T. elegans* comprise a novel clade that is distantly related to characterized MIA NMTs

To identify the enzymes for vobasine biosynthesis, we sequenced *T. elegans* leaf and root total RNA and successfully identified many putative vobasine biosynthetic enzymes using known MIA biosynthetic enzymes (Table 1). For the last step of vobasine biosynthesis, we searched the transcriptomes with six characterized MIA NMT protein sequences. Given the substantial accumulation of vobasine as a major MIA in *T. elegans* leaf, we expected the putative MT to be highly expressed in this tissue. From all hits, we selected the top three highest-expressed γTMT-like methyltransferases (MTs), namely TeMT1-3 (Genbank PP067959- PP067961), in descending order of expression levels in leaf tissue, for further investigation (Table 1). TeMT1 was the predominant MT in both leaf and root, showing transcripts per million (TPM) values 7.7 and 26.3 folds of those for TeMT2 and 3 (Table 1). In addition, TeMT1's leaf TPM (401.8) was comparable

10.3389/fpls.2024.1451298

TABLE 1 The transcripts per million (TPM) values and sequence alignment results of TePeNMT, TeMT2/3, and ten other vobasine biosynthetic genes in *Tabernaemontana elegans* leaf and root tissues.

enzyme	E-value	Amino acid identity (%)	Accession (E-value)	Leaf TPM	Root TPM
Te7DLS	0	94.3	Cr7DLS	85.6	1315.0
Te7DLH	0	89.8	Cr7DLH	14.4	324.7
Te7DLGT	0	90.2	Cr7DLGT	19.0	233.7
TeLAMT	0	83.6	CrLAMT	86.6	310.9
TeSLS	0	91.0	CrSLS	117.3	1450.0
TeSTR	2.2E-163	72.3	CrSTR	20.6	185.4
TeSGD	0	70.8	CrSGD	35.3	148.8
TeGS	0	88.7	CrGS	173.9	1430.5
TeSBE	0	83.4	RsSBE	450.2	1055.4
TeGO	0	92.5	CrGO	369.9	1080.6
TePeNMT(TeMT1)	1.7E-104	52.3	RsPiNMT	401.8	574.9
TeMT2	2.9E-110	53.9	RsPiNMT	52.5	20.0
TeMT3	2.7E-109	52.3	RsPiNMT	15.3	77.9

to other vobasine biosynthetic genes such as TeSBE (450.2) and TeGS (173.9). TeMT2 showed over 80% amino acid sequence identity to various *bona fide* γTMTs, suggesting its potential status as a γTMT. The low expression of other γTMT-like transcripts preclude them from further analyses.

Interestingly, both TeMT1 and 3 showed only 50% amino acid identity to the characterized CrPeNMT. In a phylogenetic analysis, TeMT2 clustered with other characterized γTMTs, while TeMT1 and 3 formed a distinct clade from both γTMTs and all other characterized MIA NMTs (Figure 1). Notably, these γTMT-like MTs were also distinguishable from BIA NMTs and MIA *O*-methyltransferases (OMT). In contrast, eight other vobasine biosynthetic enzymes are highly conserved (83-94% amino acid sequence identity) between *T. elegans* and two other well-studied MIA-producing species, *C. roseus* and *Rauwolfia serpentina* (Indian snakeroot) (Table 1). Notable exceptions included the putative strictosidine synthase TeSTR and putative strictosidine β-glucosidase TeSGD, which were 72% and 71% identity to *C. roseus* STR and SGD at amino acid level, respectively (Table 1). It is also worth noting that TeMT1-3 lacked chloroplast transit peptides found in *bona fide* γTMT, as predicted by TargetP 2.0 (https://services.healthtech.dtu.dk/services/TargetP-2.0/).

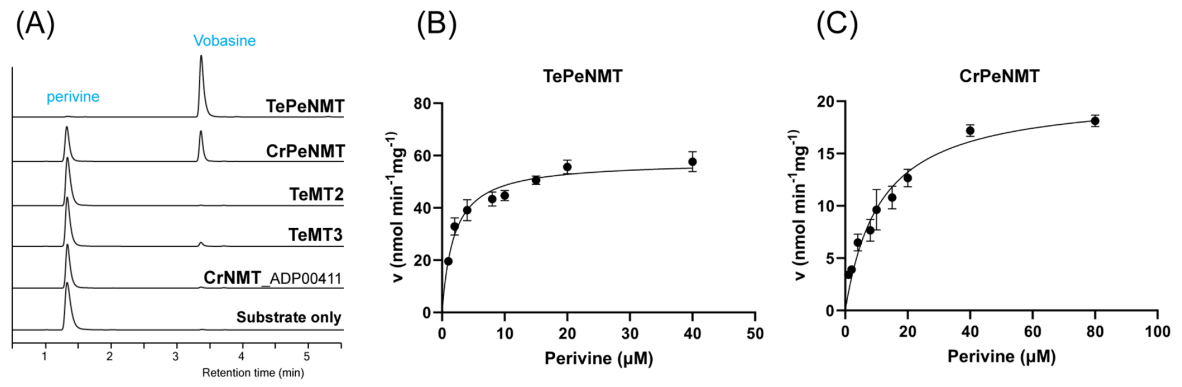


FIGURE 4
In vivo and *in vitro* biochemical characterizations identified TeMT1 as the perivine *N*β-methyltransferase in *Tabernaemontana elegans*. **(A)** *E. coli* cultures expressing CrPeNMT and TeMT1 both methylate the substrate perivine to vobasine as shown by LC-MS/MS multiple reaction monitoring (MRM) with $[M+H]^+$ *m/z* 339>166 for perivine and *m/z* 353>180 for vobasine. The MS/MS fingerprints used for selecting these parameters are included in Supplementary Figure 1. **(B, C)** The Michaelis-Menten saturation kinetics experiments for perivine substrate show superior kinetics for TePeNMT compared to CrPeNMT. Each data point shows the mean value of three technical replicates. The error bars indicate standard deviation. The curves are graphed with Prism Graphpad 9.5.0. The kinetics data are included in Table 1.

TeMT1 is the perivine *N*β-methyltransferase in *T. elegans*

To study their functions, we expressed *N*-terminal His-tagged TeMT1-3 along with the previously characterized CrPeNMT in *E. coli*. Additionally, we expressed the closest TeMT1 ortholog CrNMT-ADP00411 (Genbank ADP00411) from *C. roseus* in *E. coli* for comparison. This ortholog shared 61% amino acid identity, and clustered with TeMT1 and 3 in the phylogenetic analysis, suggesting a common evolution ancestry (Figure 1). Subsequently, we supplied perivine to cultures expressing these proteins. Consistent with previous research, CrPeNMT methylated perivine to vobasine, which co-eluted with purified vobasine standard (Figure 4A). TeMT1 also catalyzed perivine *N*β-methylation (Figure 4A). In contrast, TeMT2 did not show detectable perivine methylation, while TeMT3 and CrNMT-ADP00411 showed negligible perivine NMT activity (Figure 4A). Based on these results, we designated TeMT1 as *T. elegans* perivine *N*β-methyltransferase (TePeNMT). None of TeMT1-3 showed activity with 22 other MIAs encompassing multiple MIA skeleton families (Supplementary Table 3).

Following purification of recombinant TePeNMT and CrPeNMT through standard affinity chromatography (Supplementary Figure 25), we conducted *in vitro* enzyme kinetics assays using saturating co-substrate *S*-adenosylmethionine (SAM) and various perivine concentrations for comparison. Both recombinant enzymes methylated perivine substrate; however, TePeNMT showed superior enzyme kinetics (Figures 4B, C; Table 2). Specifically, CrPeNMT displayed a K_M value for perivine 5.9 times greater than that of TePeNMT, indicating significantly lower binding affinity towards perivine compared to TePeNMT. Moreover, TePeNMT exhibited a V_{max} 2.8 times higher than that of CrPeNMT. Consequently, the catalytic efficiency (k_{cat}/K_M) of TePeNMT, calculated from V_{max}/K_M, surpassed that of CrPeNMT by a factor of 16.6.

Homology modeling indicates significant disparities in both active site conformation and perivine substrate docking positions between TePeNMT and CrPeNMT

After observing the low identity in enzyme primary sequences, conducting phylogenetic analysis, and noting the kinetic differences, we further continued to investigate the tertiary structures of TePeNMT, CrPeNMT, TeMT3, and CrNMT-ADP00411 for their catalytic mechanisms. Using the *Arabidopsis thaliana* phosphoethanolamine NMT (PDB: 5WP5) as template, we built homology models for all four enzymes with the Molecular Operating Environment (MOE). While all NMTs shared a high degree of tertiary structure similarity, we found significant differences in their active sites, particularly in accommodating perivine binding and its *N*β-methylation via S_N2 displacement of the methylsulfonium group of SAM (Figure 5; Supplementary Figure 26).

TABLE 2 Michaelis-Menten saturation kinetics for TePeNMT and CrPeNMT.

	TePeNMT	CrPeNMT
K_M for perivine (μM)	1.91 (95% CI[a]: 1.47-2.47)	11.32 (95% CI: 8.58-14.78)
V_{max} (nmol mg protein^{-1} min^{-1})	57.83 (95% CI: 54.75-61.18)	20.70 (95% CI: 18.83-22.82)

[a]confidence interval.

For all NMTs, the SAM co-substrate was coordinated in the conserved glycine-rich region via multiple hydrogen bonds and hydrophobic interactions (Supplementary Figure 26). In TePeNMT, perivine was docked in a hydrophobic pocket primarily via van der Waals interactions with residues Trp184, Ser50/277, Val59/272, and His54 (Figure 5B; Supplementary Figure 26). The distance between SAM's methyl group and the perivine *N*β nitrogen (N···CH_3, 3.40 Å) and the optimal angle between the methyl and the *N*β nitrogen (N···CH_3–S, 175.5°) is compatible for efficient S_N2 displacement (Figures 5A, B). In comparison, perivine rotated over 90° in the active site of CrPeNMT, with the indole portion situated in a deeper binding pocket. The active site consisted of residues distinct from those found in TePeNMT. Specifically, perivine was coordinated with Asp258, Lys190/196/271, Thr262, Ile198, and Arg57 (Figure 5B; Supplementary Figure 26). This different docking pose of perivine altered both the angle between SAM's methyl group and the *N*β nitrogen (164.2°) and their distance (3.43 Å), likely contributing to reduced K_M and V_{max} for CrPeNMT. Conversely, perivine docking position was highly suboptimal for both TeMT3 and CrNMT-ADP00411 (Figure 5A). The angles between SAM's methyl group and the perivine *N*β nitrogen (137.0° and 143.1°) were not conducive for productive S_N2 methylation, aligning with their negligible perivine NMT activity.

Discussion

Monoterpenoid indole alkaloids (MIAs) stand as a fascinating alkaloid class found primarily in the families Apocynaceae, Loganiaceae, and Rubiaceae. Noted for their structural diversity and potent biological activities, an essential objective in this field of research is uncovering the biosynthetic pathways involved in the construction of these highly complex compounds. Notably, many of these biologically active molecules are either *N*-methylated themselves or are *N*-methylated as precursors in MIA biosynthesis. Methylation alters MIAs' solubility, electronic property, reactivity, and binding capacity to biological targets, thus providing a direct pathway to new bioactivities through a simple chemical transformation.

The characterization of *C. roseus* DhtNMT, a γ-tocopherol *C*-methyltransferase like enzyme involved in vindoline biosynthesis (Liscombe et al., 2010), has been a focal point in initiating the study of this interesting methyltransferase class of γ-TMT ancestry. All MIA NMTs known to date derive from γ-TMT, a lineage that includes VmPiNMT, RsPiNMT, RsNNMT, RsANMT, and CrPeNMT. The studies on these enzymes have not only elucidated their substrate specificity but also shed light on their altered subcellular localization, providing valuable clues about their metabolic roles within plant cells (Koudounas et al., 2022).

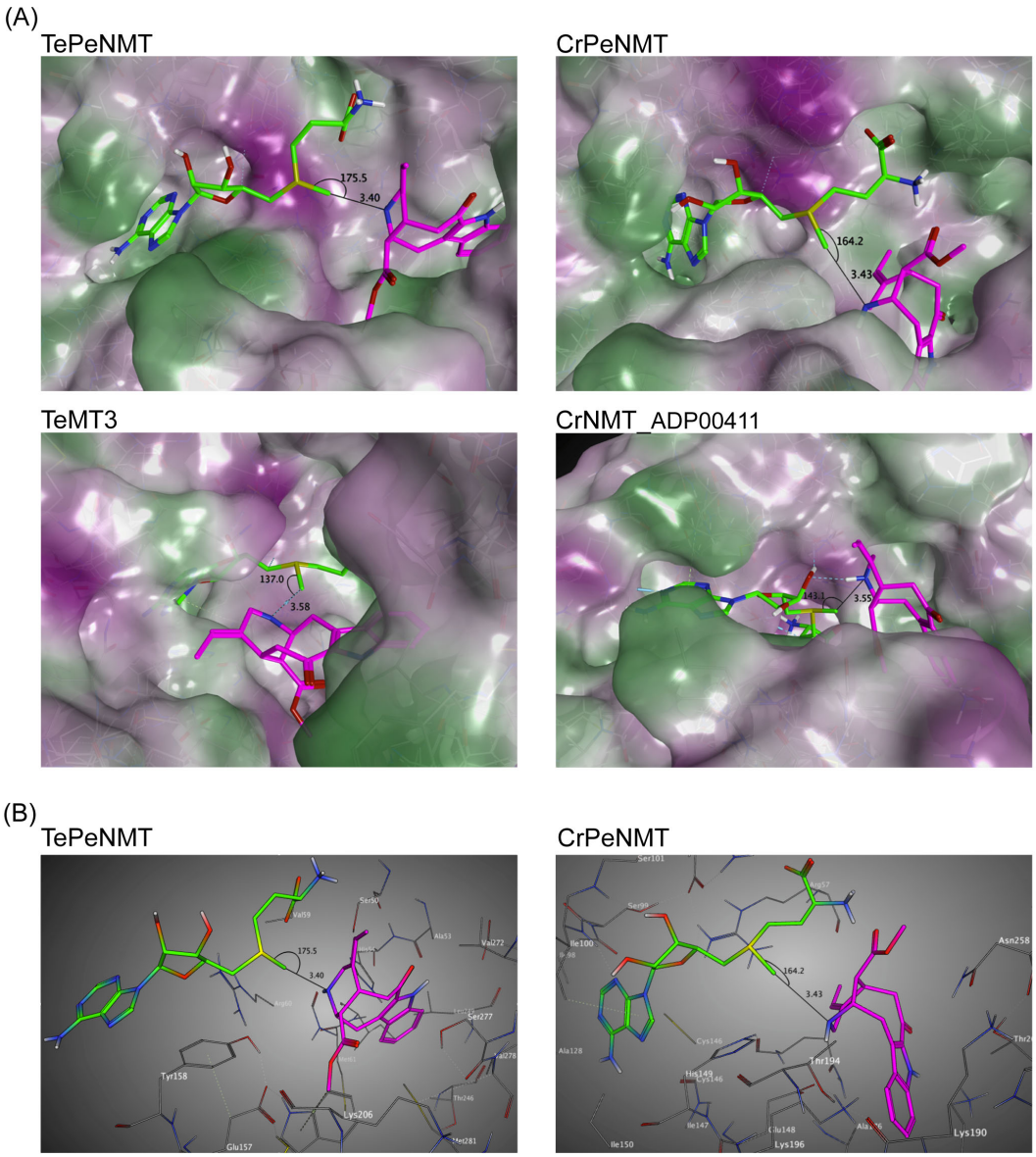


FIGURE 5
Homology modeling supports that TePeNMT's superior enzyme efficiency is contributed by significant differences in enzyme active sites and perivine binding poses between TePeNMT and CrPeNMT. Enzyme models are generated from the template *Arabidopsis thaliana* phosphoethanolamine NMT-2 (PDB: 5WP5.A) with MOE version 2022.02. *S*-adenosylmethionine (SAM) is shown in green, and perivine is shown in magenta. The distance and angel between SAM methylsulfonium donor and perivine-*N*β are illustrated. **(A)** the lipophilic mapping of four NMT enzymes, in which polar regions are shaded purple and lipophilic regions are shaded green. **(B)** The ball-and-stick diagrams of the same enzyme active sites for TePeNMT and CrPeNMT.

By investigating the metabolomes and transcriptomes of various *T. elegans* tissues, we identified a new γ-TMT-like enzyme TePeNMT, which functions as a perivine *N*β-methyltransferase. Its abundant expression in *T. elegans* leaf correlates with the high accumulation of vobasine observed in the same tissue (Figure 3, Table 1). Interestingly, while CrPeNMT has been previously characterized in *C. roseus* (Levac et al., 2022), this plant species is not known for the accumulation of vobasine or its further oxidized form, periformyline (*N*β-formylperivine), which have been documented in other Catharanthus species *C. trichophyllus* and *C. lanceus* (Maloney et al., 1965; Mukhopadhyay and Cordell, 1981). Our previous research has also demonstrated that perivine, rather than vobasine or periformyline, accumulates in the leaf tissue of *C. roseus* (Eng et al., 2022). Perhaps, a lower level of CrPeNMT or the cellular and subcellular compartmentalization of CrPeNMT and substrate perivine, could be a reason why perivine remains non-methylated in *C. roseus*. The physiological role of CrPeNMT remains to be elucidated.

Our phylogenetic analysis revealed a distinct evolutionary path for TePeNMT, setting it apart from all other characterized MIA NMTs (Figure 1). Interestingly, TePeNMT exhibits a closer relationship with bona fide γ-TMTs found in both MIA and non-

MIA producing species, such as *Arabidopsis thaliana*. Notably, an ortholog of TePeNMT in *C. roseus*, designated as CrNMT-ADP00411 and sharing 61% amino acid sequence identity, was identified in our phylogenetic study. However, this ortholog lacks perivine NMT activity, indicating that TePeNMT evolved specifically in *T. elegans* to catalyze perivine *N*β-methylation. TePeNMT and CrPeNMT only share 50% amino acid identity. The fact that eight other vobasine biosynthetic enzymes encompassing key enzymes such as geissoschizine synthase (GS) and sarpagan bridge enzyme (SBE) are highly conserved (83%-94% identity) between *T. elegans* and *C. roseus* (Table 1) further supports that the two PeNMTs evolved independently in these two species.

Moreover, our study on enzyme kinetics and homology modeling revealed significant catalytic disparities between TePeNMT and CrPeNMT (Figure 5). The low amino acid identity (50%) resulted in drastic differences in their active sites for vobasine binding. They barely shared any conserved amino acids, further supporting an independent evolution path that shaped the two perivine binding pockets. This variation in binding poses of perivine and SAM results in a superior disposition of the methylsulfonium donor and perivine *N*β in TePeNMT, facilitating a more efficient S_N2 reaction. The findings from homology modeling corroborate the kinetics data for these enzymes, wherein TePeNMT displayed superior performance over CrPeNMT in terms of both substrate affinity K_M and enzyme velocity V_{max} (Table 2).

Vobasine serves as a precursor to the potent anticancer drug (3'*R*)-hydroxytabernaelegantine C and stands as the most abundant MIA in *T. elegans* (Figures 2, 3). While it is unclear for the *in planta* role of vobasine and its derivatives tabernaemontanine and dregamine identified in this study, TePeNMT likely coevolved with the occurrence of vobasine, potentially contributing to the improved enzyme efficiency and fitness of *T. elegans*. Our phytochemical, phylogenetic, kinetic, and homology modeling data collectively suggested distinct evolutionary paths for the two PeNMTs from *T. elegans* and *C. roseus*. This presents another notable example of enzyme parallel evolution in the plant kingdom.

Materials and methods

Plant materials and alkaloid purification

Fresh leaves and stems (600 g total) of glasshouse grown *T. elegans* were submerged in methanol for a week. The extract was evaporated under vacuum and resuspended in 10 mL 1 M HCl. The suspension was extracted with ethyl acetate, and the aqueous phase was basified using NaOH to above pH 7. After extraction by ethyl acetate, the total alkaloids in organic phase were evaporated and dissolved in methanol. Total alkaloids were separated by thin layer chromatography (Sicica gel60 F254, Millipore Sigma, Rockville, MD, USA) with solvents ethyl acetate and methanol (9:1, v/v). After extracting the TLC bands with methanol, 14 mg vobasine, 8 mg tabernaemontanine, 6 mg dregamine, and 5 mg apparicine were recovered. For tissue alkaloid comparison, 100 mg tissues of leaf, root and flower tissues were ground in 2 ml methanol to generate total extract. For latex, 100 mg freshly collected latex was mixed with 2 ml methanol. The samples were analyzed by LC-MS/MS.

mRNA extraction and transcriptome assembly

T. elegans leaf and root tissues (100 mg) from glasshouse grown plants were ground into fine powder in liquid nitrogen. RNA were extracted using a Mornarch® RNA Cleanup Kit (New England Biolabs, Ipswich, MA, USA) according to the manufacture's protocol. The RNA was sequenced using Illumina NovaSeq 25M reads at the Atlantic Cancer Research Institute (Moncton, NB, Canada). The assembled transcriptomes were analyzed using CLC Genomic Workbench 20.0.4 (Qiagen, Redwood City, CA, USA).

LC-MS/MS and NMR

The samples were analyzed by Ultivo Triple Quadrupole LC-MS/MS system from Agilent (Santa Clara, CA, USA), equipped with an Avantor® ACE® UltraCore C18 2.5 Super C18 column (50×3 mm, particle size 2.5 µg), as well as a photodiode array detector and a mass spectrometer. For alkaloid analysis, the following solvent systems were used: Solvent A, Methanol: Acetonitrile: Ammonium acetate 1 M, water at 29:71:2:398; solvent B, Methanol: Acetonitrile: Ammonium acetate 1 M: water at 130:320:0.25:49.7. The following linear elution gradient was used: 0-5 min 80% A, 20% B; 5-5.8 min 1% A, 99% B; 5.8-8 min 80% A, 20% B; the flow during the analysis was constant and 0.6 ml/min. The photodiode array detector range was 200 to 500 nm. The mass spectrometer was operated with the gas temperature at 300°C and gas flow of 10 L/min. Capillary voltage was 4 kV from m/z 100 to m/z 1000 with scan time 500 ms and the fragmentor performed at 135 V with positive polarity. The MRM mode was operated as same as scan mode with the adjusted precursor and product ion. ^{1}H (400 MHz) and ^{13}C (100 MHz) NMR spectra were recorded on an Agilent 400 MR and a Bruker Avance III HD 400 MHz spectrometers in $CDCl_3$.

Protein expression and purification

All MT sequences were codon optimized, synthesized, and subcloned into pET30a(+) vector between NotI/XhoI restriction sites (Bio Basic Inc., Toronto, ON, Canada). CrPeNMT (Genbank KC708453) was amplified from *C. roseus* leaf cDNA using primers (5'-ATAGGATCCAATGGCCTCAATGGGAGAGAAGGA and 5'-ATAGTCGACTCATTTAGTTTTGCGAAATGTAACTG) and cloned in pET30b+ vector within BamHI/SalI sites. The clones were transformed and expressed in *Escherichia coli* strain BL21-(DE3). An overnight culture was used to inoculate 200 mL LB media, which was grown at 37°C to OD_{600} between 0.6-0.8. The cultures were then induced with IPTG at a final concentration of 0.1 mM and shifted to 15°C with shaking at 200 rpm. After approximately 12 hours of growth, the cells were harvested by centrifugation. The pellet cells were sonicated and resuspended in ice-cold lysis buffer (25 mM Tris-HCl

at pH 7.5, 100 mM NaCl, 10% glycerol, and 20 mM imidazole) and lysed using sonication. The lysate was centrifuged at 4°C, 6,000 g for 25 minutes. The resulting supernatant was subjected to purification using standard Ni-NTA affinity chromatography. The column was washed twice with 3 mL of wash buffer (25 mM Tris-HCl at pH 7.5, 100 mM NaCl, 10% glycerol, and 30 mM imidazole) to rid of proteins with low binding affinities. Recombinant MT proteins were eluted using buffer containing imidazole at 250 mM and desalted in the storage buffer (25 mM Tris-HCl at pH 7.5, 100 mM NaCl, 10% glycerol) using a PD-10 Desalting column (Cytiva, Marlborough, MA, USA) and stored in -80° C until analysis.

In vivo biotransformation and *in vitro* kinetics assay

For *in vivo* biotransformation experiment, an aliquot (1 mL) of induced culture was centrifuged to collect cell pellet, which was resuspended in 0.5 mL Tris-HCl at pH 7.5 supplemented with 2 µg perivine or other alkaloid substrates. The biotransformation alkaloid substrates was carried out at 37°C, 200 rpm for 5 h. An aliquot of the biotransformation broth was mixed with equal volume of methanol and analyzed by LC-MS/MS. *In vitro* enzyme kinetics assays (50 µl) contained 20 mM Tris pH 7.5, 3 µg of the purified recombinant proteins, 50µM *S*-adenosyl-methionine, and perivine concentrations ranging from 1 to 80 µM. After adding the enzyme, the reaction was incubated in a 30°C water bath for 5 min and subsequently terminated by adding 150 µL of methanol. The product formation was analyzed by LC-MS/MS. Triplicated experiments were graphed using GraphPad Prism (9.5.0) (Boston, MA, USA).

Homology modeling and substrate docking studies

All computational experiments and visualization were carried out with MOE version 2022.02 on local computers or on the Digital Research Alliance of Canada (DRAC, formerly Compute Canada) Advanced Research Computing Network (alliancecan.ca). All molecular mechanics calculations and simulations were conducted using the AMBER14:EHT forcefield with Reaction Field solvation. AM1-BCC charges were used for modeling the coenzyme and the substrates while AMBER charges were used for the enzymes. Sequence homology searches for both TePeNMT and CrPeNMT on the Protein DataBank found phosphoethanolamine *N*-methyltransferase 2 (AtPMT2) from *Arabidopsis thaliana* to be the best template for homology modeling (PDB ID: 5WP5) based on best Hidden Markov Model energy scores; *S*-adenosyl-homocysteine (SAH) was co-crystallized with the AtPMT2 enzyme. Sequence identity of TePeNMT and CrPeNMT to the relevant 5WP5 domain was 19.3%. and 19.2%, respectively. Homology models were then derived in MOE using default settings and scored using the GBVI/WSA dG method. SAM was modeled and docked into each homology model using the location of SAH in the 5WP5 enzyme as general docking region and default MOE-Dock settings. Perivine was then modeled and docked to the PeNMT-SAM complexes; for each perivine docking experiment, 100,000 docking poses were initially generated via the Triangle Matcher method and scored by the London dG function. A subset comprising the best docking poses were refined by the induced fit method where the bound ligands and active site residues were submitted to local geometry optimization and rescored by the GBVI/WSA dG scoring function. The top scoring docking poses that were geometrically compatible for β*N* methylation of perivine by SAM via S_N2 displacement were retained for subsequent analyses. Cartesian coordinates for the TePeNMT and CrPeNMT homology models and their perivine docked complexes can be found in the Supplementary Information.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI, PRJNA1148476.

Author contributions

MF: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. MR: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. DD: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. ZM: Formal analysis, Investigation, Writing – original draft. DN: Formal analysis, Investigation, Writing – original draft. GD: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Writing – original draft, Writing – review & editing. YQ: Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work is supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant RGPIN-2020-04133 to YQ.

Acknowledgments

The authors thank the Chemical Computing Group ULC (www.chemcomp.com) for MOE licenses. This research was enabled in part by support provided by ACENET (www.ace-net.ca) and the Digital Research Alliance of Canada (alliancecan.ca).

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

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1451298/full#supplementary-material

References

Beek, T. A. V., Lankhorst, P. P., and Verpoorte, R. (1985). Monogagaine, a novel dimeric indole alkaloid from Tabernaemontana chippii and Tabernaemontana dichotoma. *Z Naturforsch.* 40, 693–701. doi: 10.1515/znb-1985-0521

Bennasar, M.-L., Zulaica, E., Solé, D., and Alonso, S. (2009). The first total synthesis of (±)-apparicine. *Chem. Commun.* 0, 3372–3374. doi: 10.1039/b903577j

Cázares-Flores, P., Levac, D., and Luca, V. D. (2016). Rauvolfia serpentina N-methyltransferases involved in ajmaline and Nβ -methylajmaline biosynthesis belong to a gene family derived from γ-tocopherol C-methyltransferase. *Plant Journal : Cell Mol. Biol.* 87, 335–342. doi: 10.1111/tpj.13186

Dagnino, D., Schripsema, J., and Verpoorte, R. (1994). Terpenoid indole alkaloid biotransformation capacity of suspension cultures of Tabernaemontana divaricata. *Phytochemistry* 35, 671–676. doi: 10.1016/S0031-9422(00)90584-3

Deluca, V., Balsevich, J., Tyler, R. T., and Kurz, W. G. (1987). Characterization of a novel N-methyltransferase (NMT) from Catharanthus roseus plants : Detection of NMT and other enzymes of the indole alkaloid biosynthetic pathway in different cell suspension culture systems. *Plant Cell Rep.* 6, 458–461. doi: 10.1007/BF00272782

Eng, J. G. M., Shahsavarani, M., Smith, D. P., Hájíček, J., Luca, V. D., and Qu, Y. (2022). A Catharanthus roseus Fe(II)/α-ketoglutarate-dependent dioxygenase catalyzes a redox-neutral reaction responsible for vindolinine biosynthesis. *Nat. Commun.* 13, 3335. doi: 10.1038/s41467-022-31100-1

Facchini, P. J., and Luca, V. D. (2008). Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. *Plant Journal : Cell Mol. Biol.* 54, 763–784. doi: 10.1111/j.1365-313X.2008.03438.x

Ferreira, M.-J. U., and Paterna, A. (2019). Monoterpene indole alkaloids as leads for targeting multidrug resistant cancer cells from the African medicinal plant Tabernaemontana elegans. *Phytochem. Rev.* 18, 971–987. doi: 10.1007/s11101-019-09615-1

Guedes, J. G., Ribeiro, R., Carqueijeiro, I., Guimarães, A. L., Bispo, C., Archer, J., et al. (2023). The leaf idioblastome of the medicinal plant Catharanthus roseus is associated with stress resistance and alkaloid metabolism. *J. Exp. Bot.* 75, 274–299. doi: 10.1093/jxb/erad374

Koudounas, K., Guirimand, G., Hoyos, L. F. R., Carqueijeiro, I., Cruz, P. L., Stander, E., et al. (2022). Tonoplast and peroxisome targeting of γ-tocopherol N-methyltransferase homologs involved in the synthesis of monoterpene indole alkaloids. *Plant Cell Physiol.* 62, 200–216. doi: 10.1093/pcp/pcab160

Kutney, J. P., Horinaka, A., Ward, R. S., and Worth, B. R. (1980). Studies on the total synthesis of bisindole alkaloids within the voacamine family. *Can. J. Chem.* 58, 1829–1838. doi: 10.1139/v80-288

Lam, K. C., Ibrahim, R. K., Behdad, B., and Dayanandan, S. (2007). Structure, function, and evolution of plant O-methyltransferases. *Genome* 50, 1001–1013. doi: 10.1139/G07-077

Levac, D., Cázares, P., Yu, F., and Luca, V. D. (2016). A picrinine N-methyltransferase belongs to a new family of γ-tocopherol-like methyltransferases found in medicinal plants that make biologically active monoterpenoid indole alkaloids. *Plant Physiol.* 170, 1935–1944. doi: 10.1104/pp.15.01813

Levac, D., Flores, P. C., and Luca, V. D. (2022). Molecular and biochemical characterization of Catharanthus roseus perivine-N β -methyltransferase. *Phytochemistry* 201, 113266. doi: 10.1016/j.phytochem.2022.113266

Levac, D., Murata, J., Kim, W. S., and Luca, V. D. (2008). Application of carborundum abrasion for investigating the leaf epidermis: molecular cloning of Catharanthus roseus 16-hydroxytabersonine-16-O-methyltransferase. *Plant Journal : Cell Mol. Biol.* 53, 225–236. doi: 10.1111/j.1365-313X.2007.03337.x

Liscombe, D. K., Usera, A. R., and O'Connor, S. E. (2010). Homolog of tocopherol C methyltransferases catalyzes N methylation in anticancer alkaloid biosynthesis. *Proc. Natl. Acad. Sci.* 107, 18793–18798. doi: 10.1073/pnas.1009003107

Luca, V. D., and Cutler, A. J. (1987). Subcellular localization of enzymes involved in indole alkaloid biosynthesis in catharanthus roseus. *Plant Physiol.* 85, 1099–1102. doi: 10.1104/pp.85.4.1099

Maloney, E. M., Farnsworth, N. R., Blomster, R. N., Abraham, D. J., and Sharkey, A. G. (1965). Catharantbus Lanceus VII. Isolation of tetrahydroalstonine, lochnerinine, and periformyline. *J. Pharm. Sci.* 54, 1166–1168. doi: 10.1002/jps.2600540815

Mansoor, T. A., Ramalho, R. M., Mulhovo, S., Rodrigues, C. M. P., and Ferreira, M. J. U. (2009). Induction of apoptosis in HuH-7 cancer cells by monoterpene and β-carboline indole alkaloids isolated from the leaves of Tabernaemontana elegans. *Bioorg Med. Chem. Lett.* 19, 4255–4258. doi: 10.1016/j.bmcl.2009.05.104

Morris, J. S., and Facchini, P. J. (2019). Molecular origins of functional diversity in benzylisoquinoline alkaloid methyltransferases. *Front. Plant Sci.* 10, 1058. doi: 10.3389/fpls.2019.01058

Mukhopadhyay, S., and Cordell, G. A. (1981). Catharanthus Alkaloids. XXXVI. Isolation of Vincaleukoblastine (VLB) and Periformyline From Catharanthus trichophyllus and Pericyclivine From Catharanthus roseus. *J. Nat. Prod* 44, 335–339. doi: 10.1021/np50015a017

Paterna, A., Gomes, S. E., Borralho, P. M., Mulhovo, S., Rodrigues, C. M. P., and Ferreira, M.-J. U. (2016a). Vobasinyl-iboga alkaloids from tabernaemontana elegans: cell cycle arrest and apoptosis-inducing activity in HCT116 colon cancer cells. *J. Natural products* 79, 2624–2634. doi: 10.1021/acs.jnatprod.6b00552

Paterna, A., Gomes, S. E., Borralho, P. M., Mulhovo, S., Rodrigues, C. M. P., and Ferreira, M.-J. U. (2016b). (3′R)-hydroxytabernaelegantine C: A bisindole alkaloid with potent apoptosis inducing activity in colon (HCT116, SW620) and liver (HepG2) cancer cells. *J. Ethnopharmacol* 194, 236–244. doi: 10.1016/j.jep.2016.09.020

Paterna, A., Kincses, A., Spengler, G., Mulhovo, S., Molnár, J., and Ferreira, M.-J. U. (2017). Dregamine and tabernaemontanine derivatives as ABCB1 modulators on resistant cancer cells. *Eur. J. Med. Chem.* 128, 247–257. doi: 10.1016/j.ejmech.2017.01.044

Pereira, P. S., França, S. de C., de Oliveira, P. V. A., Breves, C. M. de S., Pereira, S. I. V., Sampaio, S. V., et al. (2008). Chemical constituents from Tabernaemontana catharinensis root bark: a brief NMR review of indole alkaloids and *in vitro* cytotoxicity. *Química Nova* 31, 20–24. doi: 10.1590/S0100-40422008000100004

Singh, B., Sharma, R. A., and Vyas, G. K. (2011). Antimicrobial, antineoplastic and cytotoxic activities of indole alkaloids from tabernaemontana divaricata (L.) R.Br. *Curr. Pharm. Anal.* 7, 125–132. doi: 10.2174/157341211795684844

Stander, E. A., Lehka, B., Carqueijeiro, I., Cuello, C., Hansson, F. G., Jansen, H. J., et al. (2023). The Rauvolfia tetraphylla genome suggests multiple distinct biosynthetic routes for yohimbane monoterpene indole alkaloids. *Commun. Biol.* 6, 1197. doi: 10.1038/s42003-023-05574-8

Uzaki, M., Mori, T., Sato, M., Wakazaki, M., Takeda-Kamiya, N., Yamamoto, K., et al. (2024). Integration of cell differentiation and initiation of monoterpenoid indole alkaloid metabolism in seed germination of Catharanthus roseus. *N Phytol.* 242, 1156–1171. doi: 10.1111/nph.19662

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 12 November 2024
DOI 10.3389/fpls.2024.1473527

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Qiang Zhang,
Shandong Agricultural University, China
Muhammad Junaid Rao,
Zhejiang Agriculture & Forestry University, China

*CORRESPONDENCE
Chenggang Liang
201503001@gznu.edu.cn

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

RECEIVED 31 July 2024
ACCEPTED 16 October 2024
PUBLISHED 12 November 2024

CITATION
Wei K, Chen X, Cheng Z, Wang H, Wang F, Yang L, Wu S, Yang Y, Tu Y, Wang Y and Liang C (2024) LC-MS/MS-based metabolic profiling: unraveling the impact of varying degrees of curing on metabolite transformations in tobacco.
Front. Plant Sci. 15:1473527.
doi: 10.3389/fpls.2024.1473527

LC-MS/MS-based metabolic profiling: unraveling the impact of varying degrees of curing on metabolite transformations in tobacco

Kesu Wei[1†], Xuling Chen[2†], Zhijun Cheng[3], Heng Wang[4], Feng Wang[1], Lei Yang[3], Shengjiang Wu[1], Yijun Yang[3], Yonggao Tu[1], Yan Wang[2] and Chenggang Liang[2*]

[1]Upland Flue-cured Tobacco Quality and Ecology Key Laboratory, China National Tobacco Corporation (CNTC), Guizhou Academy of Tobacco Science, Guiyang, China, [2]College of Life Science, Guizhou Normal University, Guiyang, China, [3]China Tobacco Hunan Industrial Co. Ltd, Changsha, China, [4]School of Biological and Environmental Engineering, Guiyang College, Guiyang, China

The curing process regulates metabolite transformations of leaves and significantly influences the formation of tobacco quality. This study investigated the major physicochemical compositions and metabolic profiles under normal curing (NC), excessive curing (EC), and insufficient curing (IC) treatments. The results indicated that the contents of nicotine, nitrogen, potassium, and chlorine remained stable among treatments, while the sugar content in EC was significantly lower than in IC. LC-MS/MS identified 845 metabolites, with flavonoids as the most abundant class. Comparative analyses identified a series of differentially expressed metabolites (DEMs) among fresh leaf, NC, EC, and IC leaves at the end of 42°C, 54°C, and 68°C, respectively. At the end of 68°C, 256 up-regulated and 241 down-regulated common DEMs across treatments were isolated in comparison to fresh leaf, underscoring the consistency of metabolic changes during curing. Notably, nonivamide varied markedly across treatments, suggesting its potential as a key curing indicator. NC_68°C displayed 11 up-regulated and 17 down-regulated unique DEMs, differing from EC_68°C and IC_68°C, suggesting their potential availability in evaluating tobacco leaf quality. KEGG pathway analysis revealed temporal shifts in metabolic pathways, particularly those involved in secondary metabolite biosynthesis (such as flavonoids, flavones, flavonols) and amino acid metabolism, during the transition from yellowing to color-fixing. Correlation analysis isolated the top 25 DEMs correlated with curing degree and stage, which might play pivotal roles in the curing process and could serve as potential biomarkers for assessing curing degree and stage. Specifically, D-(+)-cellobiose displayed the strongest negative correlation with curing degree, while 5,7-dihydroxychromone exhibited the highest positive correlation coefficient. Furthermore, curcurbitacin IIa showed the highest positive correlation with curing stage, followed by hesperetin and 8-shogaol. Additionally, random forest analysis emphasized morellic acid as a core molecular metabolite across curing degrees, suggesting its potential as a biomarker. Debiased

sparse partial correlation (DSPC) network analysis further pinpointed hispidulin as a key metabolite, underscoring its significance in elucidating flavonoid metabolism during the curing process. Collectively, this study enhances the understanding of metabolite transformations underlying tobacco curing processes and provides a valuable reference for optimizing curing strategies to achieve desired outcomes.

KEYWORDS

tobacco, curing degree, curing stage, metabolite transformation, flavonoids, sugar

1 Introduction

Tobacco (*Nicotiana tabacum* L.), a globally economically significant crop, not only plays a pivotal role in agricultural economies but also holds a unique position in aromatic industrial applications (Mavroeidis et al., 2024). Tobacco leaves undergo a series of intricate processing steps, with curing being a critical phase that significantly impacts the final quality and flavor of the tobacco (Li et al., 2022). Curing involves carefully controlled conditions of temperature, humidity, and airflow, aimed at gradually drying the leaves while promoting desired chemical reactions to enhance the aroma, reduce moisture content, and prepare the leaves for further processing and consumption (Bacon et al., 1952). Notably, the flue-curing process, which typically lasts 6-7 days, involves the application of artificial heat to elicit a cascade of biochemical reactions. These reactions transform the leaf color and chemical composition, ultimately dictating the quality and marketability of the final product (Abubakar et al., 2000; Chen et al., 2019; Meng et al., 2024).

The curing process is typically characterized by three pivotal stages: the yellowing stage, the color fixing stage, and the dry tendon stage (Zou et al., 2019; Meng et al., 2024). Specifically, the degree of yellowing during this process is a critical parameter that not only influences leaf visual appeal but also plays a pivotal role in determining the chemical profile and subsequent industrial applications (Zou et al., 2019). Therefore, the formation and accumulation of aromatic compounds, polyphenols, and other flavor-related chemicals during this process have been extensively studied, as they contribute significantly to the smoking characteristics and overall quality of the cured tobacco (Vaughan et al., 2008; Lim et al., 2022; Zou et al., 2023). It is widely acknowledged that the color of flue-cured tobacco, serving as a crucial quality indicator, exhibits a strong correlation with natural pigment content during the curing process (Meng et al., 2024). Leaf performance is commonly utilized to evaluate and guide decisions related to tobacco quality and industrial applicability (Xin et al., 2023).

The curing process of tobacco involves various aspects and factors, including drying methods, temperature, humidity, duration, draft fan control, and the use of exogenous additives. Studies on this process have elucidated the transformation law of major chemical compositions in different types of tobacco under various flue-curing methods, indicating the feasibility of combining multiple curing methods to enrich style characteristics and potentially enhance tobacco quality (Bacon et al., 1952). In comparison to oven-drying, flue-curing leads to significant decreases in four plastid pigments (lutein, chlorophyll A, chlorophyll B, β-carotene) and notable increases in six polyphenol substances (neochlorogenic acid, chlorogenic acid, caffeic acid, chrysatropic acid, rutin, kaempferol) (Zong et al., 2022). Optimization of wind speed parameters through the use of a heat pump-powered curing barn and a three-stage curing process has been shown to significantly enhance the baking quality of tobacco leaves, promoting the accumulation of aroma substances and the degradation of macromolecular substances (Sun et al., 2023). Furthermore, the pectinase preparation derived from *Bacillus amyloliquefaciens* W6-2 has been found to effectively enhance the quality of flue-cured tobacco, as evidenced by improvements in aroma, sweetness, and smoothness, along with alterations in the levels of macromolecules and volatile components (Weng et al., 2024). The application of starch-degrading bacteria has been shown to significantly improve tobacco leaf quality by reducing macromolecule content, increasing water-soluble total sugar and reducing sugar levels, and enhancing the production of crucial volatile aroma components (Gong et al., 2023). Another study revealed that the application of γ-PGA in flue-cured tobacco leaves influenced the accumulation and transformation of carbon and nitrogen compounds, regulating the carbon and nitrogen metabolic processes during leaf growth and ultimately affecting tobacco leaf quality (Gao et al., 2023). Additionally, the presence of stems has been found to alter leaf metabolism, prevent browning, and enhance starch degradation during flue-curing, ultimately resulting in lower starch content in leaves (Meng et al., 2022).

Recently, metabolomics has emerged as a powerful tool for comprehensively profiling the metabolic shifts that occur during various biological processes. For instance, GC-MS analysis has identified 128 flavor chemicals in tobacco flavor capsules, with menthol as the dominant component and the carcinogenic compound pulegone also detected, although exposure margins were below safety thresholds (Lim et al., 2022). LC-MS/MS analysis identified 259 and 178 differentially expressed metabolites (DEMs) between 0.4 M sucrose-treated and control tobacco leaves

in the early and late stages of air-curing, respectively, revealing alterations in carbohydrate and amino acid metabolism that promote normal chlorophyll degradation and mitigate the green spot phenomenon through sucrose treatment (Li et al., 2023). Furthermore, GC-MS and CE-MS analyses have uncovered significant metabolic differences between growing districts, with a complex carbon and nitrogen metabolic network modulated by environmental factors (Zhao et al., 2015). A recent study elucidated the key chemical components contributing to the honey-sweet and burnt aroma characteristics of Shandong flue-cured tobacco, identifying 29 aroma precursors positively correlated with sensory quality, thereby providing guidance for targeted improvement and precise regulation of flavor-style characteristics (Li et al., 2024). UPLC-Q-TOF MS and GC-MS identified alpha-cembratriene-diol, beta-cembratriene-diol, sucrose esters III, and cembratriene-diol oxide as primary contributors to antifungal activity, providing insights for the development of botanical pesticides and multipurpose utilization of tobacco germplasms (Liu et al., 2024).

It is widely recognized that curing technology significantly influences the transformation of leaf metabolites, which in turn directly affects tobacco quality and price. Yunyan 87, a highly regarded tobacco cultivar, is notable for its wide adaptability, robust stress resistance, stable yield, ease of curing, and superior quality, which is characterized by a striking golden to orange-yellow color, rich oil content, high gloss, and moderate thickness, all collectively contributing to its prominent market value (Li et al., 2001). However, the metabolite transformation across different degrees of curing remains incompletely understood. Typically, the assessment of whether curing meets the standard relies heavily on visual data, lacking quality indicators such as metabolites. Therefore, the primary objective of this study is to conduct a detailed analysis of metabolite profiles at three typical stages of the flue-curing process in Yunyan 87 using the LC-MS/MS platform, specifically comparing normal curing (NC), excessive curing (EC), and insufficient curing (IC) treatments, and to investigate the differential transformation of key metabolites, which is expected to provide a comprehensive understanding of how varying degrees of curing can influence the transformation of metabolites in tobacco and offer valuable insights for optimizing curing practices.

2 Results and discussion

2.1 Physicochemical composition

To identify the differences in major physicochemical compositions among various treatments, tobacco leaf samples were collected and analyzed following the completion of the 68°C curing process (Table 1). Despite the absence of statistical differences in leaf weight and leaf density across treatments, EC demonstrated slightly lower levels of both parameters compared to NC and IC. Nitrogen content in tobacco leaves exerts multifaceted effects on quality, primarily as a constituent of the nicotine molecule, which directly influences nicotine concentrations and, consequently, the overall quality of flue-cured tobacco leaves (Henry et al., 2019). As shown in Table 1, both nitrogen and nicotine content in tobacco were largely unaffected by the different treatments, suggesting minimal alterations in the conversion of nitrogen and nicotine, despite variations in curing temperature and duration. Potassium and chlorine contents also play crucial roles in determining the quality of tobacco leaves (Attoe, 1946). The potassium and chlorine contents are found to be relatively stable in air-cured, flue-cured, and sun-cured tobacco leaves (Chen et al., 2021). Similarly, these components did not exhibit significant changes in response to the treatments. It is well established that sugars in tobacco leaves play a significant role in determining quality, as they influence the formation of harmful compounds and smoking properties. High temperatures in flue and sun curing elevate final sugar content, whereas low temperatures in air curing decrease sugar levels (Banožić et al., 2020). A notable observation was the variation in sugar content among the treatments, with EC exhibiting the lowest sugar content, followed by NC and IC. The sugar content of EC was significantly lower than that of IC. Therefore, variations in sugar transformation due to different curing conditions might have implications for the final quality of tobacco.

2.2 Multivariate analysis of metabolomics

Metabolomics is a powerful tool for uncovering the rules of metabolite transformation during tobacco leaf curing (Zong et al., 2022). To gain insights into the impact of different curing degrees on metabolome dynamics in tobacco leaves, metabolomics analysis was conducted using an LC–MS/MS platform. A total of 845 metabolites were identified and predominantly categorized into 20 distinct groups, including 198 flavonoids, 147 alkaloids, 114 terpenoids, 59 amino acids and peptides, 31 phenols and phenolic acids, 31 polyketides, 30 lipids, 27 organic acids and oxygen compounds, 23 carbohydrates, 22 coumarins, 20 benzene derivatives, 19 phenylpropanoids, and aromatic compounds (Figure 1A). Principal Component Analysis (PCA) revealed distinct metabolome profiles between the NC, IC, and EC treatments at 42, 54, and 68°C, with PC1 and PC2 explaining 72.9% and 8.4% of the variance, respectively (Figure 1B). Subsequently, OPLS-DA analysis was employed to enhance the differentiation among the NC, IC, and EC treatments, confirming the suitability of the model for identifying differentially expressed metabolites (DEMs), with Component 1 and Component 2 accounting for 70.4% and 7.3% of the variance, respectively (Figure 1C).

2.3 Analysis of DEMs

The heat map demonstrated significant variations in metabolite levels during curing (Figure 2A). Pairwise comparison analyses revealed 319 up-regulated and 313 down-regulated DEMs (p-value ≤ 0.05 and VIP value ≥ 1) in the comparison of NC_68°C vs fresh leaf, 306 up-regulated and 297 down-regulated DEMs in IC_68°C vs fresh leaf, and 301 up-regulated and 301 down-regulated DEMs in EC_68°C vs fresh leaf (Figure 2B). Among these, 256 up-regulated and 241 down-regulated common DEMs were consistently observed across NC_68°C vs fresh leaf, IC_68°C vs fresh leaf, and EC_68°C vs fresh leaf (Figure 3A). Notably, 91 DEMs exhibited at

TABLE 1 The major physicochemical compositions of tobacco under different treatments.

Treatment	Leaf weight (g)	Leaf density (g/m^2)	Sugar content (%)	Nicotine content (%)	Nitrogen content (%)	Potassium content (%)	Chlorine content (%)
NC	8.5 ± 0.7a	75.7 ± 1.5a	19.0 ± 2.8ab	3.51 ± 0.08a	2.44 ± 0.17a	1.47 ± 0.02a	0.25 ± 0.03a
EC	7.8 ± 0.5a	68.7 ± 5.0a	15.9 ± 3.3b	3.42 ± 0.32a	2.33 ± 0.10a	1.52 ± 0.07a	0.27 ± 0.04a
IC	8.8 ± 0.4a	74.0 ± 6.1a	21.4 ± 1.9a	3.46 ± 0.12a	2.40 ± 0.09a	1.58 ± 0.10a	0.25 ± 0.01a

NC, normal curing; EC, excessive curing; IC, insufficient curing. Different lowercase letters indicate significant differences among treatments according to One-way ANOVA.

least a 10-fold change (82 up-regulated and 9 down-regulated), indicating consistent and significant changes during the curing process (Figure 3B; Supplementary Table S1).

The top 20 fold-change DEMs during curing under NC condition are present in Table 2. At the yellowing stage, notable increases exceeding 100-fold were detected in isochlorogenic acid B, 3,5-dicaffeoylquinic acid, 1-caffeoylquinic acid, isochlorogenic acid C, and toosendanin. Conversely, silychristin, N-acetyl-L-carnosine, isoanhydroicaritin, nonivamide, fluoxetine, and adenine exhibited decreases of more than 100-fold. During color fixing, scopolamine hydrobromide, scopolamine HBr trihydrate, brevifolincarboxylic acid, and 3-nitro-L-tyrosine increased greater than 50-fold, while acesulfame, isoschaftoside, vicenin III, and ipecoside underwent comparable declines. At the dry tendon stage, polydatin levels surged by over 50-fold, but decursinol, Robinin, apigenin 7-o-(2g-rhamnosyl) gentiobioside, scopolamine hydrobromide, luteolin 3',7-di-o-glucoside, scopolamine HBr trihydrate, 3-nitro-L-tyrosine, and D-(-)-penicillamine decreased by more than 100-fold (Table 2). Furthermore, 6 DEMs demonstrated substantial changes across two stages, with 4 of them (scopolamine hydrobromide, scopolamine HBr trihydrate, 3-nitro-L-tyrosine, and D-(-)-penicillamine) experiencing a sharp rise in the color fixing stage, followed by a steep decline in the dry tendon stage.

In contrast, each pairwise comparison identified unique DEMs, indicative of distinct metabolic responses to the various treatments. Specifically, NC_68°C vs fresh leaf showed 11 up-regulated and 17 down-regulated DEMs, IC_68°C vs fresh leaf had 12 up-regulated and 20 down-regulated, while EC_68°C vs fresh leaf exhibited 23 up-regulated and 27 down-regulated DEMs (Figure 3B; Supplementary Table S1). Further pairwise analyses uncovered shared and exclusive DEMs between treatment comparisons. For instance, 34 up-regulated and 26 down-regulated DEMs were

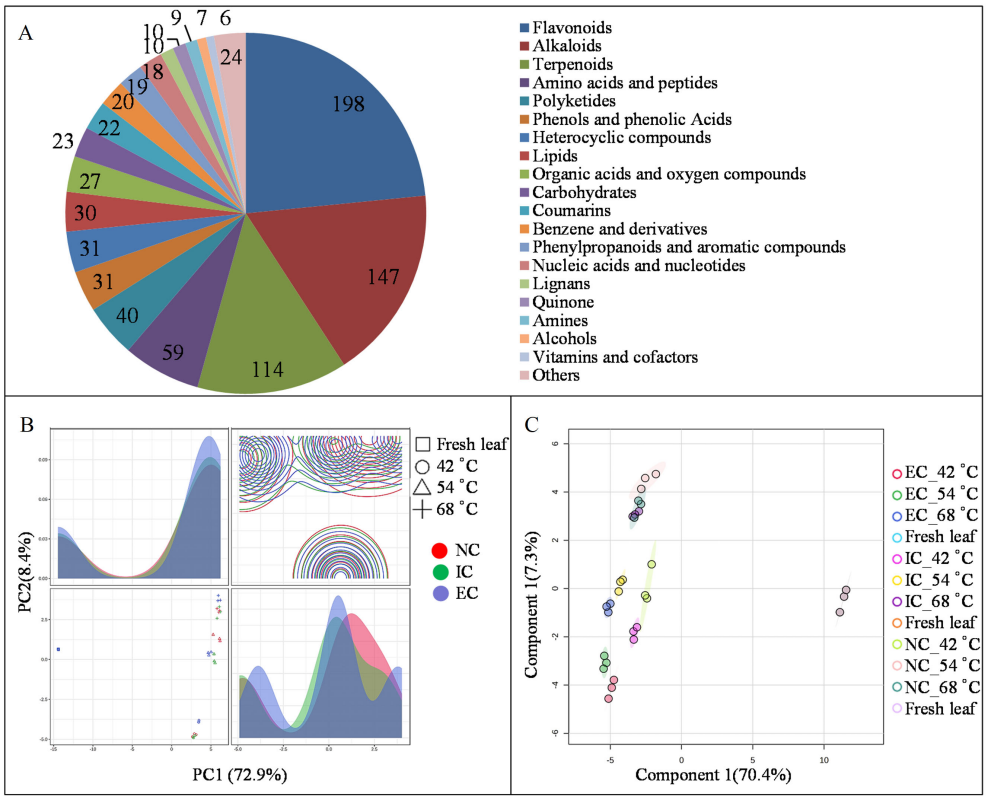


FIGURE 1
The groups of metabolite **(A)**, PCA score **(B)**, and OPLS-DA score **(C)** by LC-MS/MS analysis. EC represents excessive curing, IC represents insufficient curing, NC represents normal curing, the same as below.

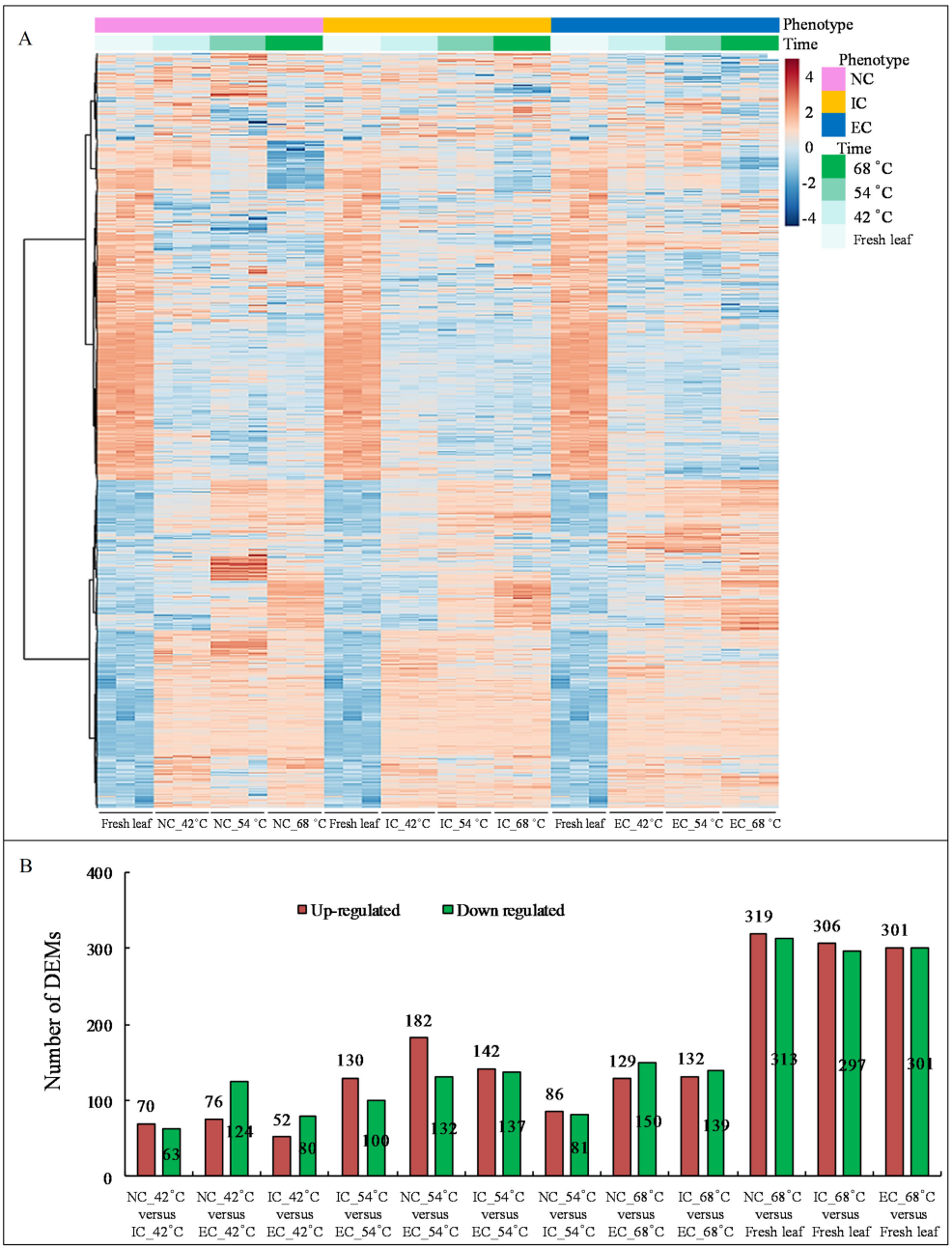


FIGURE 2
The heatmap of metabolites **(A)** and the number of differentially expressed merabolites by pairwise comparison analyses **(B)**.

common to NC_68°C vs fresh leaf and IC_68°C vs fresh leaf, while 18 up-regulated and 29 down-regulated DEMs were shared between NC_68°C vs fresh leaf and EC_68°C vs fresh leaf. The comparison between IC_68°C vs fresh leaf and EC_68°C vs fresh leaf showed 4 up-regulated and 7 down-regulated DEMs (Figure 3A). Importantly, unique sets of metabolites with more than 10-fold differential expression were identified, with corrections made to avoid repetition: 5, 5, and 17 up-regulated DEMs (each unique to a comparison), along with 1, 1, and 0 down-regulated DEMs, respectively (Figure 3B; Supplementary Table S1). Additionally, in examining the intersections among the pairwise groups, distinct patterns emerged, with 12, 11, and 6 up-regulated DEMs, and 3, 5, and 0 down-regulated common DEMs observed, respectively, providing insights into the overlap and divergence of metabolic responses under different curing degrees.

During curing, metabolite levels exhibited moderate variation across treatments at identical temperatures. Comparative analyses at 42, 54, and 68°C revealed the following differentially expressed metabolites (DEMs): 70, 130, and 86 up-regulated, and 63, 100, and 81 down-regulated DEMs in NC vs IC comparison; 76, 182, and 129 up-regulated, and 124, 137, and 150 down-regulated DEMs in NC vs EC comparison; and 52, 142, and 132 up-regulated, with 80, 137, and 139 down-regulated DEMs in IC vs EC comparison, respectively (Figure 2B).

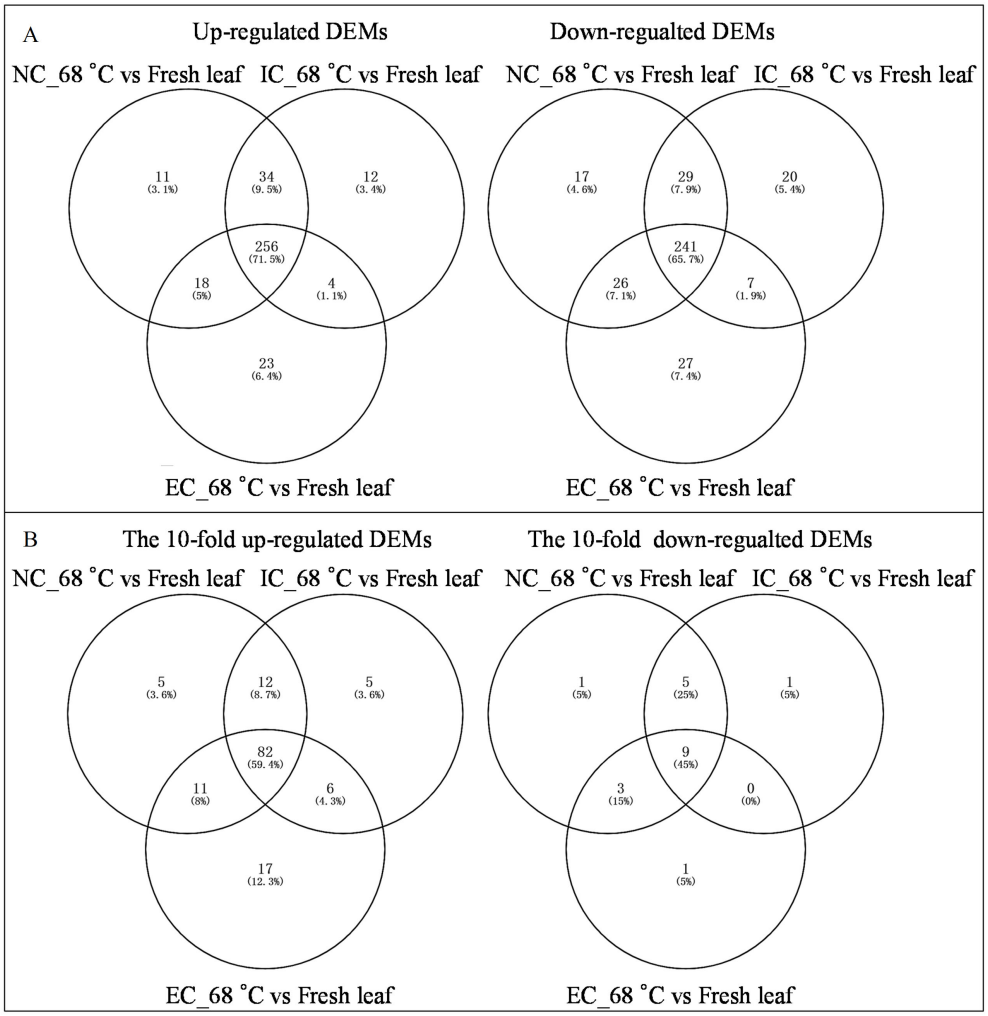


FIGURE 3
Venn diagram of differential expressed metabolites **(A)** and the 10-fold differential expressed metabolites **(B)** between tobacco cured leaves subjected to different curing treatments and fresh leaf.

Within the NC vs IC group, 5 (2 up-regulated and 3 down-regulated), 39 (25 up-regulated and 14 down-regulated), and 3 (2 up-regulated and 1 down-regulated) DEMs changed at least 3-fold at 42, 54, and 68°C, respectively. In the NC vs EC group, 23 (2 up-regulated and 21 down-regulated), 65 (41 up-regulated and 24 down-regulated), and 21 (5 up-regulated and 16 down-regulated) DEMs changed at least 3-fold at 42, 54, and 68°C, respectively. And in the IC vs EC group, 23 (2 up-regulated and 8 down-regulated), 22 (9 up-regulated and 13 down-regulated), and 31 (9 up-regulated and 22 down-regulated) DEMs changed at least 3-fold at 42, 54, and 68°C, respectively (Supplementary Table S2).

The top 20 fold-change DEMs at the end of curing among treatments are presented in Table 3. In the comparison of NC_68°C vs. IC_68°C, 7-ethyl-10-hydroxy-camptothecin (3.89-fold), prim-o-glucosylcimifugin (3.08-fold), and cimetidine (2.88-fold) were the most up-regulated DEMs, while di-o-methylquercetin (0.26-fold) was the most down-regulated. When comparing IC_68°C to EC_68°C, S-(-)-carbidopa (30.46-fold), acyclovir (8.84-fold), and acetylcorynoline (5.34-fold) showed the highest up-regulation, whereas several DEMs, including nonivamide (0.10-fold), myricetin (0.12-fold), liquiritin (0.13-fold), and valsartan (0.16-fold), were significantly down-regulated. In the NC_68°C vs. EC_68°C comparison, S-(-)-carbidopa (7.92-fold), cimetidine (6.88-fold), and acyclovir (5.23-fold) were the most up-regulated DEMs, while octyl gallate (0.14-fold), liquiritin (0.17-fold), myricetin (0.17-fold), and valsartan (0.18-fold) were the most down-regulated (Table 3). Furthermore, 15 common DEMs were detected among treatments (Table 3). Notably, nonivamide exhibited substantial variations across the three treatments, with the highest levels observed in EC_68°C, moderate levels in NC_68°C, and the lowest levels in IC_68°C, suggesting that nonivamide might be served as a key metabolite for assessing the degree of curing.

2.4 Analysis of correlation

The top 25 metabolites correlated with the phenotype (curing degree) were isolated, with 10 positively correlated and 15 negatively

TABLE 2 The top 20 fold-change DEMs at three pivotal stages under normal curing condition.

Yellowing stage (NC 42°C vs Fresh leaf)		Color fixing stage (NC 54°C vs NC 42°C)		Dry tendon stage (NC 68°C vs NC 54°C)	
Metabolite	Fold	Metabolite	Fold	Metabolite	Fold
Isochlorogenic acid B	318.95	**Scopolamine hydrobromide**	130.09	Polydatin	50.17
3,5-dicaffeoylquinic acid	279.94	**Scopolamine HBr trihydrate**	112.86	Byakangelicol	9.87
1-caffeoylquinic acid	264.48	Brevifolincarboxylic acid	89.9	Vicine	8.75
Isochlorogenic acid C	234.44	**3-nitro-L-tyrosine**	70.44	(S)-3-(allylsulphinyl)-L-alanine	7.22
Toosendanin	103.54	Ethylenediaminetetraacetic acid	34.97	Deoxyandrographolide	7.01
5-iodo-2'-deoxyuridine	94.33	Paeoniflorin	31.95	**Acesulfame**	5.49
Periplocymarin	89.38	Monocrotaline	31.64	Piperonyl acetone	5.26
7-hydroxycoumarin	88.56	Valsartan	22.18	Glycitin	5.20
Gentiopicrin	86.91	**D-(-)-penicillamine**	21.00	Lithosprmoside	5.06
Apigenin-7-o-beta-d-glucoside	78.48	Leonurine	17.47	Amarogentin	4.99
Bisdemethoxycurcumin	0.11	Ranaconitine	0.31	Marmesin	0.11
L-isoleucine	0.11	Daidzin	0.28	Polyphyllin VI	0.11
Bilobalide	0.11	Tyramine	0.28	Decursinol	0.09
p-Hydroxy-cinnamic acid	0.11	Deoxyadenosine monophosphate	0.22	Kaempferol-3-o-robinoside-7-o-rhamnoside (Robinin)	0.07
Silychristin	0.09	**Fluoxetine**	0.21	Apigenin 7-o-(2G-rhamnosyl) gentiobioside	0.07
N-acetyl-L-carnosine	0.08	Acetylcorynoline	0.21	**Scopolamine hydrobromide**	0.07
Isoanhydroicaritin	0.07	**Acesulfame**	0.17	Luteolin 3',7-di-o-glucoside	0.07
Nonivamide	0.07	Isoschaftoside	0.09	**Scopolamine HBr trihydrate**	0.06
Fluoxetine	0.07	Vicenin III	0.06	**3-Nitro-L-tyrosine**	0.04
Adenine	0.06	Ipecoside	0.05	**D-(-)-Penicillamine**	0.03

The bolded fonts represent the common DEMs among treatments.

correlated (Figure 4A). Flavonoids, the most abundant class of metabolites in tobacco leaves, possess diverse biological functions (Hu et al., 2021). The contents of flavonoid metabolites exhibited a significant decrease throughout the air-curing period (Geng et al., 2024). Notably, 5 flavonoids and 4 carbohydrates were negatively correlated, whereas 3 flavonoids showed positive correlation with the phenotype (Figure 4A). D-(+)-Cellobiose displayed the strongest negative correlation, while 5,7-Dihydroxychromone exhibited the highest positive correlation coefficient. Carbohydrates serve as vital precursors for aroma and bioactive compounds (Banožić et al., 2020). The observed negative correlation between sucrose content and phenotype, consistent with the sugar content data in Table 1, underscores the profound influence of curing degree on tobacco leaf carbohydrate metabolism.

Furthermore, the top 25 metabolites were positively correlated with the time (curing stage), encompassing a diverse array of compounds crucial to tobacco quality (Figure 4B). It is well established that terpenoids (Demole and Enggist, 1975), phenols and phenolic acids (Zou et al., 2021), heterocyclic compounds (Higashio and Shoji, 2004), amino acids and peptides (Wang et al., 2020), and alkaloids (Shoji et al., 2024) are well-known contributors to the aroma, flavor, and overall sensory quality of tobacco leaves. Notably, curcurbitacin IIa showed the highest positive correlation with time, followed closely by hesperetin and 8-shogaol, indicating their potential significance in the curing process.

2.5 Analysis of Kyoto encyclopedia of genes and genomes pathways

The KEGG analysis revealed a striking similarity in metabolic strategies employed under varying curing degrees, as evidenced by the collective enrichment of the same top 10 metabolic pathways among the DEMs derived from the comparisons of NC_68°C vs fresh leaf, EC_68°C vs fresh leaf, and IC_68°C vs fresh leaf (Figure 5). Consistent with previous findings (Yang et al., 2023), this study also observed a prominent enrichment of DEMs in the flavonoid biosynthesis pathway and amino acid metabolism. Specifically, the synthesis of various flavonoid metabolites, such as isoflavonoids, flavonoids, flavones, and flavonols, along with amino acid metabolism, were enriched. However, despite these shared enrichments, closer inspection revealed distinct patterns within the

TABLE 3 The top 20 fold-change DEMs at the end of curing among treatments.

NC_68°C vs IC_68°C		IC_68°C vs EC_68°C		NC_68°C vs EC_68°C	
Metabolite	Fold	Metabolite	Fold	Metabolite	Fold
7-ethyl-10-hydroxy-camptothecin	3.89	**S-(-)-carbidopa**	30.46	**S-(-)-Carbidopa**	7.92
Prim-o-glucosylcimifugin	3.08	**Acyclovir**	8.84	**Cimetidine**	6.88
Cimetidine	2.88	**Acetylcorynoline**	5.34	**Acyclovir**	5.23
Glabridin	2.83	Picfeltarraenin IB	4.63	Ipecoside	3.71
Laetanine	2.8	**Byakangelicol**	3.84	**Glycitin**	3.48
Bavachin	2.75	**Deoxyandrographolide**	3.69	Adenosine cyclophosphate	2.98
Isobavachalcone	2.73	14-deoxyandrographolide	3.67	Isoschaftoside	2.95
Homoplantaginin	2.70	Gentiopicrin	3.23	**Deoxyandrographolide**	2.91
Chrysoeriol 5-o-hexoside	2.70	**Glycitin**	3.11	6”-O-Acetylglycitin	2.88
Emodin-3-methyl ether/Physcion	2.47	Apigenin-7-o-beta-d-glucoside	0.21	**Byakangelicol**	2.88
Diosmetin-7-o-beta-d-glucopyranoside	2.46	Genistin	0.21	**10-Hydroxydecanoic acid**	0.25
Tectoridin	2.43	**Chrysosplenetin B**	0.21	**Casticin**	0.24
Nonivamide	2.43	**Casticin**	0.18	**Chrysosplenetin B**	0.23
Morellic acid	0.47	**Valsartan**	0.16	**Nonivamide**	0.23
Tetrahydropiperine	0.44	**7-ethyl-10-hydroxy-camptothecin**	0.15	Atropine sulfate monohydrate	0.22
Acetylcorynoline	0.43	**Octyl gallate**	0.15	Mangiferin	0.21
L-alanyl-L-phenylalanine	0.42	**Liquiritin**	0.13	**Valsartan**	0.18
Toosendanin	0.41	**Myricetin**	0.12	**Liquiritin**	0.17
S-(-)-carbidopa	0.41	**10-Hydroxydecanoic acid**	0.11	**Myricetin**	0.17
Di-o-methylquercetin	0.26	**Nonivamide**	0.10	**Octyl gallate**	0.14

The bolded fonts represent the common DEMs among treatments.

flavonoid biosynthesis pathways. Notably, the isoflavonoid biosynthesis pathway, flavone and flavonol biosynthesis pathway exhibited variations among different treatments, indicating that the flavonoid metabolism was indeed influenced by the specific curing degrees.

The top 10 metabolic pathways covered by the DEMs in the comparisons of NC_42°C vs fresh leaf, NC_54°C vs NC_42°C, and NC_68°C vs NC_54°C reflected the distinct treatment effects. The yellowing stage represents the apex of primary chemical component conversion, during which numerous metabolites within tobacco leaves undergo profound alterations (Zou et al., 2019). A comparative analysis between NC_42°C and fresh leaf, revealed a notable enrichment of up-regulated metabolic pathways pertaining to secondary metabolite biosynthesis, particularly flavonoid, flavone, and flavonol biosynthesis, as well as alkaloid biosynthesis. Conversely, pathways such as ABC transporters, tyrosine metabolism, and nucleotide metabolism were down-regulated, indicating a shift towards enhanced secondary metabolite synthesis accompanied by augmented amino acid degradation. Further investigation comparing NC_54°C to NC_42°C showed a distinct pattern, with only two pathways (isoflavonoid biosynthesis and flavone/flavonol biosynthesis) being down-regulated, while the remaining eight pathways, including those related to secondary metabolite and flavonoid biosynthesis, were up-regulated. This suggested a continuation and intensification of metabolic activity in these pathways during the transition from yellowing to color fixing. However, in the comparison between NC_68°C and NC_54°C, while isoflavonoid biosynthesis was up-regulated, pathways associated with secondary metabolite and flavonoid biosynthesis, as well as tryptophan metabolism, exhibited down-regulation. These findings implied that metabolic pathways, especially those linked to secondary metabolites and flavonoids, undergo a surge during the yellowing to color fixing transition but might experience a subsequent decline as leaves progress towards the dry tendon stage. This dynamic regulation of metabolic pathways highlighted the intricacies and temporal specificity involved in tobacco curing processes.

Additionally, the results of the comparisons between different treatments at 68°C revealed intricate regulatory patterns in metabolic pathways. Specifically, when comparing NC_68°C to IC_68°C, DEMs were enriched in up-regulated biosynthesis of secondary metabolites and down-regulated isoflavonoid biosynthesis. In contrast, NC_68°C vs EC_68°C showed DEMs enriched in up-regulated flavone and flavonol biosynthesis and down-regulated biosynthesis of secondary metabolites. Furthermore, IC_68°C vs EC_68°C showed DEMs enriched in up-regulated isoflavonoid biosynthesis and down-regulated

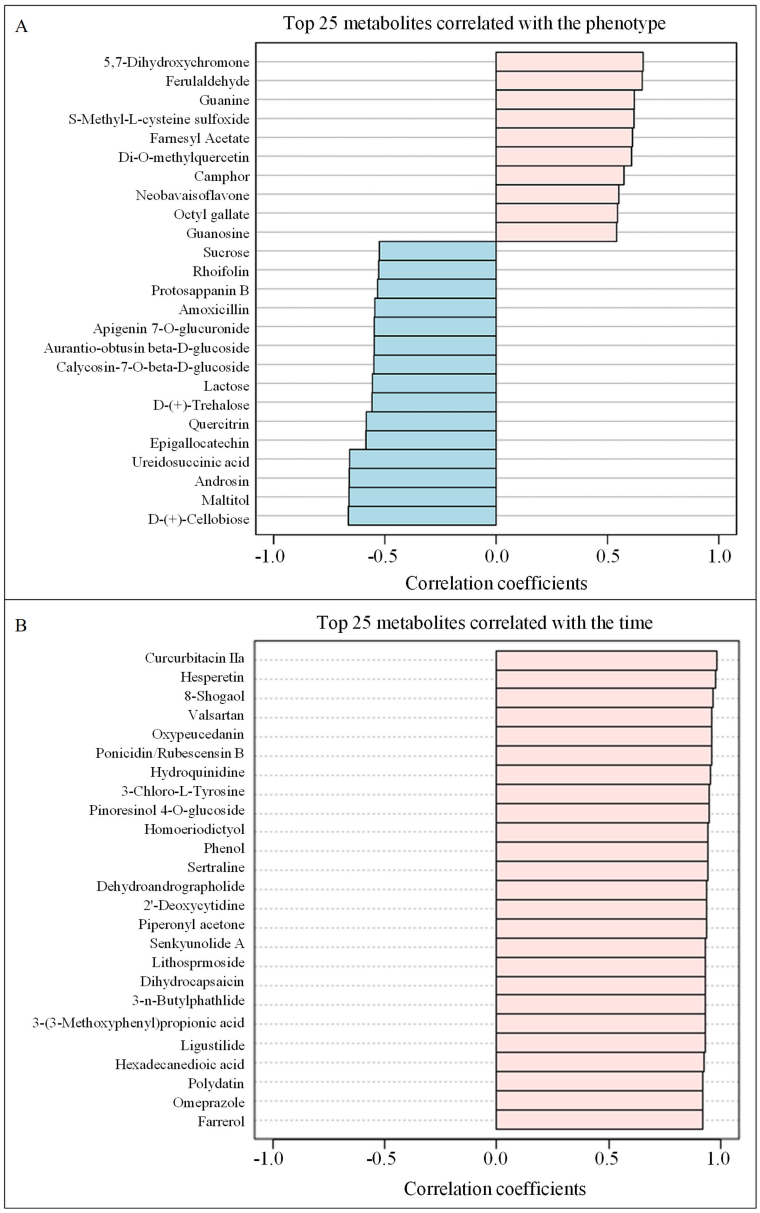


FIGURE 4
The correlation of the top 25 metabolites with the phenotype **(A)** and the time **(B)**.

biosynthesis of secondary metabolites. These findings suggested that the curing degrees significantly impacted the metabolic landscape of tobacco leaves, particularly with respect to secondary metabolite biosynthesis. The differential regulation of specific pathways, such as isoflavonoid, flavone, and flavonol biosynthesis, highlighted the intricate interplay between treatments and metabolic fluxes during tobacco curing.

2.6 Analysis of random forest

Random forest analysis, a powerful tool for identifying key variables in complex datasets, was employed to uncover the metabolic underpinnings of tobacco leaf responses to varying curing regimes (Figure 6A). Notably, morellic acid, a natural compound known for its antioxidant and anti-inflammatory properties (Aswathy et al., 2024), emerged as the core molecular metabolite across different treatment conditions and temperature variations. This finding underscores the potential of morellic acid as a biomarker for monitoring metabolic adjustments in tobacco leaves during curing.

2.7 Analysis of DSPC network

DPSC analysis provided a visualization of the correlation networks among metabolites. Flavonoids, known for their diverse biological functions, are the largest class of metabolites in tobacco

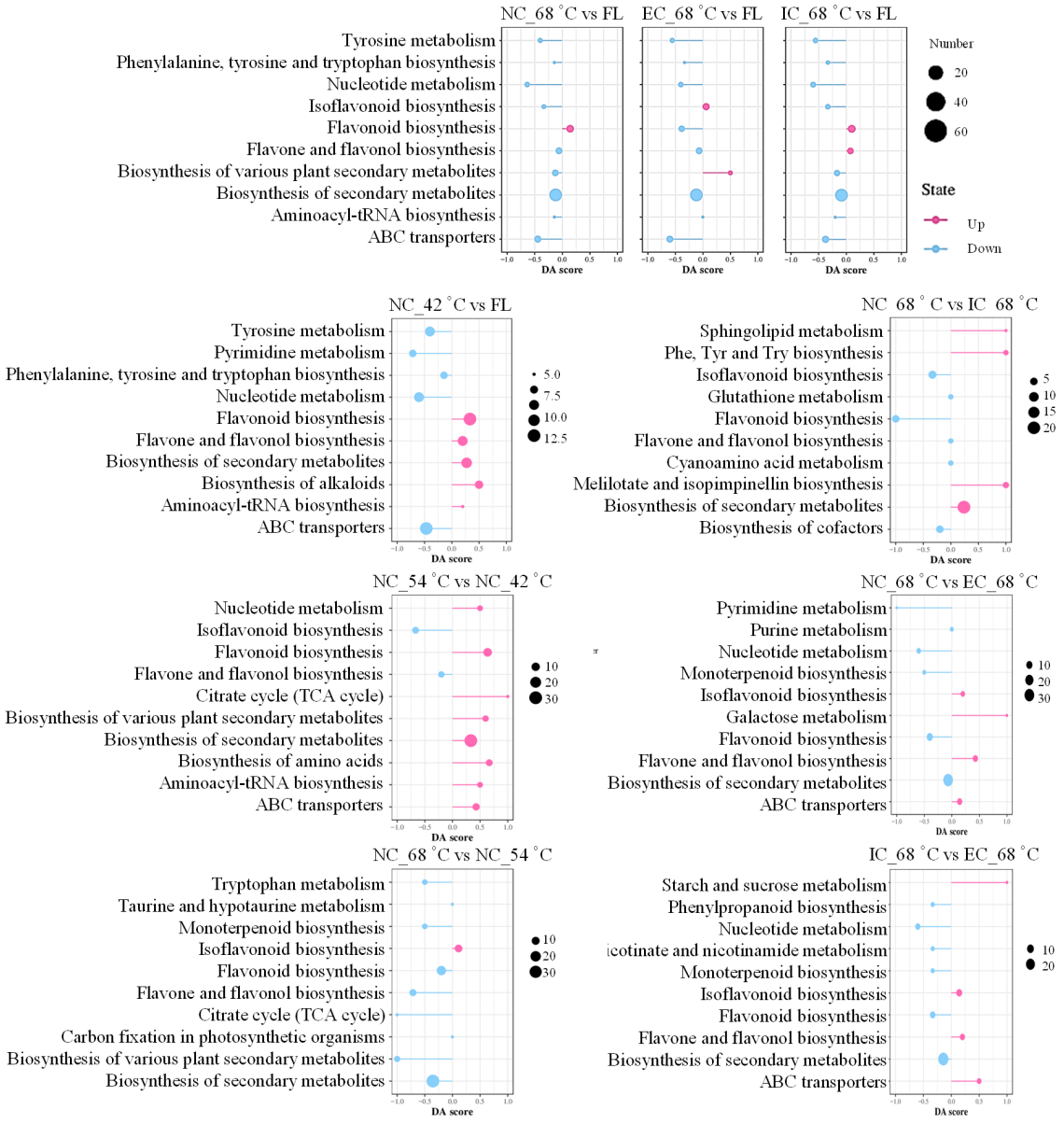


FIGURE 5
Analysis of differential expressed metabolites using KEGG pathway analysis. FL represents fresh leaf.

leaves (Hu et al., 2021). Notably, hispidulin, an anticancer flavone, was identified as the core metabolite by DPSC among the seven isolated flavonoids (Rout et al., 2024). Hispidulin showed positive correlations with four flavonoids and negative correlations with two others (Figure 6B), indicating its potential value in elucidating the metabolic pathway of flavonoids.

3 Conclusions

Curing degrees (NC, EC, and IC) minimally affected the physicochemical properties of tobacco leaves, such as leaf density, leaf weight, and the content of nicotine, nitrogen, chlorine, and potassium. However, the sugar content varied among three curing degrees, notably showing that EC exhibited significantly less sugar compared to IC. LC-MS/MS identified 845 metabolites, with flavonoids predominant. Heatmap analysis indicated significant changes in metabolites throughout curing stages and treatments. Comparative analyses isolated 256 up-regulated and 241 down-regulated DEMs consistently changing from fresh leaf to cured leaves, with 91 DEMs demonstrating at least a 10-fold alteration, highlighting the substantial metabolite transformation during curing. Notably, nonivamide varied markedly across treatments, suggesting its potential as a key curing indicator. NC displayed 11 up- and 17 down-regulated unique DEMs compared to EC and IC, indicating their potential role in tobacco quality formation. KEGG pathway analysis indicated a significant shift in metabolic pathways during the critical stages from yellowing to color fixing, particularly

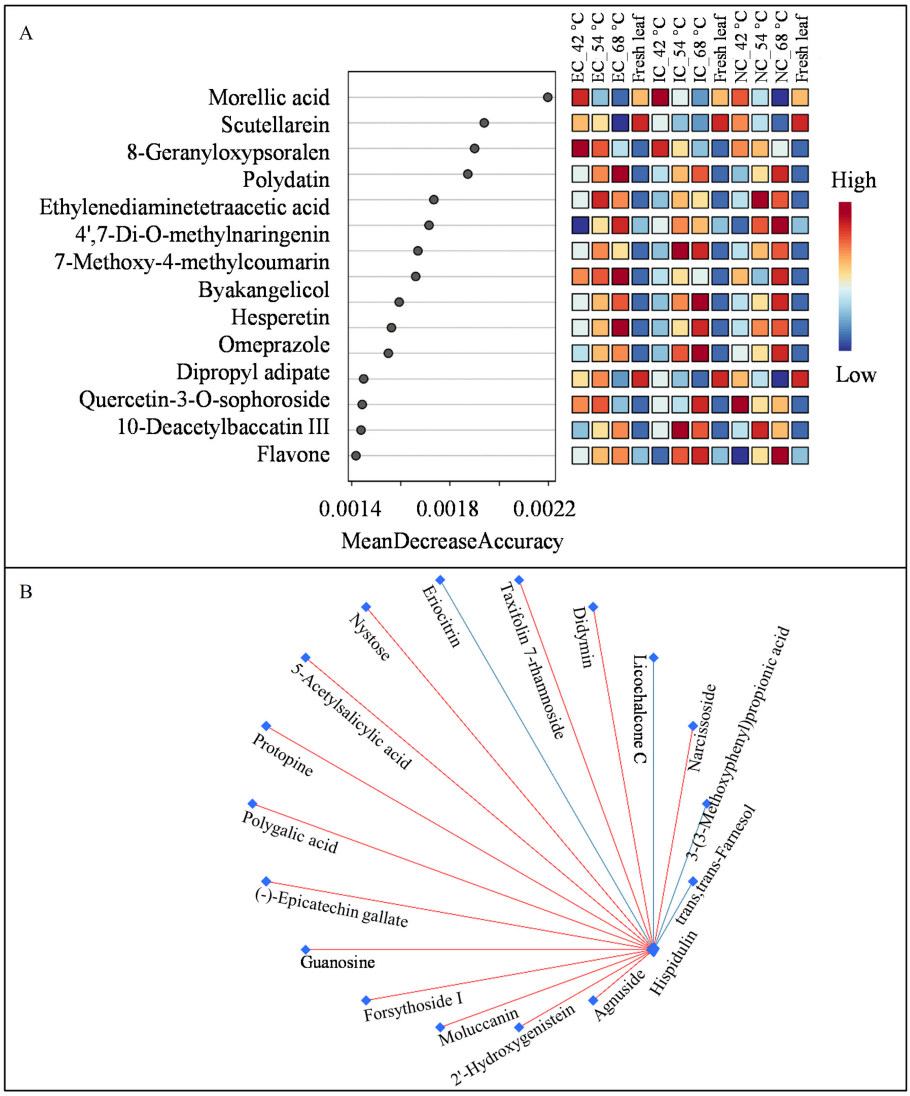


FIGURE 6
Analysis of metabolites by random forest **(A)** and debiased sparse partial correlation network **(B)**.

those related to secondary metabolite biosynthesis and amino acid metabolism, potentially contributing to the desired flavor and aroma profiles. Correlation analysis isolated the top 25 DEMs correlated with curing degree and stage, which might play pivotal roles in the curing process and could serve as potential biomarkers for assessing curing degree and stage. Specifically, D-(+)-cellobiose displayed the strongest negative correlation with curing degree, followed by maltitol and androsin. Conversely, 5,7-dihydroxychromone exhibited the highest positive correlation coefficient, followed by ferulaldehyde and guanine. Furthermore, curcurbitacin IIa showed the highest positive correlation with curing stage, followed by hesperetin and 8-shogaol. Random forest analysis highlighted morellic acid as a consistent core metabolite across curing conditions, suggesting its potential as a biomarker. Moreover, DPSC analysis pinpointed hispidulin as a key metabolite, underlining its importance in flavonoid metabolism elucidation. Collectively, this study enhances the understanding of the metabolite transformation underlying tobacco curing processes and provides valuable insights for optimizing curing strategies to achieve desired product characteristics.

4 Materials and methods

4.1 Plant growth and sampling

Tobacco variety Yunyan 87 was cultivated at the research farm of Guizhou Academy of Tobacco Science in Fuquan City, Guizhou Province, China, situated at an altitude of 1200 meters belonging to subtropical monsoon climate. The soil composition consisted of 25.32 g/kg organic matter, 138.73 mg/kg of available nitrogen (N), 36.32 mg/kg of phosphorus (P), 218.69 mg/kg of potassium (K), and a pH value of 6.2. The experiment was conducted using a randomized block design, with each plot covering an area of 121 m^2 and a planting density of 1.1 m × 0.55 m, replicated three times. Seeding was sown on January 20th, followed by transplantation on April 25th. The

average temperatures ranged from 16.13°C to 23.17°C in May, 19.23°C to 24.93°C in June, 22.87°C to 29.83°C in July, and 21.30°C to 30.37°C in August. The base fertilizer regime included 525 kg/ha of compound fertilizer (with an N:P:K ratio of 10:10:25), 450 kg/ha of fermented oilseed meal and 375 kg/ha of calcium-magnesium-phosphate fertilizer. On the day of transplantation, each plant was treated with 150-200 mL of a water-soluble fertilizer solution, containing 1% compound fertilizer and 0.28% cyhalothrin emulsifiable concentrate. Ten days after transplantation, an additional 100-150 mL of water-soluble fertilizer, formulated with 4% compound fertilizer, was applied per plant. This application was repeated 30 days post-transplantation, maintaining the same volume and concentration of fertilizer. All cultivation and management measures were uniformly maintained in consistency.

The tobacco leaves from the middle section, which exhibited similar growth patterns and uniform leaf color and size, were prelabeled before sampling. At the maturation stage, 400 labeled leaves were sampled from each repetition and allocated to three different tobacco curing rooms to undergo normal curing (NC), excessive curing (EC), and insufficient curing (IC) process treatments, with the specific details outlined in Table 4.

The middle layers of fresh leaf and curing leaf at the end of yellowing (42°C), color fixing (54°C), and dry tendon (68°C) stages were collected in triplicate, with each repetition containing 30 pieces, respectively. Samples were stored at -80°C in an ultra-low temperature refrigerator for further measurement.

4.2 Determination of physicochemical composition

At the end of the flue-curing process, the leaf weight was measured, and the leaf density was subsequently calculated as the ratio of leaf weight to leaf area. The measurement of total sugar and chlorine contents was carried out in strict adherence to the "Tobacco and Tobacco Products - Determination of Water Soluble Sugars - Continuous Flow Method" (YC/T 159-2002) and the "Tobacco and Tobacco Products - Determination of Nicotine - Continuous Flow Method" (YC/T 160-2002), respectively. The determination of total nitrogen, potassium, and chlorine elements was conducted using the "Tobacco and Tobacco Products - Determination of Total Nitrogen - Continuous Flow Method" (YC/T 161-2002), "Tobacco and Tobacco Products - Determination of Potassium - Continuous Flow Method" (YC/T 217-2007), and the "Tobacco and Tobacco Products - Determination of Chlorine - Continuous Flow Method" (YC/T 162-2011), respectively.

4.3 LC-MS/MS analysis

A 150 mg sample was precisely weighed and ground in a 2 mL thick-walled tube using 1 mL of pre-cooled (-20°C) 7:3 methanol: water extraction solution. This mixture was stored at 4°C with periodic vortex mixing (every 10 min, 3 times) followed by overnight incubation to ensure thorough extraction. For LC-MS analysis, supernatants were filtered through 0.22 μm filter membrane and analyzed using a Waters ACQUITY UPLC I-Class Plus coupled with a QTRAP 6500 Plus mass spectrometer. Chromatographic separation was achieved on an HSS T3 column (2.1 mm x 10 cm, 1.8 μm, Waters) with a gradient mobile phase comprising 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B), selected to optimize separation efficiency and peak shape. The elution gradient was set as follows: 0 - 2.00 min, 5% B; 2.00 - 22.00 min, 5% B; 22.00 - 27.00 min, 95% B; 27.00 - 27.10 min, 95% B; 27.10 - 30.00 min, 5% B,at a flow rate of 0.300 ml/min with a column temperature maintained at 40°C (Bian et al., 2023).

For the QTRAP 6500 Plus equipped with ESI Turbo ion spray, ion source parameters were optimized as follows: ion source temperature, 450°C; ion spray voltage (IS), 5500 V (positive mode) and -4500 V (negative mode); ion source gas I (GS1), gas

TABLE 4 Parameters in the curing process of different treatments.

Treatment	Parameter	YS				CFS				DTS	
NC	Dry bulb (°C)	35	38	40	42	45	48	51	54	60	68
	Wet bulb (°C)	34	35	36	36	36	36	37	38	39	40
	Heating rate (°C/h)	1	1	0.5	0.5	0.5	0.5	0.5	0.5	1	1
	Time (h)	6	22	14	14	6	8	6	14	8	28
EC	Dry bulb (°C)	35	38	40	42	45	48	51	54	60	68
	Wet bulb (°C)	34	36	36	36	36	37	38	39	40	41
	Heating rate (°C/h)	1	1	0.5	0.5	0.5	0.5	0.5	0.5	1	1
	Time (h)	6	18	18	18	10	8	6	14	8	33
IC	Dry bulb (°C)	35	38	40	42	45	48	51	54	60	68
	Wet bulb (°C)	34	35	36	36	36	36	37	38	39	40
	Heating rate (°C/h)	1	1	0.5	0.5	0.5	0.5	0.5	0.5	1	1
	Time (h)	6	16	10	14	6	8	6	14	8	28

YS, yellowing stage; CFS, color fixing stage; DTS, dry tendon stage; NC, normal curing; EC, excessive curing; IC, insufficient curing.

II (GS2), and curtain gas (CUR) set to 40, 40, and 20 psi, respectively. Multiple Reaction Monitoring (MRM) methods were configured in MRM mode, encompassing information on MRM transitions, collision energy (CE), DE clustering potential (DP), and retention time of target metabolites.

Metabolites were identified and quantified using Skyline software (version 21.1.0.146) in conjunction with the Beijing Genomics Institute (BGI)-Wide Target-Library database. Subsequent bio-informatics analysis encompassed comprehensive data preprocessing, data quality control, global analysis, and in-depth screening for inter-group differences across comparative groups using the online platforms MetaboAnalyst 6.0 (https://www.metaboanalyst.ca/) (Pang et al., 2024) and BGI (https://biosys.bgi.com/). Differential ex-pressed metabolites (DEMs) were identified by VIP ≥ 1, fold change values of ≥ 1.20 or ≤ 0.80, and p values ≤ 0.05. Hierarchical cluster analysis (HCA) was performed on samples and metabolites, with results presented as heatmaps accompanied by dendrograms. Pearson correlation coefficients (PCC) between samples were calculated using the cor function in R. Both HCA and PCC were performed using the pheatmap R package. Correlation Networks (DSPC) and Random Forest were generated by the online platform of MetaboAnalyst 6.0 (https://www.metaboanalyst.ca/) (Pang et al., 2024). Metabolites were annotated using the KEGG compound database (http://www.kegg.jp/kegg/compound/), and then mapped to the corresponding pathways in the KEGG pathway database (http://www.kegg.jp/kegg/pathway.html) (Kanehisa et al., 2016).

4.4 Data analysis

Data were presented as mean ± standard deviation based on three independent replicate experiments. Statistical significance at $p \leq 0.05$ was determined using One-way ANOVA in SPSS version 18.0.

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

KW: Writing – original draft. XC: Writing – original draft, Investigation. ZC: Resources, Writing – original draft. HW: Investigation, Writing – original draft. FW: Investigation, Writing – original draft. LY: Resources, Writing – original draft. SW: Investigation, Writing – original draft. YY: Investigation, Writing – original draft. YT: Investigation, Writing – original draft. YW: Investigation, Writing – original draft. CL: 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 Science and Technology Project of China Tobacco Company (110202202016), Science and Technology Project of Science and Technology Department of Guizhou Province (QKHZC (2024)YB159 and QKHJC-ZK (2022)YB288); Science and Technology Project of Guizhou Tobacco Company (2022XM17).

Acknowledgments

We thank Shenzhen BGI Co., Ltd for technical support in metabolite detection.

Conflict of interest

Authors ZC, LY, YY were employed by the company China Tobacco Hunan Industrial Co. Ltd.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1473527/full#supplementary-material

References

Abubakar, Y., Young, J., Johnson, W., and Weeks, W. (2000). Changes in moisture and chemical composition of flue-cured tobacco during curing. *Tob. Sci.* 44, 51–58. doi: 10.3381/0082-4623-44.1.51

Aswathy, S., Joe, I. H., Rameshkumar, K., and Ibrahim, J. M. (2024). Insights into the role of RAHB and IHB interactions for the structure-activity relationship of caged xanthone morellic acid: Spectroscopic and DFT exploration, molecular docking, and

molecular dynamics simulation studies as an Antituberculosis agent. *J. Mol. Struct.* 1303, 137429. doi: 10.1016/j.molstruc.2023.137429

Attoe, O. (1946). Leaf-burn of tobacco as influenced by content of potassium, nitrogen, and chlorine. *Agron. J.* 38, 186–196. doi: 10.2134/agronj1946.00021962003800020009x

Bacon, C. W., Wenger, R., and Bullock, J. F. (1952). Chemical changes in tobacco during flue-curing. *Indust. Eng. Chem.* 44, 292–296. doi: 10.1021/ie50506a021

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

Bian, J., Sun, J., Chang, H., Wei, Y., Cong, H., Yao, M., et al. (2023). Profile and potential role of novel metabolite biomarkers, especially indoleacrylic acid, in pathogenesis of neuromyelitis optica spectrum disorders. *Front. Pharmacol.* 14. doi: 10.3389/fphar.2023.1166085

Chen, J., Li, Y., He, X., Jiao, F., Xu, M., Hu, B., et al. (2021). Influences of different curing methods on chemical compositions in different types of tobaccos. *Ind. Crops Prod.* 167, 113534. doi: 10.1016/j.indcrop.2021.113534

Chen, Y., Ren, K., He, X., Gong, J., Hu, X., Su, J., et al. (2019). Dynamic changes in physiological and biochemical properties of flue-cured tobacco of different leaf ages during flue-curing and their effects on yield and quality. *BMC Plant Biol.* 19, 555. doi: 10.1186/s12870-019-2143-x

Demole, E., and Enggist, P. (1975). A chemical study of burley tobacco flavour (*Nicotiana tabacum* L.) VI. Identification and synthesis of four irregular terpenoids related to solanone, including a 'Prenylsolanone'. *Helv. Chim. Acta* 58, 1602–1607. doi: 10.1002/hlca.19750580614

Gao, L., Gao, J. M., Ren, X. H., Gao, Y. B., Huang, G. H., Wang, X., et al. (2023). Effects of poly-γ-glutamic acid (γ-PGA) on metabolites of flue-cured tobacco leaves based on metabolomics analysis. *Russ. J. Plant Physiol.* 70, 140. doi: 10.1134/S1021443723601647

Geng, Z., Yang, H., Gao, H., Xing, L., Hu, X., Yan, T., et al. (2024). Metabolomics reveal the chemical characteristic of cigar tobacco leaves during air-curing process. *J. Biobased. Mater. Bioenergy* 18, 621–633. doi: 10.1166/jbmb.2024.2411

Gong, Y., Li, J., Deng, X., Chen, Y., Chen, S., Huang, H., et al. (2023). Application of starch degrading bacteria from tobacco leaves in improving the flavor of flue-cured tobacco. *Front. Microbiol.* 14. doi: 10.3389/fmicb.2023.1211936

Henry, J. B., Vann, M. C., and Lewis, R. S. (2019). Agronomic practices affecting nicotine concentration in flue-cured tobacco: A review. *Agron. J.* 111, 3067–3075. doi: 10.2134/agronj2019.04.0268

Higashio, Y., and Shoji, T. (2004). Heterocyclic compounds such as pyrrole, pyridines, pyrrolidine, piperidine, indole, imidazol and pyrazines. *Appl. Catal. A Gen.* 260, 251–259. doi: 10.1016/S0926-860X(03)00197-2

Hu, Z., Pan, Z., Yang, L., Wang, K., Yang, P., Xu, Z., et al. (2021). Metabolomics analysis provides new insights into the medicinal value of flavonoids in tobacco leaves. *Mol. Omics* 17, 620–629. doi: 10.1039/D1MO00092F

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, 457–462. doi: 10.1093/nar/gkv1070

Li, J., Ma, Z., Dai, H., Li, H., Qiu, J., and Pang, X. (2024). Application of PLSR in correlating sensory and chemical properties of middle flue-cured tobacco leaves with honey-sweet and burnt flavour. *Heliyon* 10, e29547. doi: 10.1016/j.heliyon.2024.e29547

Li, N., Yu, J., Yang, J., Wang, S., Yu, L., Xu, F., et al. (2023). Metabolomic analysis reveals key metabolites alleviating green spots under exogenous sucrose spraying in air-curing cigar tobacco leaves. *Sci. Rep.* 13, 1311. doi: 10.1038/s41598-023-27968-8

Li, X., Bin, J., Yan, X., Ding, M., and Yang, M. (2022). Application of chromatographic technology to determine aromatic substances in tobacco during natural fermentation: A review. *Separations* 9, 187. doi: 10.3390/separations9080187

Li, Y., Wang, Y., Ma, W., and Tan, C. (2001). Breeding and selecting of a new flue-cured tobacco variety Yunyan87 and its characteristics. *Chin. Tobacco. Sci.* 22, 38–42. doi: 10.3969/j.issn.1007-5119.2001.04.003

Lim, H. H., Choi, K. Y., and Shin, H. S. (2022). Flavor components in tobacco capsules identified through non-targeted quantitative analysis. *J. Mass. Spectrom.* 57, e4811. doi: 10.1002/jms.4811

Liu, J., Wang, J., Du, Y., Yan, N., Han, X., Zhang, J., et al. (2024). Application and evaluation of the antifungal activities of glandular trichome secretions from air/sun-cured tobacco germplasms against *botrytis cinerea*. *Plants* 13, 1997. doi: 10.3390/plants13141997

Mavroeidis, A., Stavropoulos, P., Papadopoulos, G., Tsela, A., Roussis, I., and Kakabouki, I. (2024). Alternative crops for the European tobacco industry: A systematic review. *Plants* 13, 236. doi: 10.3390/plants13020236

Meng, L., Song, W., Chen, S., Hu, F., Pang, B., Cheng, J., et al. (2022). Widely targeted metabolomics analysis reveals the mechanism of quality improvement of flue-cured tobacco. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1074029

Meng, Y., Wang, Y., Guo, W., Lei, K., Chen, Z., Xu, H., et al. (2024). Analysis of the relationship between color and natural pigments of tobacco leaves during curing. *Sci. Rep.* 14, 166. doi: 10.1038/s41598-023-50801-1

Pang, Z., Lu, Y., Zhou, G., Hui, F., Xu, L., Viau, C., et al. (2024). MetaboAnalyst 6.0: Towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Res.* 52, gkae253. doi: 10.1093/nar/gkae253

Rout, K. K., Kar, M. K., Agarwal, P. C., and Dash, S. K. (2024). Analysis of bioactive hispidulin: an anticancer flavone of *Clerodendrum philippinum*. J. *Planar. Chromat.* 37, 49–56. doi: 10.1007/s00764-023-00267-8

Shoji, T., Hashimoto, T., and Saito, K. (2024). Genetic regulation and manipulation of nicotine biosynthesis in tobacco: strategies to eliminate addictive alkaloids. *J. Exp. Bot.* 75, 1741–1753. doi: 10.1093/jxb/erad341

Sun, C. L., Zhang, H. L., Zhou, D. B., Cheng, Z. J., Xie, Y., Rang, Z. W., et al. (2023). Based on metabolomics, the optimum wind speed process parameters of flue-cured tobacco in heat pump bulk curing barn were explored. *Sci. Rep.* 13, 21558. doi: 10.1038/s41598-023-49020-5

Vaughan, C., Stanfill, S. B., Polzin, G. M., Ashley, D. L., and Watson, C. H. (2008). Automated determination of seven phenolic compounds in mainstream tobacco smoke. *Nicotine. Tob. Res.* 10, 1261–1268. doi: 10.1080/14622200802123146

Wang, J., Liu, S. M., Long, J., Lei, D. A., and Gao, F. (2020). Derivatization method for the determination of amino acids in tobacco by gas chromatography-mass spectrometry. *J. Anal. Chem.* 75, 1046–1053. doi: 10.1134/S1061934820080171

Weng, S., Deng, M., Chen, S., Yang, R., Li, J., Zhao, X., et al. (2024). Application of pectin hydrolyzing bacteria in tobacco to improve flue-cured tobacco quality. *Front. Bioeng. Biotechnol.* 12. doi: 10.3389/fbioe.2024.1340160

Xin, X., Gong, H., Hu, R., Ding, X., Pang, S., and Che, Y. (2023). Intelligent large-scale flue-cured tobacco grading based on deep densely convolutional network. *Sci. Rep.* 13, 11119. doi: 10.1038/s41598-023-38334-z

Yang, H., Sun, G., Yin, G., Sun, H., Wang, T., Bai, T., et al. (2023). Browning mechanism of tobacco leaves during flue-curing process: Proteomics and metabolomics analysis reveals the changes in materials. *Mater. Express.* 13, 1068–1080. doi: 10.1166/mex.2023.2443

Zhao, Y., Zhao, J., Zhao, C., Zhou, H., Li, Y., Zhang, J., et al. (2015). A metabolomics study delineating geographical location-associated primary metabolic changes in the leaves of growing tobacco plants by GC-MS and CE-MS. *Sci. Rep.* 5, 16346. doi: 10.1038/srep16346

Zong, J., He, X., Lin, Z., Hu, M., Xu, A., Chen, Y., et al. (2022). Effect of two drying methods on chemical transformations in flue-cured tobacco. *Drying. Technol.* 40, 188–196. doi: 10.1080/07373937.2020.1779287

Zou, X., Bk, A., Rauf, A., Saeed, M., Al-Awthan, Y. S., Al-Duais, M. A., et al. (2021). Screening of polyphenols in tobacco (*Nicotiana tabacum*) and determination of their antioxidant activity in different tobacco varieties. *ACS Omega.* 6, 25361–25371. doi: 10.1021/acsomega.1c03275

Zou, C., Hu, X., Huang, W., Zhao, G., Yang, X., Jin, Y., et al. (2019). Different yellowing degrees and the industrial utilization of flue-cured tobacco leaves. *Sci. Agric.* 76, 1–9. doi: 10.1590/1678-992x-2017-0157

Zou, L., Su, J., Xu, T., Ji, X., Wang, T., Chen, Y., et al. (2023). Untargeted metabolomics revealing changes in aroma substances in flue-cured tobacco. *Open Chem.* 21, 20220326. doi: 10.1515/chem-2022-0326

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 28 November 2024
DOI 10.3389/fpls.2024.1452804

OPEN ACCESS

EDITED BY
Moonhyuk Kwon,
Gyeongsang National University, Republic of Korea

REVIEWED BY
Juan De Dios Franco-Navarro,
Spanish National Research Council (CSIC), Spain
Jongmin Ahn,
Korea Research Institute of Bioscience and Biotechnology (KRIBB), Republic of Korea

*CORRESPONDENCE
Biljana Filipović
biljana.nikolic@ibiss.bg.ac.rs
Marijana Skorić
mdevic@ibiss.bg.ac.rs
Uroš Gašić
uros.gasic@ibiss.bg.ac.rs
Danijela Mišić
dmisic@ibiss.bg.ac.rs

RECEIVED 21 June 2024
ACCEPTED 11 November 2024
PUBLISHED 28 November 2024

CITATION
Petrović L, Filipović B, Skorić M, Šiler B, Banjanac T, Matekalo D, Nestorović Živković J, Dmitrović S, Aničić N, Milutinović M, Božunović J, Gašić U and Mišić D (2024) Molecular background of the diverse metabolic profiles in leaves and inflorescences of naked catmint (*Nepeta nuda* L.).
Front. Plant Sci. 15:1452804.
doi: 10.3389/fpls.2024.1452804

Molecular background of the diverse metabolic profiles in leaves and inflorescences of naked catmint (*Nepeta nuda* L.)

Luka Petrović, Biljana Filipović*, Marijana Skorić*, Branislav Šiler, Tijana Banjanac, Dragana Matekalo, Jasmina Nestorović Živković, Slavica Dmitrović, Neda Aničić, Milica Milutinović, Jelena Božunović, Uroš Gašić* and Danijela Mišić*

Department of Plant Physiology, Institute for Biological Research "Siniša Stanković" - National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia

Nepeta nuda L. shares a typical secondary chemistry with other *Nepeta* species (fam. *Lamiaceae*), characterized by the tendency to intensively produce monoterpenoid iridoids, whereas the phenylpropanoid chemistry is steered towards the production of a caffeic acid ester, rosmarinic acid. Combining complementary state-of-the-art analytical techniques, *N. nuda* metabolome was here comprehensively characterized in the quest for the organ-specific composition of phenolics and terpenoids that possess well-defined functions in plant-biotic interactions as well as therapeutic potential. *N. nuda* inflorescences showed generally higher constitutive levels of specialized metabolites, as compared to leaves, and the composition of major iridoids and phenolics in reproductive organs was found to be more conserved than in leaves across 13 populations from the Central Balkans. The results suggest that *N. nuda* plants most likely invest more in constitutive than inducible biosynthesis of functional metabolites in flowers, since they are of essential importance for both pollination and defense against herbivores and pathogens. Conversely, specialized metabolism of leaves is found to be more susceptible to reprograming in response to differential growth conditions. The defense strategy of leaves, primarily functioning in CO_2 fixation during photosynthesis, more likely relies on the induction of metabolite levels following plant-environment interplay. Organ-specific biosynthesis of iridoids in *N. nuda* is found to be tightly regulated at the transcriptional level, and high constitutive levels of these compounds in inflorescences most likely result from the up-regulated expression of several key genes (*NnG8H*, *NnNEPS1*, *NnNEPS2*, and *NnNEPS3*) determining the metabolic flux through the pathway. The organ-specific content of rosmarinic acid and co-expression patterns of the corresponding biosynthetic genes were much less correlated, which suggests independent organ-specific transcriptional regulation of the iridoid and phenolic pathways. Knowledge gathered within

the present study can assist growers to select productive genotypes and manipulate phenology of *N. nuda* towards maximizing yields and facilitating its integration into pest management systems and other applications related to human health.

KEYWORDS

metabolomics, UHPLC-ESI-QToF-MS, UHPLC/DAD/(±)HESI-MS², GC/MS, phenolics, terpenes, iridoids, populations

1 Introduction

Plants produce a diverse array of specialized metabolites that have evolved specific physiological and ecological functions. These metabolites occur either as volatile or non-volatile compounds and play important roles in plant's adaptation to different environmental conditions and defense against herbivores and microorganisms, as well as in processes such as pollination and seed dispersal (Wink, 2018). Phytochemical diversity varies across different levels of biological organization, i.e., among populations, among individuals within a population, and within an individual plant (among plant organs, during development, and across seasons) (Moore et al., 2014; Wetzel and Whitehead, 2020).

Variation in several categories of specialized metabolites is well studied at the interpopulation level of numerous species, including those belonging to the genus *Nepeta*. The interpopulation chemical polymorphism detected in *Nepeta sessilifolia*, *N. heliotropifolia*, and *N. fissa* was reported to be correlated with the elevation (Talebi et al., 2019). It has been suggested that the amounts of different types of monoterpenes decreased and the amounts of oxygenated compounds increased along the altitudinal gradient. Environmental conditions are highlighted as important factors influencing the yield and the chemical composition of essential oils of *N. fissa* (Talebi et al., 2017) and *N. heliotropifolia* (Yarmoohammadi et al., 2017). In another study, 29 accessions of *N. kotschyi*, *N. menthoides*, *N. crassifolia*, and *N. cataria* were cultivated under western Tehran (Iran) environmental conditions and evaluated for their phenolics composition (Hadi et al., 2017). Diverse composition of specialized metabolites in various plant tissues and organs is an important component of the overall plant phenotype, and changes in phytochemical profiles are one of the fastest phenotypic responses to the dynamic environmental conditions (Junker, 2016). These plant-part-specific differences may also be related to the ecological functions of the vegetative and reproductive tissues.

Nepeta nuda L. (subfam. Nepetoidae, fam. Lamiaceae) is an herbaceous perennial species widespread across Europe and Asia. Its individuals, characterized by numerous erect stems, oblong-lanceolate lower leaves, and ovate upper leaves can grow 50 to 100 cm tall (Aćimović et al., 2020). Flowers are organized in lax or dense spike-like verticillasters that bloom from June to August, depending on the altitude (Aćimović et al., 2020). Health benefits of *N. nuda* are usually associated with the accumulation of terpenoids and phenolics, two major groups of specialized metabolites of the genus *Nepeta*. Terpenoids, predominately represented by iridoid compounds, are produced in most plant tissues where they primarily act either as pollinator attractants (Schultz et al., 2004) or insect repellents (Bernier et al., 2005; Birkett et al., 2011; Sparks et al., 2018; Reichert et al., 2019; reviewed in Formisano et al., 2011 and Süntar et al., 2018), but can also function as insect pheromones (Glinwood et al., 1999; Birkett and Pickett, 2003) and are appealing to felines (Uenoyama et al., 2021 and 2022). Numerous studies on the chemical composition of *N. nuda* plant extracts and essential oils revealed a high level of chemodiversity, which may be ecologically driven, as it has been proven to be strongly influenced by various environmental factors, such as geographical origin, light conditions or altitude (Sharma et al., 2021; Petrović et al., 2024). Only few studies report on variations in iridoid and phenolic composition among populations of *N. nuda* within a specific geographical region (Narimani et al., 2017; Petrova et al., 2022; Petrović et al., 2024). In a recent study, we reported low differences in the iridoid and phenolic profiles in leaves among populations in the Central Balkans, since the majority of chemical and genetic variations were found within populations (Petrović et al., 2024). Moreover, other recent studies suggested organ-specific biosynthesis and accumulation of bioactive metabolites in *N. nuda* (Petrova et al., 2022; Zaharieva et al., 2023). Thus, the content of rosmarinic acid and 1,5,9-*epi*-deoxyloganic acid in *N. nuda* plants from Mt. Pirin (Bulgaria) was significantly higher in flowers than in leaves (Petrova et al., 2022).

The objective of this study was to comparatively investigate the patterns of metabolites' variation in *N. nuda* vegetative (leaves) and reproductive (inflorescences) organs across 13 natural populations within the Central Balkans, adopting both targeted and untargeted metabolomics approaches. The determination of variation patterns in the plant metabolome in nature is crucial for predicting the persistence of populations and species in the future, understanding the species' ecological interactions, and conserving their genetic diversity (Labarrere et al., 2019). Characterization of intraspecies chemodiversity in *N. nuda* within the Central Balkans can indirectly enable mapping of areas suitable for cultivation and intendedly

assist farmers and other stakeholders to identify accessions with the highest crop potential and earmark them for future cultivation. *N. nuda* can be considered attractive medicinal crop for montane and subalpine regions (up to 2100 m a.s.l.), due to high adaptive success and production of essential oils with prominent bioactive properties (e.g., Gkinis et al., 2010; Formisano et al., 2011; Musso et al., 2017; reviewed in Sharma et al., 2021; Zaharieva et al., 2023). Considering differential but equally important ecophysiological roles of leaves and inflorescences, we intended to investigate which plant part is more susceptible to metabolic reprogramming in response to external influences and discuss the potential ecological implications arising from variations in the composition of terpenes and phenolics in these two organs of *N. nuda*. We further delved into the molecular background of the organ-specific differences in the biosynthesis and accumulation of iridoids and phenolics, thus providing evidence for their mutually independent transcriptional regulation of metabolic fluxes in different organs.

2 Materials and methods

2.1 Chemicals and reagents

Acetonitrile (Fisher Scientific UK, Leicestershire, UK) and formic acid (Merck, Darmstadt, Germany) were of MS grade. Ultra-pure deionized water was generated using a Water Purification System (New Human Power I Integrate, Human Corporation, Republic of Korea). Standards of 1,5,9-*epi*-deoxyloganic acid and 5,9-dehydronepetalactone were isolated from natural sources as previously described by Aničić et al. (2021). The standard of *cis,trans*-nepetalactone was a generous gift from Entomol Products LLC (San Francisco, CA, USA). Analytical standards of loganin, rosmarinic acid, caffeic acid, luteolin and apigenin were purchased from Sigma-Aldrich (Hamburg, Germany).

2.2 Collection of plant material

Aboveground parts of flowering *N. nuda* plants were collected in June-August 2022 from 13 populations across Serbia: 1. Mali Ljukten (43°32'35"N; 20°48'38"E); 2. Brodica (44°29'04"N; 21°50'27"E); 3. Debeli lug (44°21'45"N; 21°54'01"E); 4. Rtanj (43°44'13"N; 21° 57'03"E); 5. Straža (43°50'46"N; 21°42'03"E); 6. Židilje (44°00'42"N; 21°38'45"E); 7. Vinatovača (44°04'14"N; 21°45'36"E); 8. Vlasina (42° 41'39"N; 22°22'44"E); 9. Donji Krivodol (43°06'18"N; 22°55'41"E); 10. Topli Do (43°22'05"N; 22°37'58"E); 11. Janjska reka (43°25'33"N; 22° 31'11"E); 12. Balta Berilovac (43°24'10"N; 22°30'43"E); 13. Gornje selo (42°11'21"N; 20°56'20"E). Plants were identified in the field by the authors, and representative specimens, assigned with the voucher numbers, have been deposited in the Herbarium of the Institute of Botany and Botanical Garden, University of Belgrade, BEOU (acronym follows Thiers, 2023). Plant material was dried in the shade at room temperature until constant mass and subsequently stored in paper bags. Samples were kept in the dark at room temperature until use. *N. nuda* aboveground parts were detached to leaves and inflorescences.

2.3 Preparation of *N. nuda* methanol extracts

Samples containing either leaves or inflorescences (50 mg) of *N. nuda* were ground in liquid nitrogen and extracted overnight at 4 °C with 1 mL of 96% methanol (w:v = 1:20). After vortexing for 1 min, samples were sonicated in an ultrasonic bath (RK100, Bandelin, Berlin, Germany) for 1 h at 4°C. Supernatants were separated after centrifugation at 10,000g for 10 min and subsequently filtered through 0.2 μm-pore size cellulose filters (Agilent Technologies, Santa Clara, CA, USA). Each population was represented by 3 biological replicates.

2.4 Characterization of metabolites in *N. nuda* leaves and inflorescences using UHPLC-ESI-QToF-MS analysis

Three randomly selected *N. nuda* samples (populations 3, 8, and 9) of both inflorescences and leaves were subjected to metabolic fingerprinting using an Agilent 1290 Infinity UHPLC system in combination with a quadrupole time-of-flight mass spectrometer (6530C QToF-MS, Agilent Technologies, Inc., Santa Clara, CA, USA). The mass detector, equipped with an electrospray ionization (ESI) source, was operated in both positive and negative ionization modes, in the mass range from 100 to 1000 *m/z*. Samples were chromatographically separated on a Zorbax C18 column (2.1 × 50 mm, particle size 1.8 μm; Agilent Technologies, Inc., Santa Clara, CA, USA) thermostated at 40°C. The gradient elution program, UHPLC and MS parameters as well as ion source settings were described by Petrović et al. (2024).

The Agilent MassHunter software was used for acquisition, control, and MS data collection from the 6530C Q-ToF-MS instrument. The R Studio platform ('enviPick' and 'xcms' R packages) was used for the evaluation and presentation of MS data (Zengin et al., 2021). The metabolites were identified based on their monoisotopic masses and MS^2 fragmentation and confirmed using literature data (Alimpić Aradski et al., 2023; Aničić et al., 2021; Dienaitė et al., 2018; Goldansaz et al., 2019). Accurate masses of components and fragment ions were calculated using the ChemDraw software (version 12.0, CambridgeSoft, Cambridge, MA, USA). The CAS SciFinder-n database (https://scifinder-n.cas.org/) was used to search for chemical compounds by formulae and structures.

2.5 UHPLC/DAD/(±)HESI-MS^2 targeted metabolic profiling of *N. nuda* leaves and inflorescences

A targeted UHPLC/DAD/(±)HESI/MS^2 metabolomics approach was adopted to quantify individual phenolics (rosmarinic acid, caffeic

acid, apigenin, and luteolin) and iridoids (*cis,trans*- and *trans,trans*-nepetalactone, 5,9-dehydronepetalactone, nepetanudoside A, and 1,5,9-*epi*-deoxyloganic acid) in *N. nuda* methanol extracts of leaves and inflorescences. Analyses were performed on a Dionex UltiMate 3000 UHPLC system coupled with a DAD detector and configured with a triple quadrupole mass spectrometer (TSQ Quantum Access Max, Thermo Fisher Scientific, Basel, Switzerland). Samples were chromatographically separated on a Hypersil gold C18 analytical column (50 × 2.1 mm, 1.9 µm particle size; Thermo Fisher Scientific, Waltham, MA, USA), using the elution gradient and the flow rate previously described by Aničić et al. (2021). The mobile phase consisted of (A) water + 0.1% formic acid and (B) acetonitrile + 0.1% formic acid. The injection volume was 10 µL. The selected reaction monitoring (SRM) mode of the instrument was used for the quantification of the targeted compounds by direct comparison with the commercial standards. Nepetanudoside A was quantified based on the calibration curve of loganin, while quantification of aucubin was performed using the calibration curve for nepetaside. Calibration curves of pure standards revealed good linearity, with r^2 values exceeding 0.99 (peak areas vs. concentration). The total amount of each phenolic and iridoid compound was evaluated by the calculation of its peak area and is expressed as µg per 100 mg leaf/inflorescence DW.

2.6 GC/MS non-targeted metabolomics of methanol extracts of *N. nuda* leaves and inflorescences

Profiling of volatile compounds in methanol extracts of *N. nuda* leaves and inflorescences was performed using an Agilent 8890 gas chromatography (GC) system coupled with a Mass Selective Detector (5977B GC/MSD, Agilent Technologies, Santa Clara, CA, USA) and connected to an automated sample extraction and enrichment platform (Centri®, Markes International Ltd., Bridgend, UK). Chromatographic separation procedure is described in detail by Petrović et al. (2024). The constituents of the methanol extracts were identified by comparison of their mass spectra and retention times with those of the respective standards as well as by comparison with the NIST05 library.

2.7 Statistical analyses of metabolomics data

For hierarchical cluster analysis (HCA), the input variables were scaled to the [min, max] range. HCA was performed based on the Spearman's method of cluster agglomeration, using the Morpheus software (https://software.broadinstitute.org/morpheus). Principal component analysis (PCA) was constructed using the Past 4 software (version 4.14; Hammer and Harper, 2001). For the comparison between populations, quantitative metabolomics data were subjected to the Tukey's *post hoc* test ($p < 0.05$) of one-way ANOVA. To test the statistical difference between leaves and inflorescences, *t* test was adopted.

2.8 Comparative metabolic profiling and gene expression analysis in leaves and inflorescences of greenhouse-grown *N. nuda*

2.8.1 Plant material and metabolic profiling of iridoids and phenolics

Seeds of *N. nuda*, collected in June 2022 in the locality Debeli Lug (East Serbia) were germinated under *in vitro* conditions, as previously described (Aničić et al., 2024). Following clonal propagation of a single genotype under *in vitro* conditions, adopting previously optimized protocols for other *Nepeta* species (Aničić et al., 2024), plantlets were acclimatized in a greenhouse of the Institute for Biological Research "Siniša Stanković" - National Institute of the Republic of Serbia, University of Belgrade, Serbia. Samples of leaves and inflorescences were collected from fully-flowering plants in their third year of growth under greenhouse conditions.

Leaves and inflorescences were excised from plants and immediately frozen in liquid nitrogen (LN). Samples were mechanically grinded in LN until fine powder, weighted and aliquoted into batches of 100 mg, and subsequently stored at −80° C until use. The first group of batches containing three biological replicates, designated for metabolomics, was dried by lyophilization until constant mass, weighted, and extracted with 96% methanol (100:1 = w:v). Following vigorous vortexing for 1 min, the extraction was continued in an ultrasonic bath for 1 h, and supernatants were separated by centrifugation for 10 min at 10,000*g*. Supernatants were filtered through 0.2-mm cellulose filters (Agilent Technologies, Santa Clara, CA, USA) and stored at 4°C until use. Quantification of the targeted iridoids and phenolics in samples was performed using a UHPLC-DAD-(±)HESI/MS^2 instrument as described in section 2.5.

2.8.2 RNA extraction and qPCR assessment of the expression of biosynthetic genes

The second batch of samples was used to extract RNA from *N. nuda* leaves and inflorescences, and all the extractions were performed in three biological replicates. The RNA extraction was performed applying a modified CTAB method (Gasic et al., 2004), which was followed by the quantification using a N60 Nano-Photometer® (Implen GmbH, Munich, Germany) and a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), while the RNA integrity was electrophoretically checked. The procedure of the DNase I (Termo Fisher Scientific, USA) treatment and the protocol for cDNA synthesis were described earlier by Aničić et al. (2024).

Gene expression analyses were performed by real-time PCR using QuantStudio™ 3 Real-Time PCR System (Thermo Fisher Scientific, Carlsbad, CA, USA). Thermocycler conditions were previously described by Aničić et al. (2018). The reactions were performed using Maxima SYBR Green/ROX Master Mix (2X) (Termo Fisher Scientific, USA), cDNA corresponding to 50 ng RNA, and 0.3 µM primers, according to the manufacturer's recommendations. The expression levels of targeted candidates of iridoid- and phenolic-biosynthetic genes were calculated according

to the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) using GAPDH as a housekeeping gene. Primer pairs were designed based on candidate genes whose sequences and accession numbers are listed in Supplementary Table 1.

2.8.3 Statistical analysis of gene expression data

qPCR data were processed and visualized using the Past 4 software (version 4.16; Hammer and Harper, 2001). To test the statistical difference between leaves and inflorescences, *t* test was adopted. Correlation analysis, adopting the Ward's algorithm, was performed to examine the co-expression patterns of targeted genes in leaves and inflorescences of *N. nuda*.

Linear regression analysis was performed to visualize relations between metabolomics and gene-expression data. Coefficients of determination (R^2) were calculated using the Past 4 software (version 4.16; Hammer and Harper, 2001).

3 Results and discussion

In this study, we adopted two complementary analytical techniques to cover different parts of *N. nuda* metabolome. A UHPLC-ESI-QToF-MS instrument, operating in both positive and negative ionization modes, enabled the chromatographic separation and identification of polar molecules in methanol extracts, primarily phenolics. It enabled us to comparatively and simultaneously analyze both glycosides and aglycones of iridoids in leaves and inflorescences. The solvent of choice for the present study was methanol as being vastly used for the extraction of both phenolics (e.g., Mišić et al., 2015; Hadi et al., 2017; Dienaitė et al., 2018; Aničić et al., 2021) and terpenoids (Aničić et al., 2024), including iridoid glycosides and volatile iridoid aglycones found in *Nepeta* species (Takeda et al., 1995; Petrova et al., 2022; Zaharieva et al., 2023).

3.1 Comparison of qualitative non-targeted phenolic profiles in *N. nuda* leaves and inflorescences

Totally 38 phenolic compounds were identified, including derivatives of hydroxybenzoic and hydroxycinnamic acids as well as flavonoid glycosides and aglycones (Supplementary Table 2).

Nine derivatives of hydroxybenzoic acid were identified mainly as hexosides (Supplementary Table 2), giving a specific loss of 162 Da in MS^2 spectra. All hydroxybenzoic acid derivatives were detected in leaves, while only one compound, dihydroxybenzoic acid (**3**), was not identified in the tested inflorescences (Supplementary Table 2). The majority of compounds belonging to the group of hydroxybenzoic acids were more abundant in inflorescences than in leaves of *N. nuda* except galloyl hexoside (**1**) and dihydroxybenzoic acid hexoside (**4**) (Figure 1A).

All nineteen compounds from the group of hydroxycinnamic acids were detected in inflorescences, whereas compounds **13**, **24**, **27**, and **28** were not found in leaves (Supplementary Table 2). Interestingly, esters of caffeic acid (**23**), malic acid (**25**), tartaric acid (**11** and **21**), tartronic acid (**12** and **16**), and glycolic acid (**14**) were more abundant in leaves than in inflorescences. The majority of other identified hydroxycinnamic acids were more abundant in inflorescences (Figure 1A). Compound **24** was not previously detected in any *Nepeta* species, but its MS^2 spectra showed fragments typical for rosmarinic acid, which is characteristic for some *Nepeta* taxa (Mišić et al., 2015). This compound showed a neutral loss of 162 Da leading to the formation of a fragment ion at 359 *m/z* that corresponds to the exact mass of deprotonated rosmarinic acid. By studying other MS^2 fragments, this compound was marked as rosmarinic acid hexoside.

All identified flavonoid compounds belong to the subgroup of flavones, which is typical for the genus *Nepeta* (Jamzad et al., 2003). The identified flavonoid aglycones were generally more abundant in inflorescences than in leaves. Hexuronides of luteolin (4 compounds) and apigenin (1 compound) were also recorded in the analyzed samples of *N. nuda* (Supplementary Table 2). Inflorescences were especially rich in these compounds. Flavones (mainly apigenin and luteolin derivatives) glycosylated with hexuronic (glucuronic) acid, often with another acyl residue (acetyl, coumaroyl, caffeoyl or similar), have been previously reported as constituents of several *Nepeta* species (Rabee et al., 2020). Compound **59** was identified as luteolin 7-*O*-(2"-caffeoyl)-hexuronide. In its MS^2 spectra, ions were detected at 443 and 337 *m/z*, formed by the neutral loss of caffeic acid and luteolin, respectively (Supplementary Figure 1). This compound showed an unusual fragmentation pattern as its MS^2 base peak was not the mass of luteolin aglycone (285 *m/z*) but corresponded to the mass of caffeic acid-H-H_2O and it was previously identified in *Satureja biflora* aerial parts (Moghadam et al., 2015).

In the present study, we adopted the combination of analytical techniques previously reported to be convenient for establishing chemical characterization of *N. nuda* populations from the Central Balkans (Petrović et al., 2024). The mentioned study was focused on the chemical profiling of leaves dried in silica gel immediately after harvesting, while the samples used in this study were air-dried at room temperature until they reached a constant mass. This enabled us to compare and discuss the phytochemical composition of methanol extracts as influenced by drying procedure. Thus, vanillic acid was recorded only in silica gel-dried samples of *N. nuda* leaves (Petrović et al., 2024). Within the group of hydroxycinnamic acids, caffeic acid hexoside (**13**), ferulic acid (**27**), and nepetoidin A or B (**28**) were not present in air-dried leaves but were detected in inflorescences. The qualitative composition of flavonoids was also slightly perturbed, as luteolin (**61**) was recorded only in air-dried leaves, while chrysoeriol and acacetin were identified in leaves dried on silica-gel (Petrović et al., 2024).

3.2 Comparison of qualitative non-targeted iridoid profiles of *N. nuda* leaves and inflorescences

Twenty-seven iridoid compounds were identified in *N. nuda* inflorescences and leaves, 11 out of them being glycosides and 16 aglycones (Supplementary Table 2). One nitrogen analogue of nepetalactone, nepetalactam, was also recorded in *N. nuda* inflorescences. The presence of iridoid compounds was confirmed

10.3389/fpls.2024.1452804

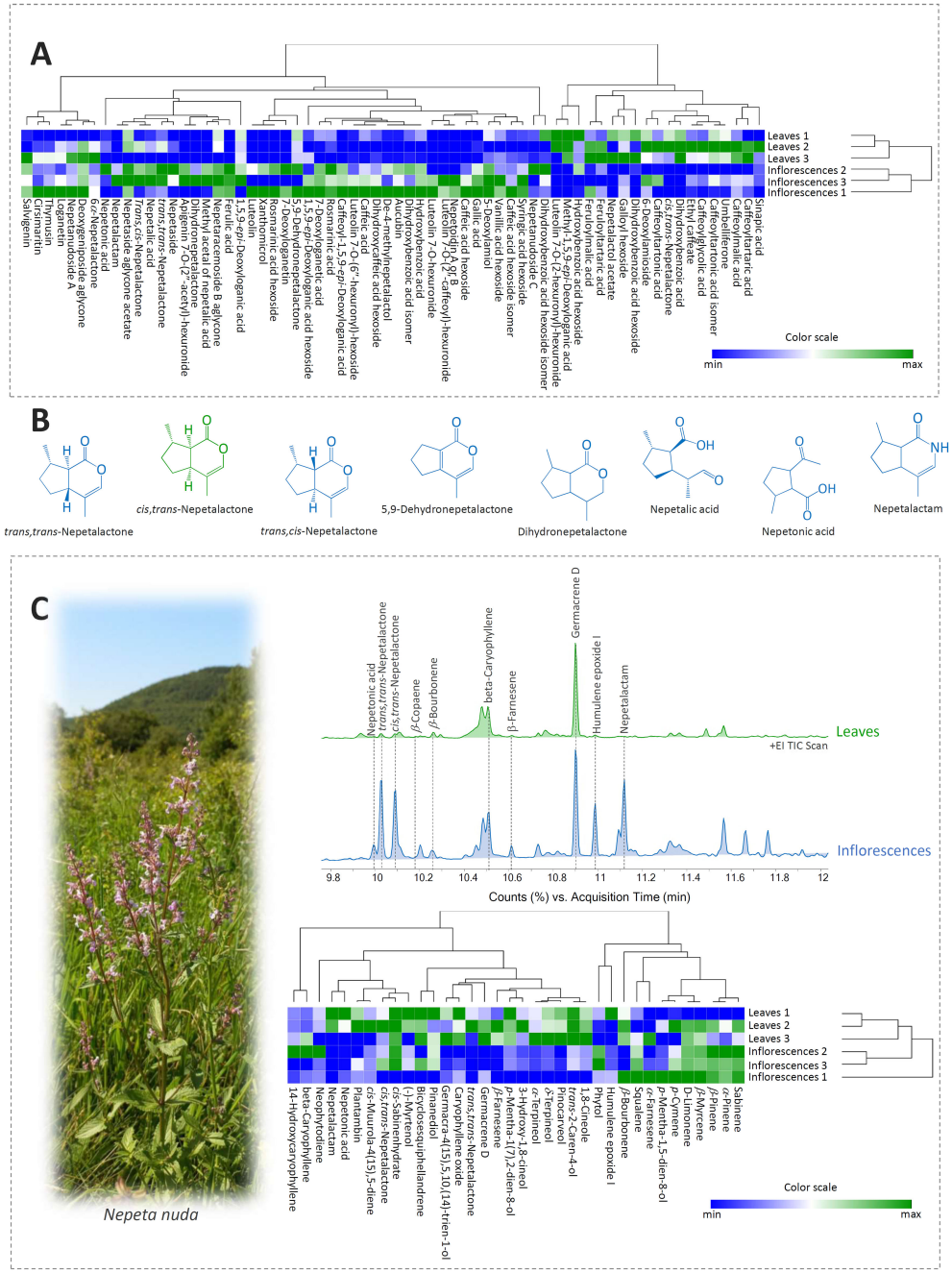


FIGURE 1
(A) Heatmap of the scaled QToF-MS data with the samples and compounds arranged according to the hierarchical cluster analysis (HCA), adopting the Spearman's method of cluster agglomeration. **(B)** Chemical structures of the characteristic iridoid aglycones identified in *Nepeta nuda* methanol extracts of leaves and inflorescences. Iridoids predominating in inflorescences are represented in blue, while those prevalent in leaves are presented in green. **(C)** Representative TIC GC/MS chromatograms of methanol extracts of leaves and inflorescences. HCA plot of the scaled GC/MS data with populations and compounds arranged according to the Spearman's method of cluster agglomeration. Plant photo denotes *N. nuda* individual captured on site (locality Straža, SE Serbia).

by the comparison with the available literature on *Nepeta* species and through the evaluation of their MS data. Some of the major iridoids identified in leaves and inflorescences of *N. nuda* are presented within Figure 1B.

Iridoid aglycones were mostly visible in the positive ionization mode due to their polarity, except for 7-deoxyloganetin, nepetalic and nepetonic acids (compounds **47**, **48**, and **50**, respectively), which gave more abundant peaks in the negative ionization mode. Among the identified iridoid aglycones, 7-deoxyloganetin (**47**) and dihydronepetalactone (**51**) were recorded in inflorescences only, while *cis*,*trans*-nepetalactone (**52**) and nepetalactol acetate (**53**) were characteristic for *N. nuda* leaves.

Nepetalactones and their derivatives were the most abundant iridoid aglycones in the analyzed *N. nuda* samples. This iridoid

monoterpenoid can usually be found in the form of four diastereoisomers, three of them, *trans*,*trans*-, *cis*,*trans*-, and *trans*, *cis*-nepetalactone, being identified in the present study. Previous studies dealing with the phytochemical characterization of *N. nuda* reported all 4 diastereoisomers: *trans*,*trans*- (De Pooter et al., 1988; Handjieva et al., 1996; Petrović et al., 2024), *cis*,*trans*- (De Pooter et al., 1988; Kökdil et al., 1999; Kobaisy et al., 2005; Mancini et al., 2009; Narimani et al., 2017; Akbaba et al., 2021; Petrović et al., 2024), *trans*,*cis*- (Regnier et al., 1967; De Pooter et al., 1988; Mancini et al., 2009; Gkinis et al., 2010; Bozok et al., 2017; Narimani et al., 2017; Akbaba et al., 2021; Aćimović et al., 2022), and *cis*,*cis*-nepetalactone (De Pooter et al., 1988; Kökdil et al., 1999; Kobaisy et al., 2005; Mancini et al., 2009; Bozari et al., 2013; Gormez et al., 2013; Zaharieva et al., 2023). In our study, *trans*,*trans*-nepetalactone was more abundant in inflorescences than in leaves (Figure 1A), while *cis*,*trans*-nepetalactone was recorded in leaves only. *trans*,*cis*-Nepetalactone was recorded in trace amounts in both leaves and inflorescences (Supplementary Table 2).

5,9-Dehydronepetalactone (**40**) is a dehydrogenated nepetalactone, isolated for the first time from *N. cataria* (Sastry et al., 1972). In *N. nuda*, this compound is primarily accumulated in inflorescences, while leaves contain considerably lower amounts (Figure 1A). This compound was previously reported in *N. nuda* (De Pooter et al., 1988; Kökdil et al., 1999; Petrović et al., 2024) and it was quantified in significant amounts in methanol extracts of *N. rtanjensis*, *N. argolica*, and *N. parnassica* (Mišić et al., 2015; Aničić et al., 2018, 2020, 2021). The content of 5,9-dehydronepetalactone in *N. teydea* EO was higher when it was prepared from flowering plants than from plants in the vegetative growth stage (Velasco-Negueruela et al., 1989), which is consistent with the findings of the present study.

Dihydronepetalactone (**51**) was identified in *N. nuda* inflorescences only (Supplementary Table 2). It is formed by nepetalactone hydrogenation (Feaster et al., 2009) and is known to be a common constituent of *N. cataria* (Regnier et al., 1967; Sharma and Cannoo, 2013; Sengupta et al., 2018; Handjeva et al., 1996; Patel et al., 2022).

Nepetalic acid (syn. 3α-hydroxy-4aα,7α,7aα-dihydronepetalactone, 3R-hydroxy-4aR,7R,7aR-dihydronepetalactone) can be synthetically converted into nepetalactone (Chauhan and Zhang, 2008), which marks this compound and plants rich in it as potentially valuable industrial resources. *N. nuda* is a good candidate as it accumulates nepetalic acid (**48**) in high amounts, especially in inflorescences (Supplementary Table 2, Figure 1A).

A derivative of nepetalic acid, its methyl isomer, nepetonic acid (2,-[(1-methyl-2-al)-ethyl-5R-methylcyclopenta,-carboxylic acid) (**50**), was identified in *N. nuda* samples in the present study (Supplementary Table 2). Methyl acetal of nepetalic acid (**54**) can be synthesized by nepetalic acid acylation (Scialdone and Liauw, 2008) and it was identified exclusively in inflorescences of *N. nuda*, adopting the UHPLC/QToF-MS technique (Supplementary Table 2).

Nepetalactam (**81**), classified under the group of tetrahydropyridine metabolites (lactams or cyclic amides), is a nitrogen analogue of (4aS,7S,7aR; *cis*,*trans*-)-nepetalactone previously identified in the essential oil of *N. cataria* (Handjeva et al., 1996; Patel et al., 2022), but also synthetized from nepetalic acid and nepetalactone (Eisenbraun et al., 1988; Chauhan et al., 2014). Interestingly, this compound was accumulated only in *N. nuda* inflorescences (Supplementary Table 2), indicating its possible contribution in plant-biotic interactions and reproductive strategy of this plant species. This compound was not identified in our previous study adopting a similar analytical approach (Petrović et al., 2024), most certainly since inflorescences were not chemically characterized.

Derivatives of iridoid glycosides were identified in the negative ionization mode (Supplementary Table 2), mostly showing $[M–H]^-$ as a molecular ion, while adducts with formic acid ($[M + HCOOH–H]^-$) were observed in three iridoid glycosides (**31**, **33**, and **34**). The majority of identified compounds from this group were present in both inflorescences and leaves. The exception was caffeoyl-1,5,9-*epi*-deoxyloganic acid (**39**), which was identified in *N. nuda* inflorescences only.

1,5,9-*epi*-Deoxyloganic acid was previously identified in many *Nepeta* species (Murai et al., 1984; Nagy et al., 1998; Aničić et al., 2021, including *N. nuda* (Kökdil et al., 1999; Petrova et al., 2022; Petrović et al., 2024). In the present study, 1,5,9- *epi*-deoxyloganic acid was detected in both leaves and inflorescences, being more abundant in the latter. Its three derivatives were also recorded. 1,5,9-*epi*-Deoxyloganic acid hexoside was very abundant in leaves and especially in inflorescences, while methyl-1,5,9-*epi*-deoxyloganic acid was found in leaves of *N. nuda*. To the best of our knowledge, caffeoyl-1,5,9-*epi*-deoxyloganic acid was not previously recorded in any *Nepeta* species, and in the present study it was identified in significant amounts only in inflorescences of *N. nuda*.

The most abundant compound from the subgroup of iridoid glycosides, compared by peak area, was nepetanudoside A (compound **34**). This compound, typical for *N. nuda* (Takeda et al., 1995; Petrović et al., 2024), was detected in both inflorescences and leaves but was more abundant in the reproductive organs (Figure 1A). The same trend was recorded for nepetanudoside C (**31**), which was previously reported for *N. nuda* (Takeda et al., 1996; Petrović et al., 2024).

Aucubin (**29**), present as a minor component in methanol extracts, was more abundant in inflorescences than in leaves of *N. nuda* and was previously reported for other *Nepeta* species, including *N. septemcrenata* (El-Moaty, 2010), *N. rtanjensis*, and *N. argolica* (Aničić et al., 2021). Lamiol derivative, 5-deoxylamiol (**32**), is an iridoid glycoside typical for *Lamium* species (Alipieva et al., 2003 and 2007) and has only recently been identified in *N. nuda* (Petrović et al., 2024). On the other hand, 6-deoxylamioside (**38**), which is here identified in both inflorescences and leaves, was previously reported for *L. amplexicaule* (Guiso and Martino, 1983).

To visualize the relations in metabolic fingerprints of *N. nuda* leaves and inflorescences, HCA was performed based on the Spearman's method of cluster agglomeration, utilizing the relative quantification data (peak areas) (Figure 1A). These two plant organs are clearly diversified based on the composition of phenolics and iridoids, as the majority of the identified compounds are more abundant in inflorescences than in leaves. Compounds, on the other hand, are divided into two major clusters, based on their predominance in leaves or inflorescences.

10.3389/fpls.2024.1452804

TABLE 1 GC/MS profiling of methanol extracts of *Nepeta nuda* inflorescences and leaves.

No.	t_R (min)	Compound assignment	Chemical formula	Inflorescences	Leaves
Monoterpene hydrocarbons				1.24	4.62
1G	6.39	*α*-Pinene	$C_{10}H_{16}$	0.08	0.34
2G	6.82	Sabinene	$C_{10}H_{16}$	0.30	1.11
3G	6.87	*β*-Pinene	$C_{10}H_{16}$	0.49	1.91
4G	6.97	*β*-Myrcene	$C_{10}H_{16}$	0.21	0.74
5G	7.33	*p*-Cymene	$C_{10}H_{14}$	0.07	0.22
6G	7.37	D-Limonene	$C_{10}H_{16}$	0.09	0.30
Oxygenated monoterpenes				65.81	37.36
7G	7.41	1,8-Cineole	$C_{10}H_{18}O$	12.67	22.56
8G	7.73	*cis*-Sabinenhydrate	$C_{10}H_{18}O$	0.48	0.98
9G	8.11	*trans*-2-Caren-4-ol	$C_{10}H_{16}O$	0.06	0.05
10G	8.39	Pinocarveol	$C_{10}H_{16}O$	0.35	0.04
11G	8.60	*δ*-Terpineol	$C_{10}H_{18}O$	1.21	0.76
12G	8.78	*α*-Terpineol	$C_{10}H_{18}O$	1.86	0.92
13G	8.84	(-)-Myrtenol	$C_{10}H_{16}O$	0.19	0.27
14G	9.04	Pinanediol	$C_{10}H_{18}O_2$	0.13	0.26
15G	9.16	3-Hydroxy-1,8-cineol	$C_{10}H_{18}O_2$	0.14	0.09
16G	9.49	*p*-Mentha-1(7),2-dien-8-ol	$C_{10}H_{16}O$	0.11	0.07
17G	9.95	*p*-Mentha-1,5-dien-8-ol	$C_{10}H_{16}O$	0.03	0.28
18G	10.03	*trans*,*trans*-Nepetalactone	$C_{10}H_{14}O_2$	44.05	2.57
19G	10.09	*cis*,*trans*-Nepetalactone	$C_{10}H_{14}O_2$	4.52	8.53
Sesquiterpene hydrocarbons				20.74	32.84
20G	10.26	*β*-Bourbonene	$C_{15}H_{24}$	1.06	2.03
21G	10.51	beta-Caryophyllene	$C_{15}H_{24}$	2.07	7.03
22G	10.56	*cis*-Muurola-4(15),5-diene	$C_{15}H_{24}$	0.14	0.27
23G	10.61	*β*-Farnesene	$C_{15}H_{24}$	2.42	0.12
24G	10.66	Bicyclosesquiphellandrene	$C_{15}H_{24}$	0.10	0.09
25G	10.90	Germacrene D	$C_{15}H_{24}$	14.88	23.03
26G	10.94	*α*-Farnesene	$C_{15}H_{24}$	0.08	0.28
Oxygenated sesquiterpenes				1.69	1.49
27G	11.57	Caryophyllene oxide	$C_{15}H_{24}O$	0.88	0.75
28G	11.72	Humulene epoxide I	$C_{15}H_{24}O$	0.06	0.13
29G	12.04	14-Hydroxycaryophyllene	$C_{15}H_{24}O$	0.07	0.25
30G	12.14	Germacra-4(15),5,10(14)-trien-1-ol	$C_{15}H_{24}O$	0.61	0.25
31G	13.11	Platambin	$C_{15}H_{26}O_2$	0.07	0.10
Diterpene hydrocarbons				0.11	0.62
32G	12.81	Neophytadiene	$C_{20}H_{38}$	0.11	0.62

(Continued)

TABLE 1 Continued

No.	t_R (min)	Compound assignment	Chemical formula	Inflorescences	Leaves
Oxygenated diterpenes				1.84	19.02
33G	14.18	Phytol	$C_{20}H_{40}O$	1.84	19.02
Oxygenated triterpenes				0.19	1.90
34G	17.30	Squalene	$C_{30}H_{5}0$	0.19	1.90
Other compounds				8.04	2.00
35G	9.99	Nepetonic acid	$C_9H_{14}O3$	4.12	2.00
36G	11.12	Nepetalactam	$C_{10}H_{15}NO$	3.92	/

Values represent means of three biological replicates and are presented relatively in % (100% refers to the total peak area of the specific compound).

3.3 Comparison of qualitative non-targeted profiles of volatile organic compounds in *N. nuda* leaves and inflorescences

In total, 36 VOCs were identified in methanol extracts of *N. nuda* inflorescences and leaves (Table 1).

The representative GC/MS TIC chromatograms are shown in Figure 1C. These VOCs belong to the groups of monoterpenes (6 monoterpenes and 13 oxygenated monoterpenoids) and sesquiterpenes (7 sesquiterpenes and 5 oxygenated sesquiterpenoids). Nepetalactam (**36G**), recorded only in inflorescences, was not classified in any of the two groups. Monoterpenoids were more abundant in leaves, while inflorescences were richer in oxygenated monoterpenoids (Table 1). Among monoterpenoids, *β*-pinene (**3G**) and sabinene (**2G**) predominated in *N. nuda* leaves. 1,8-Cineole (**7G**) was more abundant in leaves, although a significant amount of this compound was recorded in inflorescences, as well. From the group of monoterpenoid iridoids, two nepetalactone stereoisomers were detected: *trans,trans*- (**18G**) and *cis,trans*-nepetalactone (**19G**), the former predominating in inflorescences and the latter in leaves. Nepetonic acid (**35G**) was also more abundant in inflorescences. Interestingly, *trans,cis*- isomer of nepetalactone and some of nepetalactone derivatives (5,9-dehydronepetalactone, dihydronepetalactone, nepetalic acid), previously identified by the UHPLC-ESI-QToF-MS technique, were not recorded in methanol extracts of *N. nuda* samples using the GC/MS analysis. This can be, at least partially, ascribed to differential chromatography conditions (UHPLC vs. GC) and huge differences in the sensitivities of mass spectrometers.

Sesquiterpenoids and oxygenated sesquiterpenoids were more abundant in leaves than in inflorescences of *N. nuda* (Table 1). Germacrene D (**25G**) and beta-caryophyllene (**21G**) were detected as the major sesquiterpenoids in leaves, followed by *β*-bourbonene (**20G**). In inflorescences, germacrene D (**25G**) was the most abundant compound as well but was followed by *β*-farnesene (**23G**) and beta-caryophyllene (**21G**). Oxygenated diterpene phytol (**33G**) and oxygenated triterpene squalene (**34G**) were by far more abundant in leaves of *N. nuda* than in inflorescences. In sum, inflorescences and leaves of *N. nuda* were clearly distinguished based on the content of teprenoids and related compounds and can therefore be considered as different chemotypes.

Only scarce literature data reported on the composition of *N. nuda* methanol extracts revealed by GC/MS characterization. Other studies primarily focused on analyzing essential oils composition and reported the following compounds as dominant components: caryophyllene oxide (Kökdil et al., 1998), camphor and 1,8-cineole (Kilic et al., 2011), caryophyllene (Alim et al., 2009), nepetalactone (Bozok et al., 2017; Handjieva et al., 1996; Kökdil et al., 1996), 1,8-cineole and 4a*α*,7*β*,7a*α*-nepetalactone (Mamadalieva et al., 2016), nepetalactone and germacrene (Gormez et al., 2013), 1,8-cineole (Chalchat et al., 1998), and 1,8-cineole and a mixture of nepetalactones and germacrene-D (De Pooter et al., 1988). It must be considered that these differences might arise from many factors, including plant origin, growth conditions, developmental stage, plant part analyzed, a subspecies or a cultivar in question, essential oils isolation procedure, and GC/MS conditions. A study by Zaharieva et al. (2023) revealed 4a*α*,7*β*,7a*α*-nepetalactone, 1,8-cineole/eucalyptol, and germacrene D as the most abundant volatile compounds in *N. nuda* plants grown under either *in vitro* or *ex vitro* conditions. Caryophyllene, *β*-ocimene, bicyclogermacrene, *β*-pinene, myrcene, and humulene were also found in quite fair quantities.

The relative GC/MS quantitative data (peak areas) of methanol extracts of *N. nuda* were subjected to HCA to visualize the phytochemical relations between inflorescences and leaves (Figure 1C). Leaves cluster separately from inflorescences on the HCA plot constructed based on the Spearman's algorithm, confirming the clear diversification between these plant parts. Compounds are clustered based on their predominance in inflorescences or leaves.

3.4 Organ-specific variation in the quantitative content of major iridoids and phenolics among the Central Balkans' populations of *N. nuda*

The UHPLC/DAD/(±)HESI–MS^2 analysis was targeted towards totally 9 compounds belonging to the groups of iridoids (5 compounds) and phenolics (4 compounds), and the selection was based on previous studies reporting high abundance of these compounds in *N. nuda* (Mišić et al., 2015; Aras et al., 2016;

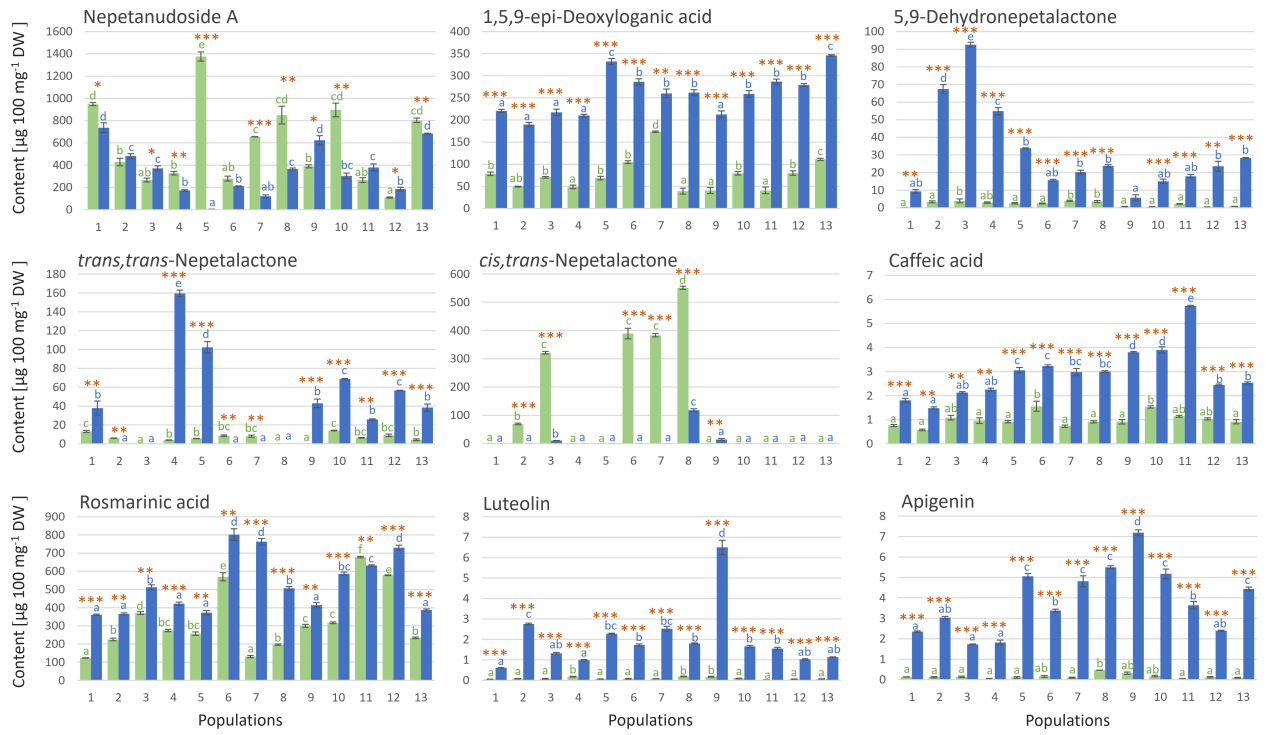


FIGURE 2
UHPLC/DAD/(±)HESI–MS^2 quantification of targeted iridoids and phenolics in methanol extracts of *Nepeta nuda* leaves (green bars) and inflorescences (blue bars), in samples originating from 13 Central Balkan populations. SE of three biological replicates is presented. Values labeled with different letters, for leaves (green letters) and inflorescences (blue letters) independently, are significantly different ($p < 0.05$) according to the Tukey's *post hoc* test of one way ANOVA. Orange asterisks above the bars denote significantly different values between leaves and inflorescences according to the t-test, *p*-values, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Abbreviations of *N. nuda* populations: 1-Mali Ljukten, 2-Brodica, 3-Debeli lug, 4-Rtanj, 5-Straža, 6-Židilje, 7-Vinatovača, 8-Vlasina, 9-Donji Krivodol, 10-Topli Do, 11-Janjska reka, 12-Balta Berilovac, 13-Gornje selo.

Smiljković et al., 2018; Petrova et al., 2022; Petrović et al., 2024). From the group of iridoids, we selected three aglycones (*cis*,*trans*- and *trans*,*cis*-nepetalactone, and 5,9-dehydronepetalactone) and two glycosides (1,5,9-*epi*-deoxyloganic acid and nepetanudoside A). The most abundant phenolics in *N. nuda* were rosmarinic acid and caffeic acid (Mišić et al., 2015; Petrova et al., 2022), and these two phenolic acids were quantified along with two flavonoids from the group of flavones, luteolin and apigenin (Figure 2). The majority of the analyzed metabolites were more abundant in inflorescences than in leaves of *N. nuda*. Literature data support these findings, as inflorescences have generally higher constitutive levels of defensive metabolites than leaves (Brown et al., 2003; Damle et al., 2005; Smallegange et al., 2007), with some exceptions (Godschalx et al., 2016). On the other hand, the amounts of these compounds were variable among the 13 analyzed populations.

The results display obvious differentiation between leaves and inflorescences (Figure 2). The major iridoid glycoside, nepetanudoside A, was present in both inflorescences and leaves of *N. nuda*, and in populations 1, 4, 5, 7, 8, 10, and 13, the content of this compound was significantly higher in leaves than in inflorescences. Population 5 was especially rich in this compound. 1,5,9-*epi*-Deoxyloganic acid was more abundant in inflorescences than in leaves of *N. nuda*, and this implies to all analyzed populations (Figure 2). The highest amounts of this compound were recorded in samples belonging to the populations 5 and 13. Of the iridoid aglycones, *trans*,*trans*-nepetalactone and 5,9-dehydronepetalactone were more abundant in flowers than in leaves of *N. nuda*. The content of *trans*,*trans*-nepetalactone varied between populations, with the population 4 and 5 being especially rich in this compound. The highest amounts of 5,9-dehydronepetalactone were recorded in samples belonging to the populations 2, 3, and 4. Interestingly, *cis*,*trans*-nepetalactone was recorded in significant amounts in only 5 populations (2, 3, 6, 7, and 8), primarily in leaves (Figure 2).

To visualize variations in the quantitative composition of the major iridoids in inflorescences and leaves of 13 N. *nuda* populations, PCA was performed adopting the Ward's method of data agglomeration (Figure 3A). The major contributors to the diversification of samples along the PC1 explaining 75.76% of the total variability are nepetanudoside A, *cis*,*trans*-nepetalactone, and 1,5,9-*epi*-deoxyloganic acid. As for the PC2, which contributes to the total variability with 17.43%, *cis*,*trans*-nepetalactone and 1,5,9-*epi*-deoxyloganic acid are the major contributors to the diversification along this component. Samples of *N. nuda* leaves are more scattered along PC1 than samples of inflorescences, implying higher uniformity in the accumulation of the targeted compounds by inflorescences.

Among phenolics, rosmarinic acid was by far the most abundant compound in the analyzed samples. It was recorded in considerable amounts in both leaves and inflorescences of *N. nuda*.

10.3389/fpls.2024.1452804

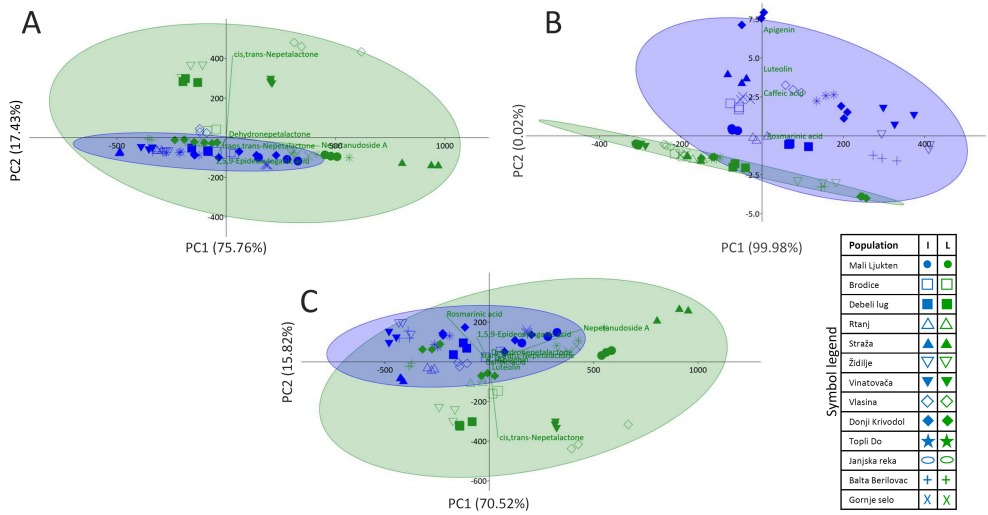


FIGURE 3
Principal component analysis (PCA) biplots constructed based on the UHPLC/DAD/(±)HESI–MS^2 quantitative data, separately for iridoids **(A)** and phenolics **(B)**, and cumulatively for both groups of metabolites **(C)**. Participation of the variables (metabolites) in the first two PCs is indicated by the corresponding loading plots. Inflorescences are labeled with blue symbols while leaves are colored green.

Its amounts in inflorescences reached ~802 μg 100 mg^{-1} dry weight (DW) (population 6), while in leaves ~678 μg 100 mg^{-1} DW was recorded in samples from the population 11 (Figure 2). In the majority of the analyzed populations, the amounts of this compound were slightly higher in inflorescences than in leaves. Caffeic acid, luteolin, and apigenin were also more abundant in inflorescences than in leaves in all analyzed populations. Population 9 was characterized by the highest amounts of luteolin (6.50 μg 100 mg^{-1} DW) and apigenin (7.19 μg 100 mg^{-1} DW), while the highest amounts of caffeic acid were recorded for the population 11 (Figure 2).

When PCA was performed with quantitative data for phenolics only, it was obvious that samples of leaves and inflorescences were clearly diversified along PC2 (Figure 3B). The major contributors to the diversification along PC1, which explains 99.98% of the total variability, and along PC2 (0.02%), are apigenin and luteolin. The third contributor is caffeic acid. Results imply that phenolics content is more conserved in leaves than in inflorescences.

PCA plot was further constructed based on the quantitative data for both iridoids and phenolics (Figure 3C). PC1 explains 70.52% of the total variability, while PC2 is responsible for 15.82% of the variability. The major contributors to the diversification along PC1 are nepetanudoside A, rosmarinic acid, and *cis*,*trans*-nepetalactone. The same compounds, but with perturbed order, are the major factors that diversify samples along PC2.

3.5 Molecular background of iridoid biosynthesis in *N. nuda* inflorescences and leaves

We further aimed to get deeper into the molecular background of the organ-specific iridoid diversity in *N. nuda*. To exclude the effect of genotype and environmental conditions on the chemical phenotype, we performed clonal propagation of a single genotype originating from the population 3 (abbreviation DL1). Plantlets having the uniform genetic background were acclimatized and grown under controlled greenhouse conditions. Samples of leaves and inflorescences were collected from fully-flowering plants, in three biological repetitions.

Targeted metabolic profiling of five iridoids (*cis*,*trans*-nepetalactol, *trans*,*trans*-nepetalactone, 5,9-dehydronepetalactone, 1,5,9-*epi*-deoxyloganic acid, and nepetanudoside A) in leaves and inflorescences of greenhouse-grown *N. nuda* in parallel with co-expression analysis of iridoid biosynthesis-related genes, revealed the molecular background of their organ-specific composition. As expected, all the quantified iridoids were more abundant in inflorescences than in the leaves of *N. nuda* plants (Figure 4A). The difference in the abundance of specific iridoids between the two plant organs was the most pronounced for 1,5,9-*epi*-deoxyloganic acid and *cis*,*trans*-nepetalactol, the latter being recorded only in inflorescences. The correlation analysis based on the quantitative data revealed significant positive correlations for the majority of comparisons (Figure 4B). These results are in accordance with the study of Gomes et al. (2024), which associated the highest concentrations of nepetalactones, nepetalic acid and nepetalactam with the floral-bud and partial-flowering stages of *N. cataria*.

The iridoid biosynthetic pathway in *Nepeta* species has been extensively researched in the past, and it is believed that nepetalactol is the common precursor of all other iridoids in these species. The biosynthesis of nepetalactol involves a series of enzymatic steps catalyzed by geranyl diphosphate synthase (GPPS), geraniol synthase (GES), geraniol 8-hydroxylase (G8H), 8-hydroxygeraniol oxidoreductase (8HGO), iridoid synthase (ISY), NAD-dependent nepetalactol-related short-chain-dehydrogenase/reductase (NEPS), and major latex protein-like (MLPL) enzymes (Figure 4C). In *Nepeta* species, ISYs are mainly responsible for the stereoselective 1,4-reduction of 8-oxogeranial (8OG) to uncyclized

10.3389/fpls.2024.1452804

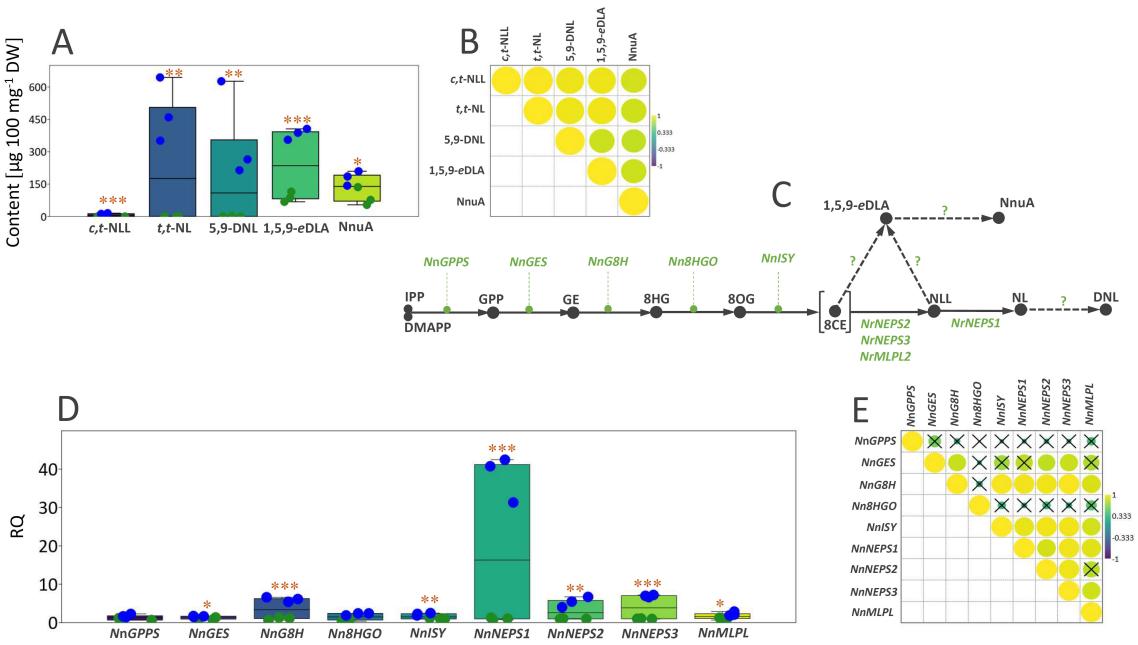


FIGURE 4
(A) Box-plot diagram depicting the distribution of the content of three iridoid aglycones (*c,t*-NLL, *t,t*-NL, 5,9-DNL) and two iridoid glycosides (1,5,9-*e*DLA and NnuA) in leaves (green dots) and inflorescences (blue dots) of *Nepeta nuda* grown under greenhouse conditions. **(B)** Pearson's correlations based on the content of five major iridoids in *N. nuda* leaves and inflorescences represented by the intensity of yellow (positive correlations) or blue (negative correlations), as indicated on the color scale. **(C)** A proposed iridoid metabolic pathway in *N. nuda*. Unknown enzymes of the biosynthetic pathway are marked with the question mark. **(D)** Box-plot diagram depicting the distribution of the transcript levels of nine iridoid biosynthesis-related genes in leaves (green dots) and inflorescences (blue dots) of *N. nuda*. **(E)** Pearson correlations based on the expression level of nine biosynthetic genes in *N. nuda* leaves and inflorescences. Non-significantly correlated values are crossed ($p > 0.05$). IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pirophosphate; GPP, geranyl pyrophosphate; GE, geraniol; 8HG, 8-hydroxygeraniol; 8OG, 8-oxogeranial; 8CE, 8-oxocitronellyl enolate; NLL, nepetalactol; NL, nepetalactone; DNL, dehydronepetalactone; *c,t*-NLL, *cis,trans*-nepetalactol; *t,t*-NL, *trans,trans*-nepetalactone; 5,9-DNL, 5,9-dehydronepetalactone; 1,5,9-*e*DLA, 1,5,9-*epi*deoxyloganic acid; NnuA, nepetanudoside A; NnGPPS, geranyl diphosphate synthase; NnGES, geraniol synthase; NnG8H, geraniol 8-hydroxylase; Nn8HGO, 8-hydroxygeraniol oxidoreductase; NnISY, iridoid synthase; NnNEPS, nepetalactol-related short-chain dehydrogenase; NnMLPL, major latex protein-like enzyme. Asterisks denote significantly different values according to the *t* test (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

and reactive 8-oxocitronellyl enol (8CE) and for determining the stereochemistry of C7 (Sherden et al., 2018; Lichman et al., 2019b and 2020; Hernaíndez Lozada et al., 2022). Enolate intermediate 8-oxocitronellyl enol in *Nepeta* is cyclized by NEPSs (Lichman et al., 2019a, 2019), but can also undergo a spontaneous cyclization to produce predominately *cis*,*trans*- steroisomer of nepetalactol (Hernaíndez Lozada et al., 2022). Although the upstream parts of the pathway have been well described, the downstream parts that are leading to various derivatives of nepetalactone, as well as those of the iridoid glycosides branch, remain largely unclear.

By BLAST-searching for the available *N. nuda* transcriptomes, candidates for 9 iridoid biosynthesis-related genes (*NnGPPS*, *NnGES*, *NnG8H*, *Nn8HGO*, *NnISY*, *NnNEPS1*, *NnNEPS2*, *NnNEPS3*, and *NnMLPL*) were identified and their expression profiles were comparatively analyzed in *N. nuda* leaves and inflorensces to more deeply investigate the molecular background of their variation between the two organs and to point to the key determinants of the organ-specific iridoid metabolism in *N. nuda*. The expression levels of *NnGPPS* and *Nn8HGO*, two genes of the early biosynthetic pathway (EBGs), did not significantly differ between leaves and inflorescences (Figure 4D). However, *NnGES*, *NnG8H*, and *NnISY* displayed around 1.6-, 6-, and 2.2-fold higher transcript levels in inflorescences than in leaves, respectively. Interestingly, genes responsible for the steps directly preceding the formation of nepetalactol and nepetalactone (*NnNEPS1*, *NnNEPS2*, *NnNEPS3*, and *NnMLPL*) showed significantly higher expression level in inflorescences than in leaves (Figure 4D). In the case of *NnNEPS1*, which belongs to the subgroup of NEPS oxidases and is responsible for converting nepetalactol into nepetalactone, around 40-fold higher expression level was recorded in inflorescences than in leaves. The NEPS oxidase (*NnNEPS1*) identified in the transcriptomes of *N. nuda* leaves, that corresponded to the *N. mussinii* NEPS1, has been reported to have *cis*,*trans*-nepetalactol dehydrogenase activity (Lichman et al., 2019a). *NnNEPS2* is a homologue of *N. sibirica* NEPS2 (*Ns*NEPS2), which acts both as a *cis-trans* cyclase and an oxidase (Lichman et al., 2019a; Hernaíndez Lozada et al., 2022). *NnNEPS3* is a homologue of *N. mussinii* NEPS3 displaying a specific 7S-*cis*,*cis*-nepetalactone cyclase activity (Lichman et al., 2019a). The composition, expression level, and ratio of NEPS enzymes are responsible for setting the stereochemistry of nepetalactones (Lichman et al., 2019a), which is most certainly the case in *N. nuda* as well. The results indicate that *NnG8H* and the 3 *NEPSs* might serve as the regulatory points, determining the flux through the iridoid pathway and thus the level of iridoid biosynthesis and accumulation in different organs. Hence, the high level of nepetalactone and other

iridoids in inflorescences is, most likely, the result of upregulated expression of *NnG8H*, *NnNEPS1*, *NnNEPS2*, and *NnNEPS3* in these organs. The majority of correlations observed between the analysed biosynthetic genes were positive and statistically significant, which suggests the regulation of iridoid biosynthesis in different organs at the transcriptional level (Figure 4E). The exception were *NnGPPS* and *Nn8HGO*, which displayed no significant correlations with other genes. Conversely, the expression of *NnGES* was significantly positively correlated with that of *NnG8H*, *NnNEPS2*, and *NnNEPS3*.

Some previous studies dealing with *Nepeta* species revealed organ-specific expression patterns of the iridoid biosynthesis-related genes (Aničić et al., 2018, and 2024, Lichman et al., 2020). Closed buds and open flowers of *N. cataria* and *N. mussinii* were more enriched with transcripts of NEPSs, G8H, and ISY than mature leaves (Lichman et al., 2020). Among NEPS enzymes, NEPS1 displayed the highest expression level in open flowers and closed buds of *N. cataria*, while NEPS5 transcripts were the most abundant in closed buds of *N. mussinnii* (Lichman et al., 2020). This is in accordance with the findings of the present study, as *NnNEPS1* displayed a significantly higher expression level in inflorescences than in mature leaves of *N. nuda* and more transcript levels than the other two NEPSs (*NnNEPS2* and *NnNEPS3*). Although we did not separately assessed closed buds from open flowers, the analyzed inflorescences contained flowers in all development stages.

The expression of iridoid biosynthesis-related genes is coordinately regulated and well correlated with metabolite pools in mature leaves and inflorescences (Table 2), supporting the hypothesis that the organ-specific biosynthesis of iridoids is controlled at the transcriptional level. In other words, gene transcripts of targeted genes were more abundant in inflorescences, where the majority of targeted iridoids accumulated. The exceptions are *NnGPPS*, *NnGES*, and *Nn8HGO*, as their expression levels were not significantly correlated with the content of targeted metabolites. The majority of significant correlations were observed between the three *NEPSs* and *G8H* on one hand and iridoid aglycones (5,9-dehydronepetalactone, *cis*,*trans*-nepetalactol, and *trans*,*trans*-nepetalactone) on the other. The most significant correlations were observed between the expression of *NnNEPS1* and *NnNEPS3* and the three iridoid aglycones. *NnNEPS1* was also positively correlated with the content of quantified iridoid aglycones, nepetanudoside A and 1,5,9-*epi*-deoxyloganic acid (Table 2), indicating possible involvement of this NEPS oxidase in the formation of iridoid glycosides. These speculations should be further investigated.

3.6 Possible ecological implications of iridoids' composition in *N. nuda* inflorescences and leaves

Due to a variety of important roles that iridoids play in plant-biotic interactions, we here discuss possible ecological implications of their organ-specific composition in *N. nuda*. Data show that *N. nuda* leaves exhibit greater diversity in terms of the targeted phenolics' and iridoids' content, and that phytochemical composition of inflorescences is more conserved among the analyzed populations. Plants likely invest more resources in the constitutive production of floral metabolites functioning in plant-biotic interactions (i.e., plant-pollinator, plant-pathogen, and plant-herbivore interplay), rather than in inducible production. It was previously hypothesized that constitutive resistance predominates in flowers (McCall and Irwin, 2006; Zangerl, 2003). Floral chemistry influences the interactions of plants with pollinators and other species and maximizes the fitness of both plants and flower visitors (Palmer-Young et al., 2019; Kessler and Kalske, 2018). The defense mechanisms of reproductive tissues, such as flowers, are crucial for plant reproduction, and their defensive strategies may differ from those of leaves, which have the primary function of fixing CO_2 during photosynthesis. According to Pichersky and Gershenzon (2002), floral volatiles serve as attractants for species-specific pollinators, whereas the volatiles emitted from vegetative parts, especially those released after herbivory, appear to protect plants by deterring herbivores and by attracting their enemies. Flowers are essentially heterotrophic as they heavily rely on organic molecules synthesized in leaves and roots. However, they biosynthesize and

TABLE 2 Correlations (R^2) between quantitative metabolomics data (for 5 selected iridoid compounds) and qPCR expression data for 9 iridoid biosynthesis-related genes in the analyzed samples of *Nepeta nuda* (*n*= 324).

Genes	Metabolites				
	Nepetanudoside A	1,5,9-*epi*-Deoxyloganic acid	Dehydronepetalactone	*c*,*t*-Nepetalactol	*t*,*t*-Nepetalactone
NnGPPS	0.05099	0.01171	0.02422	0.01530	0.02615
NnGES	0.40227	0.69953	0.62632	0.73876	0.75092
NnG8H	0.49685	0.92993*	0.65089**	0.76221**	0.84667**
Nn8HGO	0.14440	0.07399	0.05780	0.01432	0.04141
NnISY	0.47210	0.86730	0.62146	0.66582**	0.77429**
NnNEPS1	0.50028**	0.91788**	0.48505***	0.69303***	0.75252***
NnNEPS2	0.59176	0.89129	0.84279	0.81969**	0.90995**
NnNEPS3	0.61061*	0.97990**	0.70291***	0.84138**	0.90471***
NnMLPL	0.27544	0.67000	0.27289	0.38925	0.48010*

R^2, coefficient of determination; n, total number of comparisons; ***, **, * – significance levels at $p < 0.001$, $p < 0.01$, and $p < 0.05$, respectively.

accumulate complex molecules that sustain plant interactions with animal pollinators and pathogens, play roles in plant fertilization, embryo development, and the initial stages of fruit and seed set.

The chemical composition of colorful long-lasting flower spikes of *N. nuda* is a key determinant of their color, odor, and overall function in plant-biotic interactions. *Nepeta* flowers with their visual, chemical, and nutritional traits are very attractive to honeybees and other pollinators. Honeybees are the main pollinators of *N. nuda* (Bozek, 2003). The most important visitors of flowers of *N. cataria*, a species sympatric to *N. nuda*, are honeybees, solitary bees, and bumblebees (Cruden and Miller-Ward, 1981; Sih and Baltus, 1987), but also flies and butterflies (Aizen and Harder, 2009), thus it may be presumed that *N. nuda* is visited by similar groups of pollinators, as these two species share similar phytochemical profiles.

The increased content of iridoids in the reproductive organs of *N. nuda* might indicate their functional involvement in complex plant-biotic interactions. It is well documented that iridoids function as general toxicants, growth and reproductive inhibitors, repellents or oviposition-deterrents (Bernier et al., 2005; Birkett et al., 2011; Sparks et al., 2018; Reichert et al., 2019; reviewed in Formisano et al., 2011; Süntar et al., 2018). Two groups of iridoids present in *Nepeta* plants, volatile aglycones and glycosylated iridoids, most likely provide a two-level chemical defense, the former acting as repellents and/or pollinator attractants, and the latter contributing as deterrents. Both subgroups of iridoids also provide antimicrobial protection.

Constituents of *Nepeta* species, such as nepetalactone, dihydronepetalactone, and nepetalactol, are proven repellents against insects, such as mosquitos (*Aedes*, *Anopheles*, *Culex*), cockroaches (*Periplaneta*, *Blattella*), ticks (*Rhipicephalus*, *Ixodes*), flies (*Stomoxys*, *Musca*), mites (*Dermanyssus*, *Dermatophagoides*), and termites (*Reticulitermes*) (Bernier et al., 2005; Birkett et al., 2011; reviewed in Formisano et al., 2011; Sparks et al., 2018; reviewed in Süntar et al., 2018; Reichert et al., 2019). Dihydronepetalactone is reported to be two times more active than commercial N,N-diethyl-3-methylbenzamide (DEET) (Feaster et al., 2009). Notably, nepetalactone and nepetalactol mediate plant-insect interactions, but are also produced by a number of insects, most notably aphids, acting as sex pheromones (Fernández-Grandon and Poppy, 2015). In addition, dihydronepetalactones are the components of defensive secretions of some ant species (Cavill and Clark, 1967). As reported by Harney et al. (1978), nepetalic acid, along with nepetalactone and catnip essential oil, had a depressant effect in mice when administered intraperitoneally.

On the other hand, compounds from the group of iridoid glycosides present in *Nepeta* species, less volatile and more stable than nepetalactones and their derivatives, have rarely been investigated for their interaction with insects and generally for their contribution to the overall defense strategies. 1,5,9-*epi*-Deoxyloganic acid has been scarcely investigated for its biological activities, thus, only studies dealing with its antimicrobial and anti-inflammatory potential (Aničić et al., 2021), as well as enzyme inhibition and DNA protection activities (Başar et al., 2023) are presented in the literature. To the best of our knowledge, no data about its role in plant-biotic interactions are available. Despite only scarce information about the bioactivities of the major iridoid glycosides of *N. nuda*, plenty is known about some iridoid glycosides present as minor constituents. Such an example is aucubin, which displays potent insecticidal activity against various insects (Vivekanandhan et al., 2022, 2023) and antifungal activity (Marak et al., 2002). Comprehensive literature search revealed that the roles of nepetenudoside A and C, as well as of nepetaside, 5-deoxylamiol, and 6-deoxylamioside in plant-biotic interactions have not been previously documented. Due to their physico-chemical properties, iridoid glycosides of *Nepeta* sp. might provide additional values in the development of novel biopesticides.

The composition of floral volatiles, a chemically diverse group of plant metabolites with multiple functions, is shaped by environmental, ecological and evolutionary factors (Dötterl and Gershenzon, 2023). Although iridoids are by far the most abundant group of terpenes in *Nepeta* species, their synergistic effects with other terpenes must not be ruled out. The unique metabolome of each individual plant constitutes the defense machinery against biotic environmental factors and determines its success in the changing environment. Several reports showed insecticidal and repellent activity of eucalyptol, which is abundant in both leaves and inflorescences of *N. nuda*. Eucalyptol induces intoxication and has a prominent effect on locomotor activity, knock-down, and repellence on nymphs of *Triatoma infestans* and *Rhodnius prolixus* (Moretti et al., 2015). It also displays repellence against a number of insects, including house fly (Sukontason et al., 2004), *Periplaneta americana* (Scriven and Meloan, 1984), and *Aedes aegypti* (Klocke et al., 1987). On the other hand, eucalyptol is attractive to honey bees and euglosine bees (Effmert et al., 2005).

It was previously reported that the content and ratio of sesquiterpenoids β-farnesene and germacrene D in flowers of *Tanacetum vulgare* modulated both defense and pollination (Li et al., 2019). Aphid alarm pheromone β-farnesene is the key volatile cue responsible for specific attraction of beetles and avoidance by aphids at the early flowering stages (Li et al., 2019). As these terpenoids are also found in *N. nuda* inflorescences and leaves, it could be hypothesized that they can play important roles in plant-biotic interactions.

3.7 Molecular background of the biosynthesis of phenolics in *N. nuda* inflorescences and leaves

As expected, rosmarinic acid was by far the most abundant phenolic compound in leaves and inflorescences of *N. nuda* grown under greenhouse conditions (Figures 5A, B). Other phenolics quantified in samples were present in much lower amounts. All analyzed compounds were found in significantly higher amounts in inflorescences than in leaves. The amount of rosmarinic acid was significantly negatively correlated with those of caffeic acid, luteolin, and apigenin (Figure 5C). The only non-significant correlation was observed between luteolin and apigenin.

The biosynthesis of rosmarinic acid (Figure 5D) involves phenylpropanoid- and tyrosine-derived branches that use the amino acids L-phenylalanine and L-tyrosine as precursors, respectively (Matsuno et al., 2002; Petersen et al., 2009; Wang et al., 2017; Trócsányi et al., 2020; Suthar et al., 2021) and it proceeds by

10.3389/fpls.2024.1452804

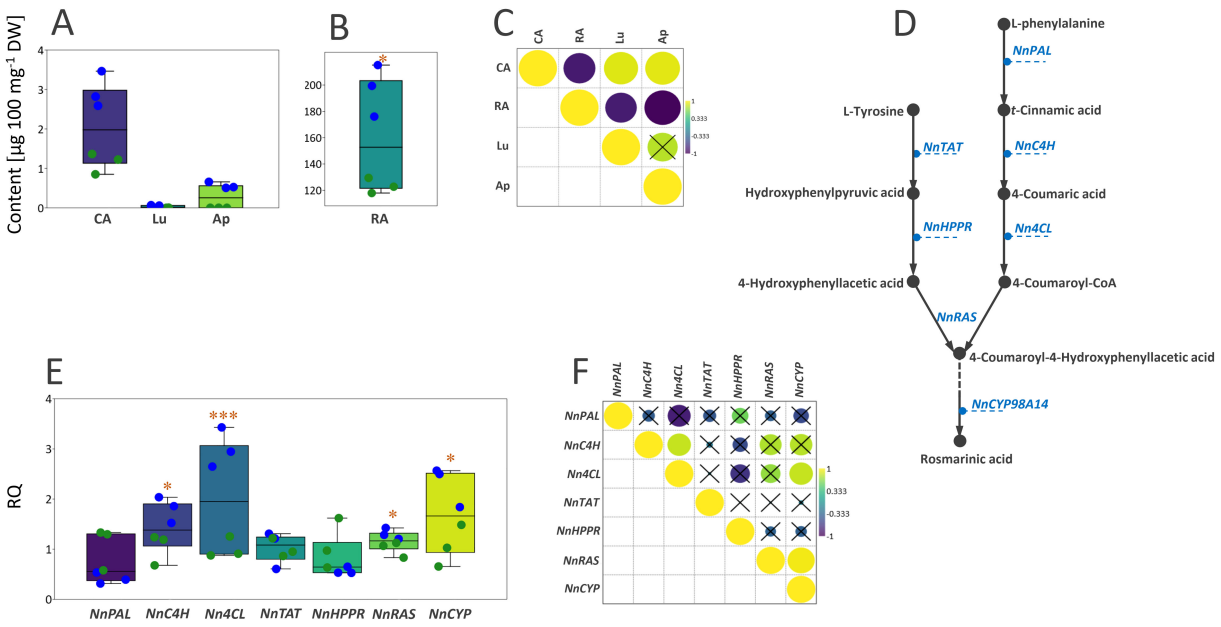


FIGURE 5
(A) Box-plot diagram depicting the distribution of the content of one phenolic acid (CA) and two flavonoids (Ap and Lu) in leaves (green dots) and inflorescences (blue dots) of *N. nuda* grown under greenhouse conditions. **(B)** RA is presented by a separate box-plot diagram. **(C)** Pearson's correlations based on the content of four major phenolics in *N. nuda* leaves and inflorescences, represented by the intensity of yellow (positive correlations) or blue (negative correlations), as indicated on the color scale. Non-significantly correlated values are crossed ($p > 0.05$). **(D)** A proposed biosynthetic pathway of rosmarinic acid in *N. nuda*. **(E)** Box-plot diagram depicting the distribution of the transcript levels of seven rosmarinic acid biosynthesis-related genes in *N. nuda* leaves and inflorescences, represented by the intensity of yellow (positive correlations) or blue (negative correlations), as indicated on the color scale. **(F)** Pearson's correlations based on the expression level of seven genes of the rosmarinic acid biosynthetic pathway in *N. nuda* leaves (green dots) and inflorescences (blue dots). Non-significantly correlated values are crossed ($p > 0.05$). CA, caffeic acid; RA, rosmarinic acid; Ap, apigenin; Lu, luteolin; PAL, phenylalanine ammonia-lyase; C4H, cinnamic acid 4-hydroxylase; 4CL, 4-coumaric acid CoA-ligase; TAT, tyrosine aminotransferase; HPPR, hydroxyphenylpyruvate reductase; RAS, 4-coumaroyl-CoA:4′-hydroxyphenyllactic acid 4-coumaroyltransferase; CYP, cytochrome P450-dependent monooxygenase. Asterisks denote significantly different values according to the *t* test (*$p < 0.05$, ***$p < 0.001$).

activating the following enzymes: phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H), 4-coumaric acid CoA-ligase (4CL), tyrosine aminotransferase (TAT), hydroxyphenylpyruvate reductase (HPPR), "rosmarinic acid synthase" (RAS; 4-coumaroyl-CoA:4′-hydroxyphenyllactic acid 4-coumaroyltransferase), and cytochrome P450-dependent monooxygenase. We here analyzed the expression profiles of 7 genes related to the biosynthesis of rosmarinic acid (*NnPAL*, *NnC4H*, *Nn4CL*, *NnTAT*, *NnHPPR*, *NnRAS*, and *NnCYP*) in *N. nuda* leaves and inflorescences. Looking at their enzymatic activities involved in this biosynthetic pathway, only two of them might be specific for rosmarinic acid biosynthesis, *NnRAS* and *NnCYP*. These two genes showed significantly higher expression levels in inflorescences than in leaves (Figure 5E), which is in accordance with a higher content of rosmarinic acid in these organs. Enzymes belonging to the general phenylpropanoid pathway (*NnPAL*, *NnC4H*, and *Nn4CL*) provide precursors for the formation of other metabolites such as lignin and other groups of phenolic compounds (e.g., flavonoids, anthocyanins). Transcript levels of *NnC4H* and *Nn4CL* were higher in inflorescences than in leaves, while *NnPAL* did not significantly differ between the two organs (Figure 5E). TAT and HPPR are also primary enzymes, and their activity results in the formation of precursors in the biosynthesis of tocopherols and plastoquinones. These two genes did not significantly differentially express between leaves and inflorescences of *N. nuda* (Figure 5E). The only significant correlations were positive correlations between the expression patterns of the analyzed rosmarinic acid-related biosynthetic genes: *Nn4CL* and *NnC4H*, *NnCYP* and *Nn4CL*, and *NnCYP* and *NnRAS* (Figure 5F).

The results further suggest that the expression levels of *NnPAL*, *NnC4H*, and *Nn4CL*, which belong to the general phenylpropanoid pathway, as well as of *NnRAS* and *NnCYP*, are not significantly correlated with the rosmarinic acid content in the two plant organs (Table 3). This is not surprising, taking into account that the majority of analyzed genes are related to the general phenylpropanoid pathway and are thus involved in providing the precursors for the biosynthesis of various classes of phenolic compounds. Furthermore, the differences in the quantitative content of the major phenolic compounds between leaves and inflorescences are less prominent than those recorded for iridoids.

3.8 Possible ecological implications of differential phenolics' composition in *N. nuda* inflorescences and leaves

The involvement of phenolic compounds in plants ranges from growth and development (e.g., cell wall thickening, hormone production, and pigmentation), reproduction (e.g., pigmentation, fruit flavoring, and fruit protection), defense against stressors (e.g.,

TABLE 3 Correlations (R^2) between quantitative metabolomics data (for 4 selected phenolic compounds) and qPCR expression data for 7 rosmarinic acid biosynthesis-related genes in mature leaves and inflorescences of *Nepeta nuda* (n= 108).

Genes	Metabolites			
	Caffeic acid	Rosmarinic acid	Luteolin	Apigenin
NnPAL	0.70436	0.37304	0.41609	0.6189
NnC4H	0.44942	0.88650	0.44946	0.77645
Nn4CL	0.80174	0.81177	0.78283	0.89404
NnTAT	0.03716	0.00140	0.04487	0.01887
NnHPPR	0.47873	0.22249	0.27942	0.42757
NnRAS	0.62464	0.74598	0.67462	0.56519
NnCYP	0.72381	0.83617	0.87822	0.66551

No significant correlations were observed.
R^2, coefficient of determination; n, total number of comparisons.

osmoregulation, UV protection, anti-herbivory roles, and antimicrobial activity) (Dixon, 2001; Singh et al., 2021), and leaf and petal senescence (Tyagi et al., 2020). The content of phenolics in plants is determined by various endogenous and environmental factors. Thus, high constitutive and inducible leaf flavonoid content has been correlated with insect and pathogen resistance (Treutter, 2005). The function of phenolics in flowers is particularly versatile: from antioxidant activity, through the protection from UV radiation, to providing diverse colors. Phenolics accumulated in flowers primarily protect from nectar robbers and enhance memory and foraging efficiency of pollinators or, as in the case of flavonoids, exhibit high antioxidant activities supporting pollinator well-being (Slavković and Bendahmane, 2023). Moreover, flavonoids can regulate biotic interactions with mutualists and antagonists and thus reduce herbivore attacks and infection (Cushnie and Lamb, 2005; Palmer-Young et al., 2019).

Rosmarinic acid, which is by far the most abundant compound in *N. nuda*, is known for its antioxidant (Erkan et al., 2008; Muñoz-Muñoz et al., 2013), antimicrobial (Bais et al., 2002; Ivanov et al., 2022), insecticidal activities (Khan et al., 2019; Darrag et al., 2022), and various health promoting effects (Amoah et al., 2016; Alagawany et al., 2017; Nadeem et al., 2019). Rosmarinic acid displays insecticidal effects against the pea aphid, *Acyrthosiphon pisum* (Khan et al., 2019), and is a highly potent feeding deterrent for the tobacco hornworm (*Manduca sexta*) (Simmonds et al., 2019). Salvigenin, nepetoidin B, and rosmarinic acid display respectable insecticidal activities against the red palm weevil (*Rhynchophorus ferrugineus*) (Darrag et al., 2022). Conversely, this compound has been reported to reduce honey bee mortality caused by pesticides (Ulziibayar et al., 2022).

4 Conclusions

Inflorescences and leaves of *Nepeta nuda* showed a differential investment in the constitutive production of phenolics and terpenes, based on their chemical profiles, which emphasizes the importance of the chemical profiling of different plant organs and encompassing both groups of metabolites to properly characterize the species' metabolome. The data strongly indicate that plants prioritize constitutive production of floral metabolites, whose qualitative and quantitative content is, to a large extent, conserved among different populations. *NnG8H* and *NnNEPS1-3*, are recognized as important regulatory points responsible for the increased content of iridoids in inflorescences.

The presented results lead to the unambiguous conclusion that specialized metabolism in leaves is more reprogrammable in response to differential growth conditions than that in inflorescences. Therefore, vegetative parts are highlighted as more convenient objects for studying the environmentally-driven chemical diversity among *N. nuda* populations. However, the specific mechanisms and the exact environmental cues associated with the organ-specific chemical composition in *N. nuda* remain largely unexplored. Due to its rich metabolite profile and a potential ornamental value, *N. nuda* can be a suitable alternative to *N. cataria*, which is the most intensively cultivated *Nepeta* species. The information about the intra-individual (organ-specific) chemical specificities of *N. nuda* provides strategic tools for growers to manipulate the phenology of this medicinal crop and achieve a higher productivity of metabolites of interest. Given the importance of iridoids as powerful insect repellents, future studies should be conducted towards optimization of agronomic practices to maximize yields in versatile geographic regions and to facilitate the integration of iridoid-containing biopesticides into pest management systems and other applications related to human health.

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

LP: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft. BF: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. MS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – original draft. BŠ: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft. TB: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. DMa: Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – review & editing. JNŽ: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. SD: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. NA: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. MM: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review &

editing. JB: Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing. UG: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft. DMi: Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research was financed by the Science Fund of the Republic of Serbia Grant No. 7749433, project acronym NEPETOME, and is also supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia [Contracts No. 451-03-47/2023-01/200007 and 451-03-66/2024-03/200007.

Acknowledgments

The authors would like to acknowledge David L. Hallahan, Ph.D. and Entomol Products LLC (GA, USA) for providing the standard of *cis*,*trans*-nepetalactone.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1452804/full#supplementary-material

References

Aćimović, M., Stanković-Jeremić, J., and Cvetković, M. (2020). Phyto-pharmacological aspects of *Nepeta nuda* L.: A systematic review. *Lek sirov.* 40, 75–83. doi: 10.5937/leksir2040075A

Aćimović, M., Lončar, B., Pezo, M., Stanković Jeremić, J., Cvetković, M., Rat, M., et al. (2022). Volatile compounds of *Nepeta nuda* L. *Rtanj Mountain (Serbia). Horticulturae.* 8, 85. doi: 10.3390/horticulturae8020085

Aizen, M. A., and Harder, L. D. (2009). The global stock of domesticated honey bees is growing slower than agricultural demand for pollination. *Curr. Biol.* 19, 915–918. doi: 10.1016/j.cub.2009.03.071

Akbaba, E., Bagci, E., Hritcu, L., and Maniu, C. (2021). Improvement of memory deficits via acetylcholinesterase inhibitory activity of Nepeta nuda ssp. nuda essential oil in rats. *Kuwait J. Sci.* 24, 48(3). doi: 10.48129/kjs.v48i3.9217

Alagawany, M., Abd El-Hack, M. E., Farag, M. R., Gopi, M., Karthik, K., Malik, Y. S., et al. (2017). Rosmarinic acid: modes of action, medicinal values and health benefits. *Anim. Health Res. Rev.* 18, 167–176. doi: 10.1017/S1466252317000081

Alim, A., Goze, I., Cetin, A., Atas, A. D., Cetinus, S. A., and Vural, N. (2009). Chemical composition and in *vitro* antimicrobial and antioxidant activities of the essential oil of *Nepeta nuda* L. subsp. Albiflora (Boiss.) gams. *Afr J. Microbiol. Res.* 3, 463–467. https://hdl.handle.net/20.500.12418/10083

Alimpić Aradski, A., Oalđe Pavlović, M., Janošević, D., Todorović, S., Gašić, U., Mišić, D., et al. (2023). Leaves micromorphology, chemical profile, and bioactivity of in *vitro*-propagated *Nepeta cyrenaica* (Lamiaceae). *Phytochem. Anal.* 34, 661–679. doi: 10.1002/pca.3257

Alipieva, K., Kokubun, T., Taskova, R., Evstatieva, L., and Handjieva, N. (2007). LC–ESI-MS analysis of iridoid glucosides in *Lamium* species. *Biochem. Syst. Ecol.* 35, 17–22. doi: 10.1016/j.bse.2006.07.004

Alipieva, K. I., Taskova, R. M., Evstatieva, L. N., Handjieva, N. V., and Popov, S. S. (2003). Benzoxazinoids and iridoid glucosides from four *Lamium* species. *Phytochem.* 64, 1413–1417. doi: 10.1016/j.phytochem.2003.08.001

Amoah, S., Sandjo, L., Kratz, J., and Biavatti, M. (2016). Rosmarinic acid–pharmaceutical and clinical aspects. *Planta medica.* 82, 388–406. doi: 10.1055/s-0035-1568274

Aničić, N., Gašić, U., Lu, F., Ćirić, A., Ivanov, M., Jevtić, B., et al. (2021). Antimicrobial and immunomodulating activities of two endemic *nepeta* species and their major iridoids isolated from natural sources. *Pharmaceuticals.* 14, 414. doi: 10.3390/ph14050414

Aničić, N., Matekalo, D., Skorić, M., Gašić, U., Nestorović, Ž.J., Dmitrović, S., et al. (2024). Functional iridoid synthases from iridoid producing and non-producing *Nepeta* species (subfam. Nepetoidae, fam. Lamiaceae). *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1211453

Aničić, N., Matekalo, D., Skorić, M., Pećinar, I., Brkušanin, M., Živković, J. N., et al. (2018). Trichome-specific and developmentally regulated biosynthesis of nepetalactones in leaves of cultivated *Nepeta rtanjensis* plants. *Ind. Crops Prod.* 117, 347–358. doi: 10.1016/j.indcrop.2018.03.019

Aničić, N., Matekalo, D., Skorić, M., Živković, J. N., Petrović, L., Dragićević, M., et al. (2020). Alterations in nepetalactone metabolism during polyethylene glycol (PEG)-induced dehydration stress in two *Nepeta* species. *Phytochem.* 174, 112340. doi: 10.1016/j.phytochem.2020.112340

Aras, A., Bursal, E., and Dogru, M. (2016). UHPLC-ESI-MS/MS analyses for quantification of phenolic compounds of *Nepeta nuda* subsp. Lydiae. *J. App Pharm. Sci.* 6, 009–013. doi: 10.7324/JAPS.2016.601102

Bais, H. P., Walker, T. S., Schweizer, H. P., and Vivanco, J. M. (2002). Root specific elicitation and antimicrobial activity of rosmarinic acid in hairy root cultures of *Ocimum basilicum*. *Plant Physiol. Biochem.* 40, 983–995. doi: 10.1016/S0981-9428(02)01460-2

Başar, Y., Yenigün, S., İpek, Y., Behçet, L., Gül, F., Özen, T., et al. (2023). DNA protection, molecular docking, enzyme inhibition and enzyme kinetic studies of 1, 5, 9-epideoxyloganic acid isolated from *Nepeta aristata* with bio-guided fractionation. *J. Biomol. Struct. Dyn.* 42, 1–14. doi: 10.1080/07391102.2023.2250461

Bernier, U. R., Furman, K. D., Kline, D. L., Allan, S. A., and Barnard, D. R. (2005). Comparison of contact and spatial repellency of catnip oil and N, N-diethyl-3-methylbenzamide (DEET) against mosquitoes. *J. Med. Entomol.* 42, 306–311. doi: 10.1603/0022-2585(2005)042[0306:cocasr]2.0.co;2

Birkett, M. A., Hassanali, A., Hoglund, S., Pettersson, J., and Pickett, J. A. (2011). Repellent activity of catmint, *Nepeta cataria*, and iridoid nepetalactone isomers against Afro-tropical mosquitoes, ixodid ticks and red poultry mites. *Phytochem.* 72, 109–114. doi: 10.1016/j.phytochem.2010.09.016

Birkett, M. A., and Pickett, J. A. (2003). Aphid sex pheromones: from discovery to commercial production. *Phytochemistry* 62, 651–656. doi: 10.1016/s0031-9422(02)00568-x

Bozari, S., Agar, G., Aksakal, O., Erturk, F. A., and Yanmis, D. (2013). Determination of chemical composition and genotoxic effects of essential oil obtained from *Nepeta*

nuda on *Zea mays* seedlings. *Toxicol. Ind. Health* 29, 339–348. doi: 10.1177/0748233711433939

Bozek, M. (2003). Pollen efficiency and foraging by insect pollinators in three catnip [*Nepeta* L.] species. *J. Apic Sci.* 47, 2. bwmeta1.element.agro-article-c15f86b6-2561-4147-b5ac-f924510b96c8.

Bozok, F., Cenet, M., Sezer, G., and Ulukanli, Z. (2017). Essential oil and bioherbicidal potential of the aerial parts of *Nepeta nuda* subsp. albiflora (Lamiaceae). *J. Essent. Oil-Bear. Plants.* 20, 148–154. doi: 10.1080/0972060X.2016.1264279

Brown, P. D., Tokuhisa, J. G., Reichelt, M., and Gershenzon, J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochem.* 62, 471–481. doi: 10.1016/s0031-9422(02)00549-6

Cavill, G. W. K., and Clark, D. V. (1967). Insect venoms, attractants, and repellents—VIII. Isodihydronepetalactone. *J. Insect Physiol.* 13, 131–135. doi: 10.1016/0022-1910(67)90009-1

Chalchat, J. C., Garry, R. P., and Michet, A. (1998). Essential oils of myrtle (*Myrtus communis* L.) of the Mediterranean littoral. *J. Essent. Oil Res.* 10, 613–617. doi: 10.1080/10412905.1998.9700988

Chauhan, K. R., Khanna, H., Bathini, N. B., Le, T. C., and Grieco, J. (2014). Biobased lactams as novel arthropod repellents. *Nat. Prod. Commun.* 9, 1934578X1400901201. doi: 10.1177/1934578X1400901201

Chauhan, K. R., and Zhang, A. (2008). Methods of separating Z, E-nepetalactone and Z, E-nepetalactone from catnip oil.

Cruden, R. W., and Miller-Ward, S. (1981). Pollen-ovule ratio, pollen size, and the ratio of stigmatic area to the pollen-bearing area of the pollinator: an hypothesis. *Evolution.* 35, 964–974. doi: 10.2307/2407867

Cushnie, T. T., and Lamb, A. J. (2005). Antimicrobial activity of flavonoids. *Int. J. Antimicrob. Agents.* 26, 343–356. doi: 10.1016/j.ijantimicag.2005.09.002

Damle, M. S., Giri, A. P., Sainani, M. N., and Gupta, V. S. (2005). Higher accumulation of proteinase inhibitors in flowers than leaves and fruits as a possible basis for differential feeding preference of *Helicoverpa armigera* on tomato (*Lycopersicon esculentum* Mill, Cv. Dhanashree). *Phytochem.* 66, 2659–2667. doi: 10.1016/j.phytochem.2005.09.006

Darrag, H. M., Almuhanna, H. T., and Hakami, E. H. (2022). Secondary metabolites in basil, bio-insecticide, inhibition effect, and in silico molecular docking against proteolytic enzymes of the red palm weevil (*Rhynchophorus ferrugineus*). *Plants.* 11, 1087. doi: 10.3390/plants11081087

De Pooter, H. L., Nicolai, B., De Laet, J., De Buyck, L. F., Schamp, N. M., and Goetghebeur, P. (1988). The essential oils of five *Nepeta* species. A preliminary evaluation of their use in chemotaxonomy by cluster analysis. *Flavour Fragr J.* 3, 155–159. doi: 10.1002/ffj.2730030404

Dienaitė, L., Pukalskienė, M., Matias, A. A., Pereira, C. V., Pukalskas, A., and Venskutonis, P. R. (2018). Valorization of six *Nepeta* species by assessing the antioxidant potential, phytochemical composition and bioactivity of their extracts in cell cultures. *J. Func. Foods.* 45, 12–522. doi: 10.1016/j.jff.2018.04.004

Dixon, R. A. (2001). Natural products and plant disease resistance. *Nature* 411, 843–847. doi: 10.1038/35081178

Dötterl, S., and Gershenzon, J. (2023). Chemistry, biosynthesis and biology of floral volatiles: roles in pollination and other functions. *Nat. Prod. Rep.* 40, 1901–1937. doi: 10.1039/D3NP00024A

Effmert, U., Große, J., Röse, U. S. R., Ehrig, F., Kägi, R., and Piechulla, B. (2005). Volatile composition, emission pattern, and localization of floral scent emission in Mirabilis Jalapa (Nyctaginaceae). *Am. J. Bot.* 92, 2–12. doi: 10.3732/ajb.92.1.2

Eisenbraun, E. J., Sullins, D. W., Browne, C. E., and Shoolery, J. N. (1988). (4aS, 7S, 7aR)-Nepetalactam and (4aS, 7S, 7aR)-2-[(3R, 4R, 4aR, 7S, 7aR)-octahydro-4, 7-dimethyl-1-oxocyclopenta [c] pyran-3-yl] nepetalactam. Nitrogen analogs of nepetalactone and nepetalic. psi.-anhydride. *J. Org. Chem.* 53, 3968–3972. doi: 10.1021/jo00252a016

El-Moaty, H. I. A. (2010). Essential oil and iridoid glycosides of *Nepeta septemcrenata* Erenb. *J. Natural Products* 3, 103–111. Available at: https://www.scinapse.io/papers/2188484169.

Erkan, N., Ayranci, G., and Ayranci, E. (2008). Antioxidant activities of rosemary (*Rosmarinus Officinalis* L.) extract, blackseed (*Nigella sativa* L.) essential oil, carnosic acid, rosmarinic acid and sesamol. *Food Chem.* 110, 76–82. doi: 10.1016/j.foodchem.2008.01.058

Feaster, J. E., Scialdone, M. A., Todd, R. G., Gonzalez, Y. I., Foster, J. P., and Hallahan, D. L. (2009). Dihydronepetalactones deter feeding activity by mosquitoes, stable flies, and deer ticks. *J. Med. Entomol.* 46, 832–840. doi: 10.1603/033.046.0413

Fernández-Grandon, G. M., and Poppy, G. M. (2015). Response of Aphidius colemani to aphid sex pheromone varies depending on plant synergy and prior experience. *Bull. Entom Res.* 105, 507–514. doi: 10.1017/S0007485315000371

Formisano, C., Rigano, D., and Senatore, F. (2011). Chemical constituents and biological activities of Nepeta species. *Chem. Biodivers.* 8, 1783–1818. doi: 10.1002/cbdv.201000191

Gasic, K., Hernandez, A., and Korban, S. S. (2004). RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. *Plant Mol. Biol. Rep.* 22, 437–438. doi: 10.1007/BF02772687

Gkinis, G., Bozin, B., Mimica-Dukic, N., and Tzakou, O. (2010). Antioxidant activity of *Nepeta nuda* L. ssp. nuda essential oil rich in nepetalactones from Greece. *J. medicinal Food* 13, 1176–1181. doi: 10.1089/jmf.2009.0218

Glinwood, R. T., Du, Y. J., Smiley, D. W. M., and Powell, W. (1999). Comparative responses of parasitoids to synthetic and plant-extracted nepetalactone component of aphid sex pheromones. *J. Chem. Ecol.* 25, 1481–1488. doi: 10.1023/A:1020872412891

Godschalx, A. L., Stady, L., Watzig, B., and Ballhorn, D. J. (2016). Is protection against florivory consistent with the optimal defense hypothesis? *BMC Plant Biol.* 16, 1–9. doi: 10.1186/s12870-016-0719-2

Goldansaz, S. M., Festa, C., Pagano, E., De Marino, S., Finamore, C., Parisi, O. A., et al. (2019). Phytochemical and biological studies of *Nepeta asterotricha* Rech. f. (Lamiaceae): Isolation of nepetamoside. *Molecules* 24, 1684. doi: 10.3390/molecules24091684

Gomes, E. N., Yuan, B., Patel, H. K., Lockhart, A., Wyenandt, C. A., Wu, Q., et al. (2024). Implications of the Propagation Method for the Phytochemistry of *Nepeta cataria* L. throughout a Growing Season. *Molecules* 29, 2001. doi: 10.3390/molecules29092001

Gormez, A., Bozari, S., Yanmis, D., Gulluce, M., Agar, G., and Sahin, F. (2013). Antibacterial activity and chemical composition of essential oil obtained from *Nepeta nuda* against phytopathogenic bacteria. *J. Essent. Oil Res.* 25, 149–153. doi: 10.1080/10412905.2012.751060

Guiso, M., and Martino, C. (1983). 6-Deoxylamioside, a new iridoid glucoside from *Lamium amplexicaule*. *J. Nat. Prod.* 46, 157–160. doi: 10.1021/np50026a002

Hadi, N., Sefidkon, F., Shojaeiyan, A., Šiler, B., Jafari, A. A., Aničić, N., et al. (2017). Phenolics' composition in four endemic *Nepeta* species from Iran cultivated under experimental field conditions: The possibility of the exploitation of *Nepeta* germplasm. *Ind. Crops Prod.* 95, 475–484. doi: 10.1016/j.indcrop.2016.10.059

Hammer, Ø., and Harper, D. A. (2001). Past: paleontological statistics software package for educaton and data anlysis. *Palaeontologia electronica* 4, 1. Available at: http://palaeo-electronica.org/2001_1/past/issue1_01.htm.

Handjieva, N. V., Popov, S. S., and Evstatieva, L. N. (1996). Constituents of Essential Oils from *Nepeta cataria* L., N. grandiflora MB and *N. nuda* L. *J. Essent. Oil Res.* 8, 639–643. doi: 10.1080/10412905.1996.9701032

Harney, J. W., Barofsky, I. M., and Leary, J. D. (1978). Behavioral and toxicological studies of cyclopentanoid monoterpenes from *Nepeta cataria*. *Lloydia* 41, 367–374.

Hernaíndez Lozada, N. J., Hong, B., Wood, J. C., Caputi, L., Basquin, J., Chuang, L., et al. (2022). Biocatalytic routes to stereo-divergent iridoids. *Nat. Commun.* 13, 1–13. doi: 10.1038/s41467-022-32414-w

Ivanov, M., Kostić, M., Stojković, D., and Soković, M. (2022). Rosmarinic acid–modes of antimicrobial and antibiofilm activities of a common plant polyphenol. *S AFR J. Bot.* 146, 521–527. doi: 10.1016/j.sajb.2021.11.050

Jamzad, Z., Grayer, R. J., Kite, G. C., Simmonds, M. S., Ingrouille, M., and Jalili, A. (2003). Leaf surface flavonoids in Iranian species of *Nepeta* (Lamiaceae) and some related genera. *Biochem. Syst. Ecol.* 31, 587–600. doi: 10.1016/S0305-1978(02)00221-1

Junker, R. R. (2016). "Multifunctional and diverse floral scents mediate biotic interactions embedded in communities," in *Deciphering Chemical Language of Plant Communication. Signaling and Communication in Plants*. Eds. J. Blande and R. Glinwood (Springer, Cham). doi: 10.1007/978-3-319-33498-1_11

Kessler, A., and Kalske, A. (2018). Plant secondary metabolite diversity and species interactions. *Annu. Rev. Ecol. Evol. Syst.* 49, 115–138. doi: 10.1146/annurev-ecolsys-110617-062406

Khan, S., Taning, C. N. T., Bonneure, E., Mangelinckx, S., Smagghe, G., Ahmad, R., et al. (2019). Bioactivity-guided isolation of rosmarinic acid as the principle bioactive compound from the butanol extract of Isodon rugosus against the pea aphid, *Acyrthosiphon pisum*. *PLoS One* 14, e0215048. doi: 10.1371/journal.pone.0215048

Kilic, O., Hayta, S., and Bagci, E. (2011). Chemical composition of essential oil of *Nepeta nuda* L. subsp. nuda (Lamiaceae) from Turkey. *Asian J. Chem.* 23, 2788.

Klocke, J. A., Darlington, M. V., and Balandrin, M. F. (1987). 1, 8-Cineole (Eucalyptol), a mosquito feeding and ovipositional repellent from volatile oil of *Hemizonia fitchii* (Asteraceae). *J. Chem. Ecol.* 13, 2131–2141. doi: 10.1007/BF01012562

Kobaisy, M., Tellez, M. R., Dayan, F. E., Mamonov, L. K., Mukanova, G. S., Sitpaeva, G. T., et al. (2005). Composition and phytotoxic activity of *Nepeta pannonica* L. essential oil. *J. Essent. Oil Res.* 17, 704–707. doi: 10.1080/10412905.2005.9699037

Kökdil, G., Kurucu, S., and Topçu, G. (1996). Composition of the essential oil of *Nepeta nuda* L. ssp. albiflora (Boiss.) Gams. *Flavour Fragr J.* 11, 167–169. doi: 10.1002/(SICI)1099-1026(199605)11:3%3C167::AID-FFJ567%3E3.0.CO,2-A

Kökdil, G., Kurucu, S., and Yıldız, A. (1998). Essential oil composition of *Nepeta nuda* L. ssp. nuda. *Flavour Fragr J.* 13, 233–234. doi: 10.1002/(SICI)1099-1026(1998070)13:4%3C233::AID-FFJ730%3E3.0.CO,2-7

Kökdil, G., Yalcin, S. M., and Topcu, G.Ü.L.A. Ç. T. I. (1999). Nepetalactones and other constituents of Nepeta nuda ssp. albiflora. *Turk. J. Chem.* 23, 99–104. https://journals.tubitak.gov.tr/chem/vol23/iss1/14

Labarrere, B., Prinzing, A., Dorey, T., Chesneau, E., and Hennion, F. (2019). Variations of secondary metabolites among natural populations of sub-Antarctic ranunculus species suggest functional redundancy and versatility. *Plants* 8, 234. doi: 10.3390/plants8070234

Li, J., Hu, H., Mao, J., Yu, L., Stoopen, G., Wang, M., et al. (2019). Defense of pyrethrum flowers: repelling herbivores and recruiting carnivores by producing aphid alarm pheromone. *New phytologist.* 223, 1607–1620. doi: 10.1111/nph.15869

Lichman, B. R., Godden, G. T., Hamilton, J. P., Palmer, L., Kamileen, M. O., Zhao, D., et al. (2020). The evolutionary origins of the cat attractant nepetalactone in catnip. *Sci. Adv.* 6, eaba0721. doi: 10.1126/sciadv.aba0721

Lichman, B. R., Kamileen, M. O., Titchiner, G. R., Saalbach, G., Stevenson, C. E. M., Lawson, D. M., et al. (2019a). 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., O'Connor, S. E., and Kries, H. (2019b). Biocatalytic strategies towards [4 + 2] cycloadditions. *Chem. - A Eur. J.* 25, 6864–6877. doi: 10.1002/chem.201805412

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2– ΔΔCT method. *methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Mamadalieva, N. Z., Sharopov, F. S., Satyal, P., Azimova, S. S., and Wink, M. (2016). Chemical composition of the essential oils of some Central Asian *Nepeta* species (Lamiaceae) by GLC-MS. *Nat. Prod. Commun.* 11, 1934578X1601101229. doi: 10.1177/1934578X1601101229

Mancini, E., Apostolides Arnold, N., De Feo, V., Formisano, C., Rigano, D., Piozzi, F., et al. (2009). Phytotoxic effects of essential oils of *Nepeta curviflora* Boiss. and *Nepeta nuda* L. subsp. albiflora growing wild in Lebanon. *J. Plant Interact.* 4, 253–259. doi: 10.1080/17429140903225507

Marak, H. B., Biere, A., and Van Damme, J. M. (2002). Two herbivore-deterrent iridoid glycosides reduce the *in-vitro* growth of a specialist but not of a generalist pathogenic fungus of *Plantago lanceolata* L. *Chemoecology.* 12, 185–192. Available at: https://www.jstor.org/stable/3448829.

Matsuno, M., Nagatsu, A., Ogihara, Y., Ellis, B. E., and Mizukami, H. (2002). CYP98A6 from Lithospermum erythrorhizon encodes 4-coumaroyl-4′-hydroxyphenyllactic acid 3-hydroxylase involved in rosmarinic acid biosynthesis. *FEBS Lett.* 514, 219–224. doi: 10.1016/S0014-5793(02)02368-2

McCall, A. C., and Irwin, R. E. (2006). Florivory: the intersection of pollination and herbivory. *Ecol.* 9, 1351–1365. doi: 10.1111/j.1461-0248.2006.00975.x

Mišić, D., Siler, B., Gašić, U., Avramov, S., Zivković, S., Nestorović Živković, J., et al. (2015). Simultaneous UHPLC/DAD/(+/-)HESI-MS/MS analysis of phenolic acids and nepetalactones in methanol extracts of Nepeta species: a possible application in chemotaxonomic studies. *Phytochem. Anal.* 26, 72–85. doi: 10.1002/pca.2538

Moghadam, S. E., Ebrahimi, S. N., Gafner, F., Ochola, J. B., Marubu, R. M., Lwande, W., et al. (2015). Metabolite profiling for caffeic acid oligomers in Satureja biflora. *Ind. Crops Prod.* 76, 892–899. doi: 10.1016/j.indcrop.2015.07.059

Moore, B. D., Andrew, R. L., Külheim, C., and Foley, W. J. (2014). Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytol.* 201, 733–750. doi: 10.1111/nph.12526

Moretti, A. N., Zerba, E. N., and Alzogaray, R. A. (2015). Lethal and sublethal effects of eucalyptol on T riatoma infestans and R hodnius prolixus, vectors of C hagas disease. *Entomol. Exp. Appl.* 154, 62–70. doi: 10.1111/eea.12256

Muñoz-Muñoz, J. L., Garcia-Molina, F., Ros, E., Tudela, J., García-Canovas, F., and Rodriguez-Lopez, J. N. (2013). Prooxidant and antioxidant activities of rosmarinic acid. *J. Food Biochem.* 37, 396–408. doi: 10.1111/j.1745-4514.2011.00639.x

Murai, F., Tagawa, M., Damtoft, S., Jensen, S. R., and Nielsen, B. J. (1984). (1R, 5R, 8S, 9S)-Deoxyloganic acid from *Nepeta cataria*. Chem. *Pharm. Bull.* 32, 2809–2814. doi: 10.1248/cpb.32.2809

Musso, L., Scaglia, B., Haj, G. A., Arnold, N. A., Adani, F., Scarì, G., et al. (2017). Chemical characterization and nematicidal activity of the essential oil of *Nepeta nuda* L. ssp. pubescens and *Nepeta curviflora* Boiss. from Lebanon. *J. Essent. Oil-Bear. Plants.* 20, 1424–1433. doi: 10.1080/0972060X.2017.1407678

Nadeem, M., Imran, M., Aslam Gondal, T., Imran, A., Shahbaz, M., Muhammad Amir, R., et al. (2019). Therapeutic potential of rosmarinic acid: A comprehensive review. *Appl. Sci.* 9, 3139. doi: 10.3390/app9153139

Nagy, T., Kocsis, Á., Morvai, M., Szabó, L., Podányi, B., Gergely, A., et al. (1998). 2′-, 4′-, and 6′-O-substituted 1, 5, 9-epideoxyloganic acids from Nepeta grandiflora. *Phytochemistry.* 47, 1067–1072. doi: 10.1016/S0031-9422(98)80074-5

Narimani, R., Moghaddam, M., Ghasemi Pirbalouti, A., and Mojarab, S. (2017). Essential oil composition of seven populations belonging to two *Nepeta* species from Northwestern Iran. *Int. J. Food Prop.* 20, 2272–2279. doi: 10.1080/10942912.2017.1369104

Palmer-Young, E. C., Farrell, I. W., Adler, L. S., Milano, N. J., Egan, P. A., Junker, R. R., et al. (2019). Chemistry of Floral Rewards: Intra-and Interspecific Variability of Nectar and Pollen Secondary Metabolites across Taxa. *Ecol. Monogr.* 89, 1–23. Available at: https://www.jstor.org/stable/26641226.

Patel, H., Gomes, E. N., Yuan, B., Lyu, W., Wu, Q., and Simon, J. E. (2022). Investigation of volatile iridoid terpenes in *Nepeta cataria* L. (Catnip) Genotypes. *Molecules.* 27, 7057. doi: 10.3390/molecules27207057

Petersen, M., Abdullah, Y., Benner, J., Eberle, D., Gehlen, K., Hücherig, S., et al. (2009). Evolution of rosmarinic acid biosynthesis. *Phytochemistry.* 70, 1663–1679. doi: 10.1016/j.phytochem.2009.05.010

Petrova, D., Gašić, U., Yocheva, L., Hinkov, A., Yordanova, Z., Chaneva, G., et al. (2022). Catmint (*Nepeta nuda* L.) phylogenetics and metabolic responses in variable growth conditions. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.866777

Petrović, L., Skorić, M., Šiler, B., Banjanac, T., Gašić, U., Matekalo, D., et al. (2024). Patterns of genetic variation of *nepeta nuda* L. from the central balkans: understanding drivers of chemical diversity. *Plants.* 13, 1483. doi: 10.3390/plants13111483

Pichersky, E., and Gershenzon, J. (2002). The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Curr. Opin. Plant Biol.* 5, 237–243. doi: 10.1016/s1369-5266(02)00251-0

Rabee, M., Andersen, Ø.M., Fossen, T., Enerstvedt, K. H., Abu Ali, H., and Rayyan, S. (2020). Acylated flavone O-glucuronides from the aerial parts of *nepeta curviflora*. *Molecules.* 25, 3782. doi: 10.3390/molecules25173782

Regnier, F. E., Waller, G. R., and Eisenbraun, E. J. (1967). Studies on the composition of the essential oils of three *Nepeta* species. *Phytochemistry.* 6, 1281–1289. doi: 10.1016/S0031-9422(00)86089-6

Reichert, W., Ejercito, J., Guda, T., Dong, X., Wu, Q., Ray, A., et al. (2019). Repellency assessment of *Nepeta cataria* essential oils and isolated nepetalactones on *Aedes aEgypti*. *Sci. Rep.* 9, 1524. doi: 10.1038/s41598-018-36814-1

Sastry, S. D., Springstube, W. R., and Waller, G. R. (1972). Identification of 5, 9-dehydronepetalactone, a new monoterpene from *Nepeta cataria*. *Phytochem.* 11, 453–455. doi: 10.1016/S0031-9422(00)90053-0

Schultz, G., Simbro, E., Belden, J., Zhu, J., and Coats, J. (2004). Catnip, *Nepeta cataria* (Lamiales: Lamiaceae)—A closer look: Seasonal occurrence of nepetalactone isomers and comparative repellency of three terpenoids to insects. *Environ. Entomol.* 33, 1562–1569. doi: 10.1603/0046-225X-33.6.1562

Scialdone, M. A., and Liauw, A. Y. (2008). Acetals of Nepetalic acid and method of preparation.

Scriven, R., and Meloan, C. E. (1984).Determining the active component in 1, 3, 3-trimethyl-2-oxabicyclo {2, 2, 2} octane (cineole) that repels the American cockroach, *Periplaneta americana*. Available online at: http://hdl.handle.net/1811/23001 (Accesed May 05, 2024).

Sengupta, S. K., Hutchenson, K. W., Hallahan, D. L., Gonzalez, Y. I., Manzer, L. E., Jackson, S. C., et al. (2018). Hydrogenation of naturally-derived nepetalactone as a topical insect repellent. *ACS Sustain. Chem. Eng.* 6, 9628–9639. doi: 10.1021/acssuschemeng.7b04521

Sharma, A., and Cannoo, D. S. (2013). Phytochemical composition of essential oils isolated from different species of genus *Nepeta* of Labiatae family: a review. *Pharmacophore* 4 (6), 181–211. Available at: https://pharmacophorejournal.com/1sHUSCy.

Sharma, A., Cooper, R., Bhardwaj, G., and Cannoo, D. S. (2021). The genus *Nepeta*: Traditional uses, phytochemicals and pharmacological properties. *J. Ethnopharmacol.* 268, 113679. doi: 10.1016/j.jep.2020.113679

Sherden, N. H., Lichman, B., Caputi, L., Zhao, D., Kamileen, M. O., Buell, C. R., et al. (2018). Identification of iridoid synthases from *Nepeta* species: Iridoid cyclization does not determine nepetalactone stereochemistry. *Phytochem.* 145, 48–56. doi: 10.1016/j.phytochem.2017.10.004

Sih, A., and Baltus, M. S. (1987). Patch size, pollinator behavior, and pollinator limitation in catnip. *Ecology.* 68, 1679–1690. doi: 10.2307/1939860

Simmonds, M. S., Stevenson, P. C., and Hanson, F. E. (2019). Rosmarinic acid in *Canna generalis* activates the medial deterrent chemosensory neurone and deters feeding in the tobacco hornworm *Manduca sexta*. *Physiol. Entomol.* 44, 140–147. doi: 10.1111/phen.12284

Singh, S., Kaur, I., and Kariyat, R. (2021). The multifunctional roles of polyphenols in plant-herbivore interactions. *Int. J. Mol. Sci.* 22, 1442. doi: 10.3390/ijms22031442

Slavković, F., and Bendahmane, A. (2023). Floral phytochemistry: impact of volatile organic compounds and nectar secondary metabolites on pollinator behavior and health. *Chem. Biodivers*, e202201139. doi: 10.1002/cbdv.202201139

Smallegange, R. C., van Loon, J. J. A., Blatt, S. E., Harvey, J. A., Agerbirk, N., and Dicke, M. (2007). Flower vs. leaf feeding by *Pieris brassicae*: glucosinolate-rich flower tissues are preferred and sustain higher growth rate. *J. Chem. Ecol.* 33, 1831–1844. doi: 10.1007/s10886-007-9350-x

Smiljković, M., Dias, M. I., Stojkovic, D., Barros, L., Bukvički, D., and Ferreira, I. C. F. R. (2018). Characterization of phenolic compounds in tincture of edible *Nepeta nuda*: development of antimicrobial mouthwash. *Food Funct.* 9, 5417–5425. doi: 10.1039/c8fo01466c

Sparks, J. T., Botsko, G., Swale, D. R., Boland, L. M., Patel, S. S., and Dickens, J. C. (2018). Membrane proteins mediating reception and transduction in chemosensory neurons in mosquitoes. *Front. Physiol.* 9. doi: 10.3389/fphys.2018.01309

Sukontason, K. L., Boonchu, N., Sukontason, K., and Choochote, W. (2004). Effects of eucalyptol on house fly (Diptera: Muscidae) and blow fly (Diptera: Calliphoridae). *Rev. Inst Med. Trop. Sao Paulo.* 46, 97–101. doi: 10.1590/S0036-46652004000200008

Süntar, I., Nabavi, S. M., Barreca, D., Fischer, N., and Efferth, T. (2018). Pharmacological and chemical features of *Nepeta* L. genus: Its importance as a therapeutic agent. *Phytother. Res.* 32, 185–198. doi: 10.1002/ptr.5946

Suthar, M. K., Purohit, P. M., and Saran, P. L. (2021). Molecular cloning, expression and in-silico characterization of rosmarinic acid synthase from *Ocimum tenuiflorum* L. *J. Plant Biochem. Biotechnol.* 30, 317–325. doi: 10.1007/s13562-020-00591-8

Takeda, Y., Morimoto, Y., Matsumoto, T., Honda, G., Tabata, M., Fujita, T., et al. (1995). Nepetanudoside, an iridoid glucoside with an unusual stereostructure from Nepeta nuda ssp. *Albiflora*. *J. Nat. Prod.* 58, 1217–1221. doi: 10.1021/np50122a009

Takeda, Y., Yagi, T., Matsumoto, T., Honda, G., Tabata, M., Fujita, T., et al. (1996). Nepetanudosides and iridoid glucosides having novel stereochemistry from Nepeta nuda ssp. albiflora. *Phytochem.* 42, 1085–1088. doi: 10.1016/0031-9422(96)00074-X

Talebi, S. M., Nohooji, M. G., and Yarmohammadi, M. (2017). Infraspecific variations in essential oil compositions of *Nepeta fissa* from Iran. *Nusantara. Bioscience.* 9, 318–321. doi: 10.13057/nusbiosci/n090313

Talebi, S. M., Nohooji, M. G., Yarmohammadi, M., Khani, M., and Matsyura, A. (2019). Effect of altitude on essential oil composition and on glandular trichome density in three *Nepeta* species (N. sessilifolia, N. heliotropifolia and *N. fissa*). *Mediterr. Botany.* 40, 81–93. doi: 10.5209/mbot.59730

Thiers, B. (2023). *Index Herbariorum: A Global Directory of Public Herbaria and Associated Staff* (New York Botanical Garden's Virtual Herbarium).

Treutter, D. (2005). Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biol.* 7, 581–591. doi: 10.1055/s-2005-873009

Trócsányi, E., György, Z., and Zámboriné-Németh, É. (2020). New insights into rosmarinic acid biosynthesis based on molecular studies. *Curr. Plant Biol.* 23, 100162. doi: 10.1016/j.cpb.2020.100162

Tyagi, K., Shukla, P., Rohela, G. K., Shabnam, A. A., and Gautam, R. (2020). Plant phenolics: their biosynthesis, regulation, evolutionary significance, and role in senescence. *Plant Phenolics Sustain. Agriculture.* 1, 431–449. doi: 10.1007/978-981-19-6426-8_11

Uenoyama, R., Miyazaki, T., Adachi, M., Nishikawa, T., Hurst, J. L., and Miyazaki, M. (2022). Domestic cats damage silver vine and catnip plants containing iridoids and enhance chemical pest defence. *iScience.* 25, 104455. doi: 10.1016/j.isci.2022.104455

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

Ulziibayar, D., Begna, T., Bisrat, D., and Jung, C. (2022). Longevity-enhancing Effects of Rosmarinic Acid Feeding on Honey bees (*Apis mellifera* L.) after Exposure to Some Pesticides Used in Strawberry Greenhouse. *J. Apic.* 2, 135–141. doi: 10.17519/apiculture.2022.06.37.2.135

Velasco-Negueruela, A., Perez-Alonso, M. J., and Rodriquez, A. B. (1989). Essential oil analysis of *Nepeta teydea* Webb. & Berth. *Flavour Fragr J.* 4, 197–199. doi: 10.1002/ffj.2730040409

Vivekanandhan, P., Swathy, K., Bedini, S., and Shivakumar, M. S. (2023). Bioprospecting of *Metarhizium anisopliae* derived crude extract: a ecofriendly insecticide against insect pest. *Int. J. Trop. Insect Sci.* 43, 1–12. doi: 10.1007/s42690-022-00935-y

Vivekanandhan, P., Swathy, K., and Shivakumar, M. S. (2022). Identification of insecticidal molecule aucubin from *Metarhizium anisopliae* ethyl acetate crude extract against disease mosquito vector. *Int. J. Trop. Insect Sci.* 42, 3303–3318. doi: 10.1007/s42690-022-00828-0

Wang, G.-Q., Chen, J.-F., YI, B., Tan, H.-X., Zhang, L., and Chen, W.-S. N. (2017). HPPR encodes the hydroxyphenylpyruvate reductase required for the biosynthesis of hydrophilic phenolic acids in *Salvia miltiorrhiza*. *Chin. J. Nat. Med.* 15, 917–927. doi: 10.1016/S1875-5364(18)30008-6

Wetzel, W. C., and Whitehead, S. R. (2020). The many dimensions of phytochemical diversity: linking theory to practice. *Ecol. Lett.* 23, 16–32. doi: 10.1111/ele.13422

Wink, M. (2018). Plant secondary metabolites modulate insect behavior-steps toward addiction? *Front. Physiol.* 9. doi: 10.3389/fphys.2018.00364

Yarmoohammadi, M., Talebi, S. M., and Nohooji, M. G. (2017). Infraspecific variations in essential oil and glandular trichomes in *Nepeta heliotropifolia*. *Biodiversitas J. Biol. Diversity.* 18, 964–970. doi: 10.13057/biodiv/d180314

Zaharieva, A., Rusanov, K., Rusanova, M., Paunov, M., Yordanova, Z., Mantovska, D., et al. (2023). Uncovering the interrelation between metabolite profiles and bioactivity of *in vitro*- and wild-grown catmint (*Nepeta nuda* L.). *Metabolites.* 13, 1099. doi: 10.3390/metabo13101099

Zangerl, A. R. (2003). Evolution of induced plant responses to herbivores. *Basic Appl. Ecol.* 4, 91–103. doi: 10.1078/1439-1791-00135

Zengin, G., Mahomoodally, M. F., Aktumsek, A., Jekő, J., Cziáky, Z., Rodrigues, M. J., et al. (2021). Chemical profiling and biological evaluation of *Nepeta baytopii* extracts and essential oil: An endemic plant from Turkey. *Plants.* 10, 1176. doi: 10.3390/plants10061176

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 13 December 2024
DOI 10.3389/fpls.2024.1498737

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Yourong Chai,
Southwest University, China
Chaorong Tang,
Hainan University, China
Riza-Arief Putranto,
Universitas Esa Unggul, Indonesia

*CORRESPONDENCE
Christian Schulze Gronover
christian.schulze.gronover@ime.fraunhofer.de

†PRESENT ADDRESS
Natalie Laibach,
Faculty of Life Sciences, Rhine-Waal University of Applied Sciences, Kleve, Germany

RECEIVED 19 September 2024
ACCEPTED 05 November 2024
PUBLISHED 13 December 2024

CITATION
Wolters SM, Laibach N, Riekötter J, Roelfs K-U, Müller B, Eirich J, Twyman RM, Finkemeier I, Prüfer D and Schulze Gronover C (2024) The interaction networks of small rubber particle proteins in the latex of *Taraxacum koksaghyz* reveal diverse functions in stress responses and secondary metabolism. *Front. Plant Sci.* 15:1498737. doi: 10.3389/fpls.2024.1498737

The interaction networks of small rubber particle proteins in the latex of *Taraxacum koksaghyz* reveal diverse functions in stress responses and secondary metabolism

Silva Melissa Wolters[1], Natalie Laibach[1†], Jenny Riekötter[2], Kai-Uwe Roelfs[2], Boje Müller[1], Jürgen Eirich[2], Richard M. Twyman[3], Iris Finkemeier[2], Dirk Prüfer[1,2] and Christian Schulze Gronover[1*]

[1]Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Münster, Germany, [2]Institute of Plant Biology and Biotechnology, University of Münster, Münster, Germany, [3]TRM Ltd, Scarborough, United Kingdom

The Russian dandelion (*Taraxacum koksaghyz*) is a promising source of natural rubber (NR). The synthesis of NR takes place on the surface of organelles known as rubber particles, which are found in latex – the cytoplasm of specialized cells known as laticifers. As well as the enzymes directly responsible for NR synthesis, the rubber particles also contain small rubber particle proteins (SRPPs), the most abundant of which are SRPP3, 4 and 5. These three proteins support NR synthesis by maintaining rubber particle stability. We used homology-based searches to identify the whole *TkSRPP* gene family and qPCR to create their spatial expression profiles. Affinity enrichment-mass spectrometry was applied to identify TkSRPP3/4/5 protein interaction partners in *T. koksaghyz* latex and selected interaction partners were analyzed using qPCR, confocal laser scanning microscopy and heterologous expression in yeast. We identified 17 SRPP-like sequences in the *T. koksaghyz* genome, including three apparent pseudogenes, 10 paralogs arranged as an inverted repeat in a cluster with *TkSRPP3/4/5*, and one separate gene (*TkSRPP6*). Their sequence diversity and different expression profiles indicated distinct functions and the latex interactomes obtained for TkSRPP3/4/5 suggested that TkSRPP4 is a promiscuous hub protein that binds many partners from different compartments, whereas TkSRPP3 and 5 have more focused interactomes. Two interactors shared by TkSRPP3/4/5 (TkSRPP6 and TkUGT80B1) were chosen for independent validation and detailed characterization. TkUGT80B1 triterpenoid glycosylating activity provided first evidence for triterpenoid saponin synthesis in *T. koksaghyz* latex. Based on its identified interaction partners, TkSRPP4 appears to play a special role in the endoplasmic reticulum, interacting with lipidmodifying enzymes that may facilitate rubber particle formation. TkSRPP5 appears to be involved in GTPase-dependent signaling and TkSRPP3 may act as part of a kinase signaling cascade, with roles in stress tolerance. TkSRPP interaction with TkUGT80B1 draws a new connection between TkSRPPs and triterpenoid saponin synthesis in *T. koksaghyz* latex. Our data contribute to the

functional differentiation between TkSRPP paralogs and demonstrate unexpected interactions that will help to further elucidate the network of proteins linking TkSRPPs, stress responses and NR biosynthesis within the cellular complexity of latex.

KEYWORDS

SRPPs, small rubber particle proteins, natural rubber, latex, *Taraxacum koksaghyz*, rubber elongation factor family, triterpenoid saponins, stress response

1 Introduction

The Russian dandelion *Taraxacum koksaghyz* produces large amounts of natural rubber (NR) in its roots and is promising as a new crop for the rubber industry (Salehi et al., 2021). NR is mainly composed of poly(*cis*-1,4-isoprene) produced in the latex, the cytoplasm of specialized cells known as laticifers. Within the latex, NR is stored in organelles known as rubber particles comprising a protein-decorated phospholipid monolayer surrounding a dense NR core (Bae et al., 2020; Collins-Silva et al., 2012; Cornish et al., 1999; Hillebrand et al., 2012; Schmidt et al., 2010b; Wood and Cornish, 2000). Proteins on the rubber particle surface contribute to NR synthesis in *T. koksaghyz* and its close relative *T. brevicorniculatum*, which produces small amounts of NR (Benninghaus et al., 2020; Collins-Silva et al., 2012; Hillebrand et al., 2012; Laibach et al., 2015; Niephaus et al., 2019). These proteins include small rubber particle proteins (SRPPs), the most abundant of which are SRPP3–5 (Collins-Silva et al., 2012; Hillebrand et al., 2012; Schmidt et al., 2010a; Wahler et al., 2012), correlating with the high levels of *SRPP3–5* mRNA in the latex (Benninghaus et al., 2020; Lin et al., 2022; Niephaus et al., 2019). *SRPP* gene silencing in *T. koksaghyz* and *T. brevicorniculatum* caused the depletion of NR and reduced rubber particle stability or NR molecular mass, confirming that SRPPs are needed for efficient NR biosynthesis (Collins-Silva et al., 2012; Hillebrand et al., 2012). Accordingly, several *SRPP* genes are upregulated in the roots of plants that produce large amounts of NR (Panara et al., 2018), and *T. koksaghyz* plants overexpressing the transcription factor MYC2 that induces *SRPP* transcription also accumulate more NR than controls (Wu et al., 2024). Dandelion SRPPs promote NR synthesis by contributing to rubber particle stability and dispersity via steric hindrance, and/or potentially by promoting *cis*-prenyltransferase (CPT) long chain polymerization (Collins-Silva et al., 2012; Hillebrand et al., 2012). The recently published genome assemblies of *T. koksaghyz* revealed 11 *SRPP* paralogs (Lin et al., 2018, 2022) but few studies have considered the entire TkSRPP family (He et al., 2024; Wu et al., 2024). Given that only TkSRPPs 3/4/5 are abundant in rubber particles, these paralogs may have the greatest impact on NR biosynthesis while the others may be involved in stress responses (Laibach et al., 2018).

Despite the clear link between TkSRPPs and NR biosynthesis, detailed information about the functions of individual TkSRPPs is limited and many studies do not refer to specific TkSRPPs and/or use inconsistent nomenclatures (Dong et al., 2023b; He et al., 2024; Mofidi et al., 2024). The functions of TkSRPPs in latex have been proposed based mostly on the characterization of TbSRPPs 1/2/3/4/5. In *Nicotiana benthamiana* cells, TbSRPP1/2/3/4/5 localized to lipid droplets (LDs) and the endoplasmic reticulum (ER), and TbSRPP1 and TbSRPP3 additionally to the cytosol, supporting the hypothesis that rubber particles, like LDs, bud from the ER and that SRPPs might be involved in this process (Cornish et al., 1999; Herman, 2008; Laibach et al., 2018; Wilfling et al., 2014). Furthermore, the expression of *TbSRPP4* and *TbSRPP5* increased LD number and size, respectively (Laibach et al., 2018). A similar effect was observed for LD size in *Arabidopsis thaliana* (Arabidopsis) overexpressing stress-related proteins (SRPs) 1–3, which are homologous to SRPPs (Kim et al., 2016). TbSRPP1–5 form homodimers and heterodimers, and at least TbSRPP3–5 bind to negatively charged lipids, suggesting they interact with rubber particles via pockets of unsaturated phosphatidylcholine (PC) (Laibach et al., 2018). SRPPs may also influence the formation and growth of rubber particles by binding to the minor lipid component phosphatidylinositol (PI) (Bae et al., 2020; Laibach et al., 2018), which causes positive membrane curvature (Harayama and Riezman, 2018).

Dandelion SRPPs belong to the rubber elongation factor (REF) superfamily and share a conserved REF domain, whose function remains unknown, with canonical REF proteins (Dennis and Light, 1989; Laibach et al., 2015; Oh et al., 1999). REF proteins are widespread in plants, even those without latex (Gidda et al., 2013; Horn et al., 2013; Kim et al., 2010, 2004; Seo et al., 2010). In several NR-producing plants other than dandelion, REF family proteins have been associated with NR biosynthesis (Berthelot et al., 2014; Dai et al., 2013; Dennis and Light, 1989; Schmidt et al., 2010b). The *T. brevicorniculatum* major REF protein has a higher molecular mass than TbSRPPs, but is also located on rubber particles. The downregulation of *TbREF* caused NR depletion but did not affect rubber particle stability, therefore suggesting a role in rubber particle biogenesis (Laibach et al., 2015).

In non-rubber plants, like Arabidopsis and avocado (*Persea americana*), REF proteins associate with non-seed LDs (Gidda et al.,

2013, 2016; Horn et al., 2013; Kim et al., 2016). LDs have an architecture similar to rubber particles, but store lipids other than NR, such as triacylglycerol (TAG) or sterols (Fernández-Santos et al., 2020; Gidda et al., 2016; Huang, 2018; Kretzschmar et al., 2020; Murphy, 2011; Slocombe et al., 2009). REF proteins from NR-producers and other plants are also involved in stress responses. SRPPs from *T. brevicorniculatum*, sweet potato (*Ipomoea batatas*), pepper (*Capsicum annuum*) and Arabidopsis conferred drought stress tolerance when overexpressed in tobacco (*Nicotiana tabacum*) or Arabidopsis (Kim et al., 2010, 2016; Kim et al., 2012; Laibach et al., 2018; Seo et al., 2010). The genes are induced by drought or other forms of abiotic stress (Guo et al., 2014; Kim et al., 2010, 2016; Kim et al., 2012; Laibach et al., 2018; Seo et al., 2010), and in some cases also by methyl jasmonate (MeJA), abscisic acid (ABA), ethylene, salicylic acid or wounding, via stress and hormone response elements in the promoter (Cao et al., 2017; Dong et al., 2023a; Fricke et al., 2013; He et al., 2024; Wu et al., 2024).

To characterize the function of TkSRPPs in more detail, we identified their interaction partners, providing insight into their roles in NR biosynthesis, stress tolerance, rubber particle composition and biogenesis, and the metabolic and regulatory networks in latex. We therefore determined the spatial expression patterns of all *TkSRPP* genes, followed by affinity enrichment-mass spectrometry (AE-MS) for TkSRPP3/4/5. Two interactors shared by TkSRPP3/4/5 (TkSRPP6 and TkUGT80B1) were chosen for independent validation and detailed characterization.

2 Materials and methods

2.1 Plant cultivation and tissue processing

We cultivated *T. koksaghyz* and *N. benthamiana* plants under controlled greenhouse conditions (18°C, 16-h photoperiod, 260 PPFD high-pressure sodium lamps with enhanced yellow and red spectrum) as previously described (Unland et al., 2018). *T. koksaghyz* tissues were harvested separately for expression analysis and immediately flash-frozen in liquid nitrogen. After lyophilization, root tissues were pulverized using a ZM 200 Ultra Centrifugal Mill (Retsch, Germany), and leaf tissues were ground under liquid nitrogen with a pestle and mortar. Latex was transferred from cut root surfaces to rubber extraction buffer (REB) [100 mM Tris-HCl pH 7.8, 350 mM sorbitol, 10 mM NaCl, 5 mM $MgCl_2$, 5 mM dithiothreitol (DTT)], flash-frozen and used for RNA extraction without further processing.

2.2 Heterologous production of TkSRPP3/4/5

SRPPs were expressed in *Escherichia coli* BL21Ai (DE3) cells (Thermo Fisher Scientific, USA) transformed with expression vector pET23a(+) containing codon-optimized sequences of *TkSRPP3/4/5* (Supplementary Table S1). Protein expression and purification were carried out as previously described (Laibach et al., 2018).

2.3 AE-MS

Latex was harvested from the roots of 12-week-old *T. koksaghyz* plants (line 203-1-ST) as previously described (Post et al., 2012; Niephaus et al., 2019) with slight modifications (Supplementary Methods S1 in Supplementary Data Sheet 1). For affinity enrichment, 100 µL Ni-NTA agarose (Qiagen, Germany) in a 2-mL tube was washed three times with 500 µL Ni-NTA lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) and centrifuged (500 g, 5 min, 4°C). We dissolved 200 µg recombinant TkSRPP3/4/5 in 1 mL Ni-NTA lysis buffer containing a protease inhibitor cocktail (diluted 1:10), added this to the Ni-NTA agarose beads, and incubated the mixture for 1 h at room temperature (RT), shaking at 200 rpm. The samples were then centrifuged (500 g, 5 min, RT) and the beads were washed three times with Ni-NTA wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0). We diluted 400 µL of the latex or its fractions with 500 µL Ni–NTA wash buffer, added this to the Ni-NTA agarose beads, and repeated the incubation, centrifugation and washing steps as above. After seven further washes with 1 mL Ni-NTA wash buffer, 500 µL of the wash buffer was added to the beads and the samples were analyzed by LC-MS. Beads loaded only with recombinant protein or latex fractions were used as controls. Three replicate samples were prepared.

2.4 LC-MS/MS-based quantitative proteomics

LC-MS/MS data acquisition and processing steps were carried out as previously described (Lassowskat et al., 2017). Briefly, proteins were extracted and digested using a modified filter-assisted sample preparation protocol (FASP). After reduction and alkylation, the samples were digested with trypsin, followed by LC-MS/MS analysis using an EASY-nLC 1200 device coupled to a Q Exactive HF mass spectrometer (both from Thermo Fisher Scientific). For details see Supplementary Methods S2 in Supplementary Data Sheet 1.

Raw data were processed using MaxQuant v1.6.9.0 (Cox and Mann, 2008). MS/MS spectra were assigned to the *T. koksaghyz* proteome (Lin et al., 2018). The sequences of 248 common contaminant proteins and decoy sequences were automatically added. Trypsin specificity was required and a maximum of two missed cleavages was allowed. We set cysteine carbamidomethylation as a fixed modification and methionine oxidation, deamidation of N and Q and protein N-terminal acetylation as variable modifications. We applied a false discovery rate of 1% for peptide spectrum matches and proteins, and enabled matching between runs, label-free quantification (LFQ) and iBAQ.

2.5 Proteomic data analysis and annotation

Proteins were annotated by BLASTP searching against the NCBI non-redundant and UniProt databases ($e \geq 1 \times 10^{-3}$). Data were processed using Perseus v1.6.0.7 and v2.0.11 (Tyanova et al.,

2016). We removed proteins that were only identified by site, reverse hits, or potential contaminants. LFQ intensities were $\log_2$ transformed and proteins were filtered for those with a mean LFQ intensity in the no-prey control lower than the 10th percentile of all data (19.45 for TkSRPP3 runs; 19.91 for TkSRPP4/5 runs) or undetected in this control, as well as proteins quantified in at least two replicate AE-MS runs. Missing values were imputed using the quantile regression imputation of left-censored missing data (QRILC) algorithm in the package imputeLCMD v2.1 (Lazar and Burger, 2022; R Core Team, 2021) and normal distribution was confirmed by consulting histograms (Supplementary Figure S1). Data reproducibility was examined by principal component analysis (PCA). Replicates formed clusters separately from control samples except for rubber phase control 1 from the TkSRPP3 run and pellet phase control 3 from the TkSRPP4/5 run, which were therefore excluded from further analysis (Supplementary Figure S2). For enrichment analysis, we applied two sample *t*-tests between AE-MS and no-bait control samples. A permutation-based FDR (q-value) was calculated to correct for multiple testing. To visualize enriched proteins in volcano plots, q-values equal to zero, representing high confidence, were replaced with the next even value smaller than the smallest calculated q-value within the same approach. Volcano plots and Gene Ontology (GO) heat maps were produced using the ggplot2 v3.5.1 R package (Wickham, 2016). Venn diagrams were generated using InteractiVenn (Heberle et al., 2015). GO terms were assigned using eggNOG-mapper (Cantalapiedra et al., 2021). GO enrichment analysis was carried out using the topGO R package v2.50.0 (Alexa and Rahnenfuhrer, 2022). Protein classes were determined using PANTHER 19.0 (Thomas et al., 2022). Protein abundance data (Benninghaus et al., 2020) were processed and heat maps generated in Perseus v2.0.11 (Tyanova et al., 2016).

2.6 *In silico* sequence analysis

Protein domains were predicted using Interpro (Paysan-Lafosse et al., 2023). Phosphorylation and *N*-glycosylation sites were predicted using CLC Main Workbench v23.0.3 (Qiagen). Phylogenetic trees were constructed using MEGA11 (Tamura et al., 2021) and *cis*-acting regulatory elements were detected using NSITE-PL (Shahmuradov and Solovyev, 2015). The chromosome map was created using MapChart (Voorrips, 2002). Isoelectric points (*pI*) and protein charges were determined using Prot Pi v. 2.2.29.152 (https://www.protpi.ch/).

2.7 Amplification and cloning of *TkSRPP3/4/5* and candidate interactor genes

TkSRPP and *TkGUT80B1* coding sequences were amplified from *T. koksaghyz* latex cDNA with flanking primers based on the genomic sequences (Lin et al., 2018) (Supplementary Table S2). Amplified fragments were digested with restriction enzymes indicated in the primer names and ligated into the Gateway pENTR 4 entry vector (Thermo Fisher Scientific).

For the split-ubiquitin membrane yeast two-hybrid (SUY2H) assays, plasmids pRS313 and pRS314 (Sikorski and Hieter, 1989) were modified to form Gateway destination vectors. The Gateway cassette was amplified from pAG304-P_{GAL1}-*ccdB* (Addgene, USA) using primers M13 rev and attR *Bgl*II fw, and overhangs were prepared by digestion with *Bgl*II and *Kpn*I. The pRS314 interim vector was digested with *Bam*HI and *Kpn*I and ligated with the Gateway cassette. For pRS313, the Gateway cassette was amplified from pBatTL (Jach et al., 2006) using primers attR *Spe*I fw and attR *Age*I rev, and ligated into the digested pRS313 vector. *TkSRPPs, TkUGT80B1* and *mEmerald* were introduced into pRS313-*ccdB-CRU* and pRS314-*Nua-ccdB* using Gateway LR Clonase II mix (Thermo Fisher Scientific). For co-immunoprecipitation (co–IP), *TkSRPP3/4/5* and *TkUGT80B1* were transferred to pAG425-P_{GPD}-*ccdB-Cerulean* and pAG423-P_{GPD}-*ccdB*-HA (Addgene), respectively, by Gateway cloning. For *N. benthamiana* transient expression experiments, genes were introduced into the Gateway-compatible vector pBatTL-*ccdB-Cerulean* as previously described (Epping et al., 2015; Unland et al., 2018). To test TkUGT80B1 activity in yeast, the *TkUGT80B1* coding sequence was inserted into pAG423-P_{GAL1}-*ccdB* (Addgene) by Gateway cloning. All constructs were validated by Sanger sequencing.

2.8 RNA extraction, cDNA synthesis and quantitative PCR

RNA extraction, cDNA synthesis and qPCR were carried out as previously described (Niephaus et al., 2019) with slight modifications (Supplementary Methods S3 in Supplementary Data Sheet 1). Briefly, normalized expression was calculated using the ΔC_q method (Equation 1) relative to the mean C_q value of the reference genes. To account for different primer efficiencies, mean Cq values of technical replicates were adjusted by multiplication with an adjustment coefficient based on primer efficiency (Equation 2).

$$normalized\ \ expression = 2^{mean\ \ adjC_q(reference\ \ genes) - adjC_q(target\ \ gene)} \quad (1)$$

$$adjC_q = C_q \times \left(\frac{-1}{slope} \times \log_2 10 \right) \quad (2)$$

2.9 Subcellular localization studies

Transient expression in *N. benthamiana* leaves was carried out using pBatTL constructs in which the protein of interest was fused to the N-terminus of the blue fluorescent protein Cerulean (Müller et al., 2010). Monomeric red fluorescent protein (mRFP) C-terminally fused to the N-terminal sequence of CYP51G1 (CYP51G1-mRFP) was used to mark the cytosolic surface of the ER (Bassard et al., 2012). Two-pore-channel 1 (TPC1) was fused to orange fluorescent protein (OFP) as a tonoplast marker (Batistič et al., 2010). To test LD localization, we co-expressed Arabidopsis *LEAFY COTYLEDON 2* (*AtLEC2*) to induce LD formation, and infiltrated the leaves with Nile red solution (Santos Mendoza et al., 2005). *AtLEC2* was amplified

from *A. thaliana* cDNA that was obtained as previously described (Jekat et al., 2013) using flanking primers introducing restriction sites for cloning into pENTR4. This construct was used for Gateway cloning into pBatTL-*ccdB* using Gateway LR Clonase II mix (Thermo Fisher Scientific) (Müller et al., 2010). Leaf discs were analyzed by confocal laser scanning microscopy (CLSM) using a Stellaris 8 microscope (Leica Microsystems, Germany). Cerulean fluorescence was detected at 445–550 nm (excitation at 440 nm), mRFP fluorescence at 570–648 nm (excitation at 555 nm), and Nile red fluorescence at 571–587 nm (excitation at 541 nm).

2.10 SUY2H assay

The *Saccharomyces cerevisiae* strain InvSc1 (Thermo Fisher Scientific) was transformed with combinations of pRS313 and pRS314 using the lithium-acetate method (Agatep et al., 1998). Positive clones were identified by colony PCR using gene-specific and vector primers. They were grown for 5–6 h at 30°C in 1 mL synthetic defined (SD) medium containing 50 µM $CuSO_4$ and lacking histidine, tryptophan and methionine. Cultures were centrifuged (19,000 *g*, 1 min, RT) and the OD_{600} was adjusted to 1 using 1× TE. We transferred 10 µL of three serial dilutions to SD medium lacking histidine and tryptophan, or to selective media containing (1) 300 µM methionine, 50 µM $CuSO_4$ and lacking histidine, tryptophan and uracil, or (2) 300 µM methionine, 50 µM $CuSO_4$ and 1 g/L 5-fluoroorotic acid (5-FOA), and lacking histidine and tryptophan. The plates were incubated at 30°C for 2–3 days.

2.11 Co-immunoprecipitation

Yeast strain InvSc1 was transformed with each of the pAG425-P_{GPD}-*TkSRPP*-Cerulean constructs and pAG423-P_{GPD}-*TkUGT80B1*-HA alone or with the three combinations of *TkUGT80B1* and each *TkSRPP*. Positive transformants were identified by colony PCR using gene-specific and vector primers. To test for the interactions between TkUGT80B1, TkSRPP3 and TkSRPP5, 5 mL of SD medium lacking histidine, leucine or both were inoculated with a single colony of each genotype, incubated at 30°C until the OD_{600} reached 3 and harvested by centrifugation (4,000 *g*, 10 min, RT). Cells expressing *TkSRPP4* and a control expressing only *TkUGT80B1* were cultivated at 20°C to enable sufficient recombinant protein synthesis. Overnight cultures were used to inoculate 50 mL SD medium to an OD_{600} of 0.3. The main cultures were then cultivated at 20°C, shaking at 140 rpm. They were harvested by centrifugation after 42 h (4,000 *g*, 10 min, RT) and washed once with 10 mL 1× TE. Proteins were extracted with 1 mL cell lysis buffer comprising 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton X-100, 2.5 mM $Na_4P_2O_7$, 1 mM β–glycerophosphate, 1 mM Na_3VO_4 and 1 µg/ml leupeptin (Cell Signaling Technology, USA) supplemented with 1 mM phenylmethylsulfonylfluoride (PMSF) and cOmplete EDTA-free protease inhibitor cocktail (Merck, Germany) for 1–5 min at 30 Hz in an MM400 bead mill (Retsch) followed by centrifugation (11,000 *g*, 5 min, 4°C). ChromoTek GFP-Trap magnetic agarose (Proteintech Group, USA) was used for immunoprecipitation according to the manufacturer's instructions. Extracts were incubated with the beads for 1 h at 4°C on a platform rocker. Proteins were eluted in 50 µL 5× SDS loading buffer containing 100 mM DTT. Samples were separated by SDS-PAGE on 10% SDS polyacrylamide gels and analyzed as previously described (Niephaus et al., 2019) with modifications. Briefly, membranes were incubated with either an anti-GFP primary antibody (Clontech Laboratories, USA; #632380) diluted 1:2,000, or with an anti-HA primary antibody (Merck; #H3663) diluted 1:1,000–2,000, followed by washing and incubation with either a goat anti-mouse IgG coupled to alkaline phosphate (Merck; #A3562) diluted 1:10,000, followed by detection using SIGMAFAST BCIP/NBT tablets (Merck), or a goat anti-mouse IgG coupled to horseradish peroxidase (Thermo Fisher Scientific; #32430) diluted 1:1,500, followed by detection using SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific).

2.12 TkUGT80B1 activity in yeast

We transformed a *S. cerevisiae* strain, previously engineered for enhanced triterpenoid production and expressing *T. koksaghyz lupeol synthase* (*TkLup*) under the control of a galactose-inducible promoter (Bröker et al., 2018), with pAG423-P_{GAL1}-*TkUGT80B1* or the empty vector using the lithium-acetate method (Agatep et al., 1998). As a control, the same strain lacking *TkLUP* was transformed with *TkUGT80B1* or the empty vector. Positive transformants were identified by colony PCR using gene-specific and vector primers (or two vector primers for the empty vector). For yeast cultivation, we inoculated 5 mL SD medium lacking histidine and tryptophan with a single colony of each genotype and incubated it overnight at 30°C on a rolling wheel. We inoculated 50 mL of the same medium, supplemented with 150 µM $CuSO_4$ to repress sterol synthesis, with the overnight cultures to an OD_{600} of 0.2 and incubated them at 30°C, shaking at 140 rpm. When the cultures reached an OD_{600} of 0.5–0.6, inducible gene expression was activated by switching to SD medium containing 2% galactose instead of glucose. Cells were harvested when cultures reached an OD_{600} of 4. Metabolites were extracted from lyophilized yeast pellets by adding glass beads and 1 mL ethyl acetate, followed by lysis for 30 min at 30 Hz in an MM400 bead mill. After centrifugation (11,000 *g*, 1 min), the supernatant was transferred to a fresh tube and the extraction was repeated twice with 0.5 mL ethyl acetate, each time vortexing for 15 min. The extracts were analyzed by LC-MS/MS using an UltiMate 3000 Rapid Separation System (Thermo Fisher Scientific) and amazon speed ion trap MS (Bruker Corporation, USA) or using an Acquity Premier LC system (Waters Corporation, UK) coupled to a Synapt XS 4k (Waters Corporation) ion mobility time-of-flight mass spectrometer. Extracts were separated using a Reprosil Pur Basic C18 (5 µm particle size, 4×250 mm) (Analytik Altmann, Germany) by isocratic elution with 10% mobile phase A (90:10 *v/v* isopropanol:water + 10 mM ammonium formate) and 90% mobile phase B (methanol + 10 mM ammonium formate). The flow rate was 0.4 mL/min and the run time 45 min. The column temperature was set to 23°C. For the identification of lupeol, an authentic standard was analyzed separately.

10.3389/fpls.2024.1498737

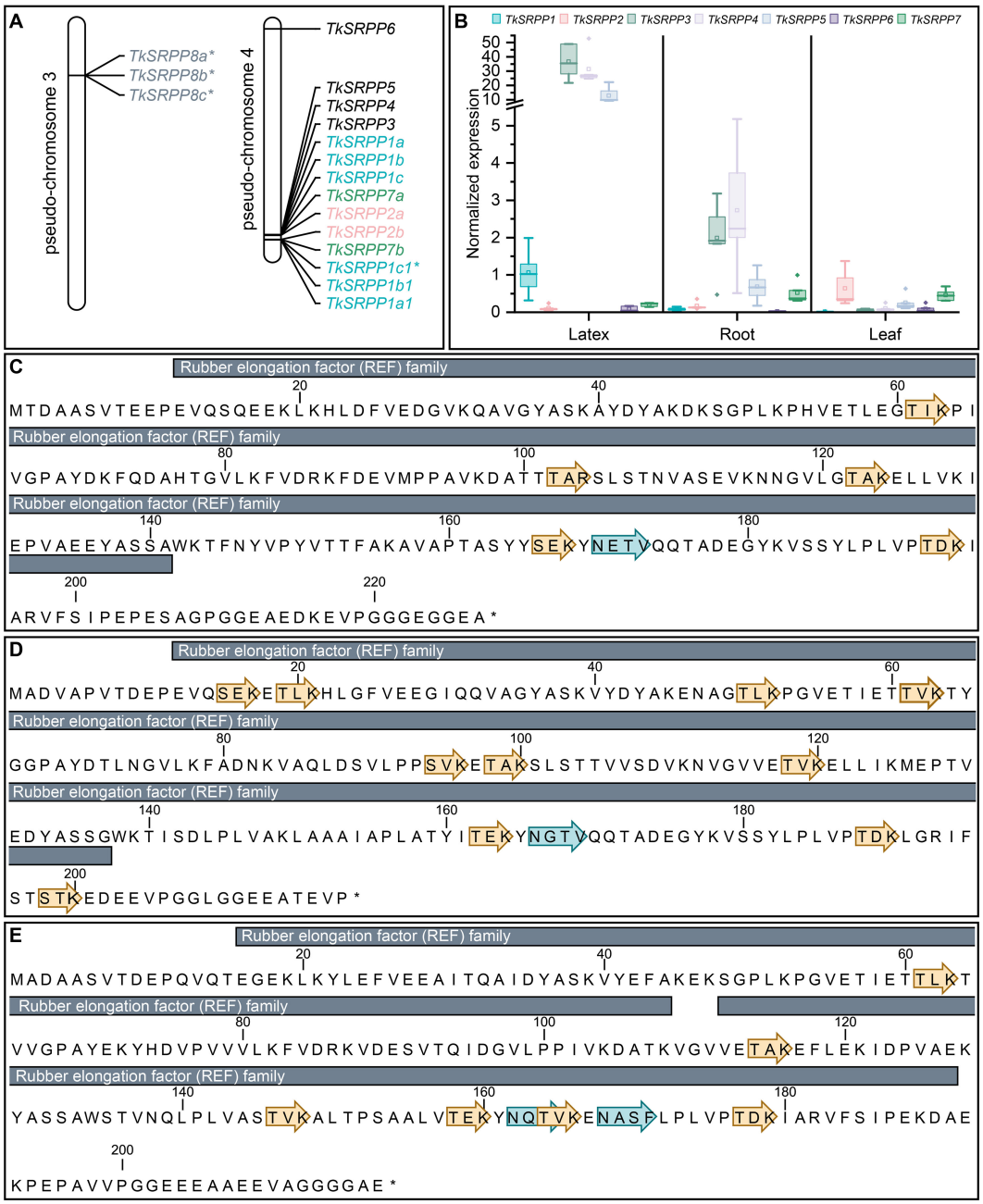


FIGURE 1
The *TkSRPP* gene family. **(A)** Pseudo-chromosome maps showing the loci for the *TkSRPP* family. Asterisks indicate *TkSRPPs* with premature stop codons. **(B)** *TkSRPPs* have different spatial gene expression patterns. Normalized gene expression levels in different tissues of 10-week-old wild-type *T. koksaghyz* plants. Box plots represent data from four or five individual plants. Expression levels of *TkSRPP1, 2* and *7* represent transcripts of all corresponding gene copies. Expression levels were normalized against *elongation factor-1 α* (*TkEF1α*) and *ribosomal protein L27* (*TkRP*). **(C-E)** Protein sequences of TkSRPP3 **(C)**, TkSRPP4 **(D)** and TkSRPP5 **(E)** showing predicted protein domains (InterPro), phosphorylation sites (yellow arrows) and *N*-glycosylation sites (blue arrows*). Asterisks mark the end of the amino acid sequences.

3 Results

3.1 Analysis of the *TkSRPP* gene family in *T. koksaghyz*

We identified 17 *SRPP*-like sequences in the *T. koksaghyz* genome (Lin et al., 2018, 2022). Twelve of them cluster on pseudo-chromosome 4 (Figure 1A). The nomenclature we applied is based on sequence similarity to known orthologs (mainly from *T. brevicorniculatum*) and the first publicly available annotations (Collins-Silva et al., 2012; Hillebrand et al., 2012; Laibach et al., 2018; Schmidt et al., 2010a). Corresponding sequence IDs and comparisons with other published nomenclatures are summarized in Supplementary Table S3. We identified one copy each of *TkSRPP3/4/5* and *TkSRPP6*, the latter being the only full *TkSRPP* gene located outside the cluster, on the other arm of pseudo-chromosome 4. The

most recent genome annotation (Lin et al., 2022) lists *TkSRPP5* as a pseudogene, but the earlier version (Lin et al., 2018) includes an open reading frame, which we confirmed by amplification from cDNA. Similarly, the 2022 genome assembly contains a premature stop codon for TkSRPP6, but amplification of an open reading frame from cDNA confirmed its integrity. We found that *TkSRPP3/4/5* form a contiguous set along with 10 additional paralogs. This indicates an inverted duplication in which the adjacent copies of the first five genes run in the opposite direction. Because the first three contiguous genes are more closely related to each other than other paralogs, we named them *TkSRPP1a-c* and their copies with sequence identities of 93–100% *TkSRPP1a1/b1/c1* respectively (Table 1). However, *TkSRPP1c1* contains a stop codon after 30 bp and has only 93% identity to *TkSRPP1c*, suggesting that functional redundancy has allowed sequence divergence in this case. The next two genes of the cluster were designated *TkSRPP7a* and *TkSRPP2a*, and their copies *TkSRPP7b* and *TkSRPP2b*. *TkSRPP2a* is not annotated in the 2022 genome release, but was identified manually by homology searches. Three additional *TkSRPP* sequences sharing 99.9% identity form another cluster on pseudo-chromosome 3 and were designated *TkSRPP8a-c*.These sequences appeared to originate from incomplete gene transposition and duplication of TkSRPP2a/b, because they contain only the first 654 bp (exon I, intron I, part of exon II) of TkSRPP2a/b which could encode a 50 aa peptide. The high conservation between *TkSRPP8a/b/c* suggests a relatively recent duplication event. Because of this truncation resulting in an incomplete REF domain, we excluded *TkSRPP8* from further analysis. *TkSRPP* gene duplications have been proposed before (He et al., 2024), but our classification differs in that we identified five additional genes and established a nomenclature based on previous publications. The TkSRPPs, sequence identities and predicted protein properties are summarized in Table 1.

We designed primer pairs to quantify *TkSRPP* gene expression in the tissues of wild-type *T. koksaghyz* plants. No discriminating primers could be designed for the proposed gene duplications, so the data for *TkSRPP1*, *TkSRPP2* and *TkSRPP7* reflect the expression of all duplicates. We observed extremely strong *TkSRPP3/4/5* expression in latex (Figure 1B), in agreement with published RNA-Seq data and previously reported high protein levels (Collins-Silva et al., 2012; Lin et al., 2018, 2022; Niephaus et al., 2019). *TkSRPP3* and *TkSRPP4* showed similarly high transcript levels, each about twice the level of *TkSRPP5*. A comparable profile was observed in the roots, which contain considerable amounts of latex (Figure 1B). In leaves, *TkSRPP3/4/5* expression was low. *TkSRPP1* was also predominantly expressed in latex, albeit at level of one tenth or less of that of *TkSRPP3/4/5*. *TkSRPP2* expression notably differed from the others, with the highest level in leaves. *TkSRPP6* was expressed at low levels and *TkSRPP7* at moderate levels in all tissues. These diverging sequences and spatial expression profiles indicate specialized, tissue-specific functions. Because latex is of special interest in *T. koksaghyz*, the remarkably strong expression of *TkSRPP3/4/5* and their contribution to NR biosynthesis prompted us to screen for protein interaction partners of these paralogs in latex (Collins-Silva et al., 2012). We also identified several *N*-glycosylation and phosphorylation sites that could play a role in these interactions and their biological functions (Figures 1C–E).

3.2 Identification of TkSRPP3/4/5 protein interaction partners by AE-MS

Recombinant TkSRPPs with His_6 tags were produced in *E. coli* and coupled to agarose beads before mixing with whole latex as well as the separate rubber phase (RP), interphase (IP) and pellet phase (PP) obtained by centrifugation (Figure 2). The latex phases are thought to comprise different cellular fractions, with the RP mostly containing rubber particles (Collins-Silva et al., 2012; Schmidt et al., 2010b), the PP containing most of the membrane and organelle components of the latex, and the IP containing most of the cytosol. This was supported by the enrichment of GO terms of the cellular component (CC) category in one fraction relative to all detected latex proteins (Supplementary Table S4), enabling us to infer interaction sites in the laticifers and potentially detect interactions with less-abundant proteins that are enriched in particular phases. This is interesting because of potential TkSRPP functions other than rubber particle stabilization and TbSRPP3/4/5 were localized in the ER and cytosol in *N. benthamiana* (Laibach et al., 2018). We used two controls to ensure qualitatively reliable results. First, we used TkSRPP-His_6-coupled beads that were not loaded with latex as no-prey controls for each paralog. Second, uncoupled agarose beads were loaded with each latex fraction as no-bait controls. Proteins were identified by LC-MS/MS and quantified by LFQ using the MaxQuant software suite. The no-bait controls were used to calculate protein enrichments (Figure 2).

Proteins were considered as potential interaction partners when they were significantly enriched compared to the no-bait control (LFQ $\log_2$ fold change (FC) ≥ 1; $q < 0.05$) (Figure 3A). Given the nature of the experiment, the MS data contained a relatively large number of missing values that were mainly assumed to be missing not at random (MNAR). Accordingly, missing values were imputed using QRLIC (Supplementary Figure S1 in Supplementary Data Sheet 1), which performs best for left-censored MNAR data (Wei et al., 2018). Imputation is necessary to enable quantitative statistical analysis, but it can influence data analysis and interpretation. Therefore, we did not strictly exclude proteins from the group of potential interaction partners if they fell outside the defined thresholds for $\log_2$FC and q-value but were subjected to imputation (Figure 3A, blue dots). Our candidate lists thus contain proteins meeting the defined threshold criteria and represent potential interaction partners with the highest confidence according to our data analysis strategy, but we do not limit TkSRPP interaction partners to these proteins (Supplementary Tables S7–S9 in Supplementary Data Sheet 1).

3.3 Characterization of the interactome datasets

For each TkSRPP, four independent datasets were obtained representing whole latex and its three fractions, revealing different numbers of interaction partners in partially overlapping sets (Figures 3B, C). The volcano plots show that the respective bait TkSRPP was one of the most strongly enriched proteins in all AE-MS experiments (Figure 3A). The number of interaction partners identified for TkSRPP5 (83) was much lower than for TkSRPP3

TABLE 1 Comparison of TkSRPP sequences and protein properties.

DNA / aa	TkSRPP1a	TkSRPP1a1	TkSRPP1b	TkSRPP1b1	TkSRPP1c	TkSRPP1c1*	TkSRPP2a	TkSRPP2b	TkSRPP3	TkSRPP4	TkSRPP5	TkSRPP6	TkSRPP7a	TkSRPP7b	TkSRPP8a*	TkSRPP8b*	TkSRPP8c*
TkSRPP1a	100	100	94.7	97.4	96.4	90.1	74	74.4	63.5	56.9	62.5	52.4	73.5	73.5	40.7	40.7	40.7
TkSRPP1a1	100	100	94.7	97.4	96.4	90.1	74	74.4	63.5	56.9	62.5	52.4	73.5	73.5	40.7	40.7	40.7
TkSRPP1b	96.1	96.1	100	96.1	95.1	91	73.5	74	62.6	55.8	62.1	52.4	73.6	73.6	40.7	40.7	40.7
TkSRPP1b1	97.4	97.4	97.8	100	97	90.4	73.5	74	63.7	57.2	62.5	52.9	73.6	73.8	40.9	40.9	40.9
TkSRPP1c	96.6	96.6	95.7	97	100	93.1	74	74.3	62.8	56.5	62.5	54.3	73.5	73.6	40.3	40.3	40.3
TkSRPP1c1*	x	x	x	x	x	100	70	69.8	59.4	54	60	51.3	69.3	69.3	37.7	37.7	37.7
TkSRPP2a	67.3	67.3	67.8	67.3	67.8	x	100	99.4	64.7	60.3	66	55.6	88.2	87.1	45	45	45
TkSRPP2b	67.8	67.8	68.3	67.8	68.3	x	99.5	100	65.2	60.6	66.5	55.2	88.3	87.2	45.2	45.2	45.2
TkSRPP3	50.7	50.7	49.8	51.2	49.3	x	51	51.5	100	71.8	72.7	47.9	66.0	65.7	40.6	40.6	40.6
TkSRPP4	44.6	44.6	43.8	44.6	43.1	x	45	45	64.5	100	68.5	46.3	61.1	61.3	40.2	40.2	40.2
TkSRPP5	51	51	51.4	51.9	52.4	x	55.6	56.2	61.3	59.9	100	52.6	65.6	65.6	41.2	41.2	41.2
TkSRPP6	45.8	45.8	45.6	45.3	44.9	x	48.6	48.6	38.4	37.2	40.2	100	55.4	55	39.9	39.9	39.9
TkSRPP7a	70.2	70.2	70.7	70.2	69.7	x	88.1	87.6	54	48	56.7	51	100	98.6	44.1	44.1	44.1
TkSRPP7b	69.7	69.7	70.2	69.7	69.2	x	86.7	86.2	54.5	48	56.7	51	98.6	100	44.5	44.5	44.5
TkSRPP8a*	66	66	64	64	62	x	98	96	49	46.9	51	44	74	74	100	99.9	99.9
TkSRPP8b*	66	66	64	64	62	x	98	96	49	46.9	51	44	74	74	100	100	99.9
TkSRPP8c*	66	66	64	64	62	x	98	96	49	46.9	51	44	74	74	100	100	100
pI	5.4	5.4	5.4	5.5	5.8	x	8.3	8.3	4.8	4.5	4.6	5.7	8.6	8.3	4.8	4.8	4.8
Charge pH 7.4	-9.3	-9.3	-10,4	-9.3	-9.1	x	1.3	1.3	-13.7	-14.8	-15.8	-6.9	2.3	1.3	-3.7	-3.7	-3.7

Percent identities are shown for TkSRPP DNA and protein sequences. The isoelectric point (*pI*) and protein charge at pH 7.4 are predicted for all paralogs. Sequence identities were determined using Clustal Omega, whereas *pI* and charge were predicted using Prot pi. Asterisks indicate either paralogs containing a premature stop codon (*TkSRPP1c1*) or partial genes (*TkSRPP8a/b/c*) (see text for details).

10.3389/fpls.2024.1498737

10.3389/fpls.2024.1498737

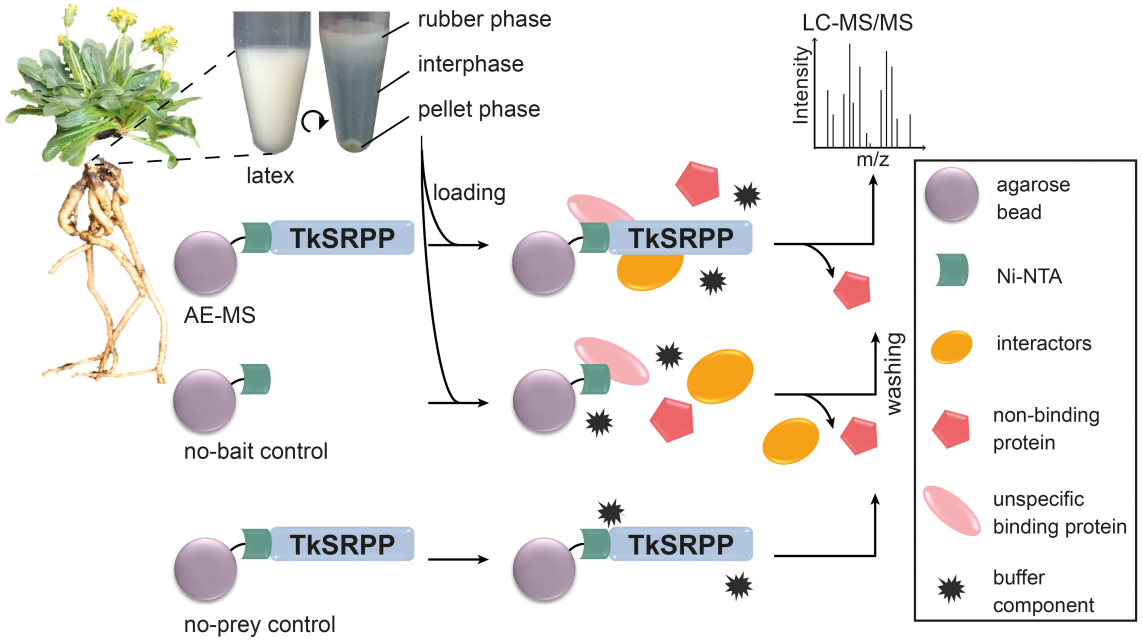


FIGURE 2
Overview of AE-MS experimental design. Latex was harvested from the roots of wild-type *T. koksaghyz* plants and separated by centrifugation. Recombinant TkSRPP3/4/5 expressed in *E. coli* were bound to Ni-NTA agarose beads and loaded with the four separate latex fractions. Fresh agarose beads were loaded with the separate latex fractions as no-bait controls. Agarose beads bound to TkSRPP3/4/5 without exposure to latex fractions served as background (no-sample) controls. All samples were washed and analyzed by LC-MS/MS for the identification of enriched proteins.

(232) and TkSRPP4 (662), this could have been caused by different affinities of TkSRPP3/4/5 to the Ni-NTA agarose beads or the protein stability on the Ni-NTA agarose. Further, it must be distinguished between the total number of hits and the number of different proteins detected within each latex phase. Some of the interactors (16% for TkSRPP3, 19% for TkSRPP4 and 23% for TkSRPP5) were detected under more than one condition, providing more confidence in their veracity (Figure 3B). The large number of interacting proteins enriched exclusively in one phase confirmed that the use of separate latex phases allowed the identification of more specialized interactions and low-abundance interactors. For TkSRPP3, most interacting proteins were enriched from the whole latex (114), followed by the PP (87), IP (50) and RP (25). In contrast, most TkSRPP4 interactors were enriched from the IP (318), followed by the RP (229), PP (185) and whole latex (66, 18 of which were exclusive to whole latex). We found 75 proteins enriched from both the RP and PP, suggesting they are not exclusive to rubber particles but are also found in other organelles. For TkSRPP5, only three interactors were enriched from the PP, with one also enriched from whole latex. Most TkSRPP5 interactors were enriched from whole latex (48) and the IP (39), with 12 enriched from both. The interactomes indicated that TkSRPP4 is a promiscuous hub protein that binds many partners from different compartments, whereas TkSRPP5 interacts more specifically, primarily with proteins present in rubber particles or the cytosol. We also found 10 interactors common to TkSRPP3, 4 and 5 (Figure 3C). TkSRPP4 shared the most interactors with the other paralogs, probably reflecting the presence of more interactors overall. More than half of the proteins interacting with TkSRPP5 also interacted with TkSRPP3 or TkSRPP4, leaving only 37 unique to TkSRPP5. In contrast, most TkSRPP3 and TkSRPP4 interactors were exclusive.

Because RP interactors are of particular interest for the elucidation of the role of TkSRPP3/4/5 in NR biosynthesis, compared the TkSRPP3/4/5 RP interactomes and found that four interactors are shared by TkSRPP4 and TkSRPP5 (Supplementary Table S6) whereas all TkSRPP3 RP interactors are exclusive, again highlighting the functional specialization on the rubber particle, as previously shown on LDs in *N. benthamiana* for TbSRPP4 and TbSRPP5 but not TbSRPP3 (Laibach et al., 2018). The shared TkSRPP4/TkSRPP5 RP interactors comprised Ras-related protein Rab11C, a CBL-interacting protein kinase, α-ketoglutarate-dependent dioxygenase, and a protein similar to an uncharacterized protein from lettuce. Notably, TkSRPP5 was identified as an interaction partner of TkSRPP3 by co-enrichment from whole latex (log_2FC 1.1) and the IP (log_2FC 2.7), but TkSRPP3 was not identified as an interactor when TkSRPP5 was the bait. For the closely related homologs TbSRPP3/4/5, all pairwise interactions have been shown by bimolecular fluorescence complementation (BiFC) (Laibach et al., 2018).

To gain insight into the specific functions of each TkSRPP in latex, we screened for GO terms enriched within each subset of paralog-specific interactors (Figure 4). Interestingly, 'chloroplast stroma' (12), 'thylakoid' (5) and 'cytosol' (30) associated proteins were significantly enriched in the CC category among the exclusive TkSRPP3 interactors (compared to all TkSRPP interactors) whereas the TkSRPP4 interactors were significantly enriched for 'polysomal ribosome' (15) and 'cytosolic large ribosomal subunit' (19), 'ER

10.3389/fpls.2024.1498737

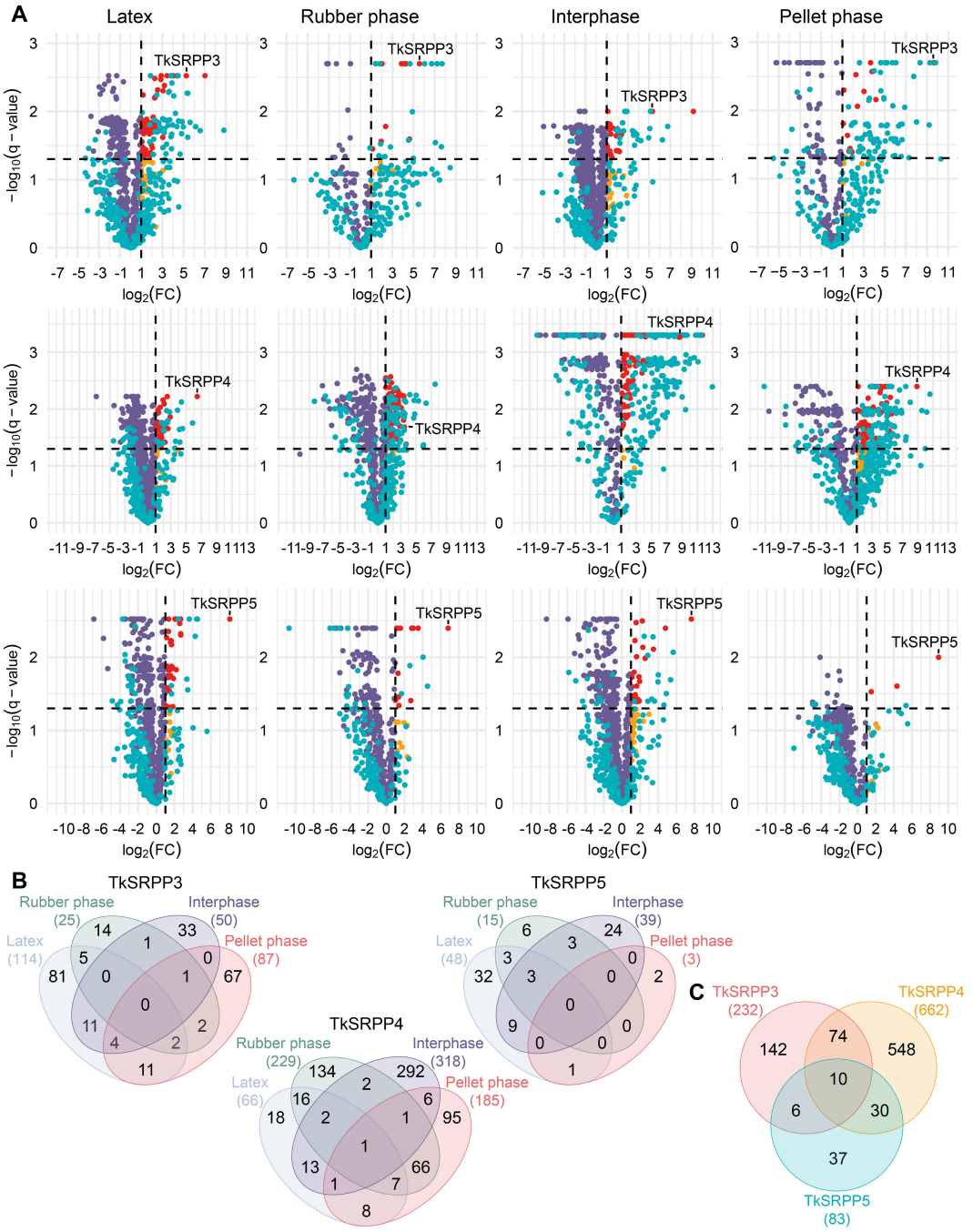


FIGURE 3
AE-MS reveals the enrichment of overlapping sets of proteins with each TkSRPP from each latex fraction. **(A)** Volcano plots showing the enrichment of proteins from four different latex fractions – latex, rubber phase (RP), interphase (IP), pellet phase (PP) – by TkSRPP3/4/5 compared to no-bait controls. The log_2FC values are plotted against the $-log_{10}$(q-value). Dashed lines show threshold values for proteins considered as interactors ($log_2FC \geq 1$; $-log_{10}$(q-value) > 1.3. Red dots highlight proteins in this area. Orange dots represent enriched proteins with higher q-values. Violet dots represent proteins that are not enriched due to their log_2FC. Blue dots mark proteins that were not detected in all replicates of the AE-MS and the control so that missing LFQ values were generated by imputation. These include both, proteins considered as interactors and those outside the thresholds. The latter may be of interest if they are close to the threshold values because imputation can affect enrichment factors and significance levels by sample variance. **(B)** Venn diagrams showing the number of total proteins enriched from each latex fraction with each TkSRPP and the overlaps between fractions. **(C)** The Venn diagram shows the total numbers of different proteins significantly enriched with one TkSRPP from all latex fractions.

membrane' (13) and 'membrane' (181). Accordingly, in the molecular function (MF) category, 'structural constituent of ribosome' (30) was enriched along with 'inorganic molecular entity transmembrane transporter activity' (26). For the TkSRPP5 interactome, the enrichment of 'cytosolic small ribosomal subunit' (3) in the CC category aligns with the enrichment of 'ribosomal small subunit biogenesis' (2) and 'rRNA processing' (2) in the biological process (BP) category. GO analysis thus indicated that

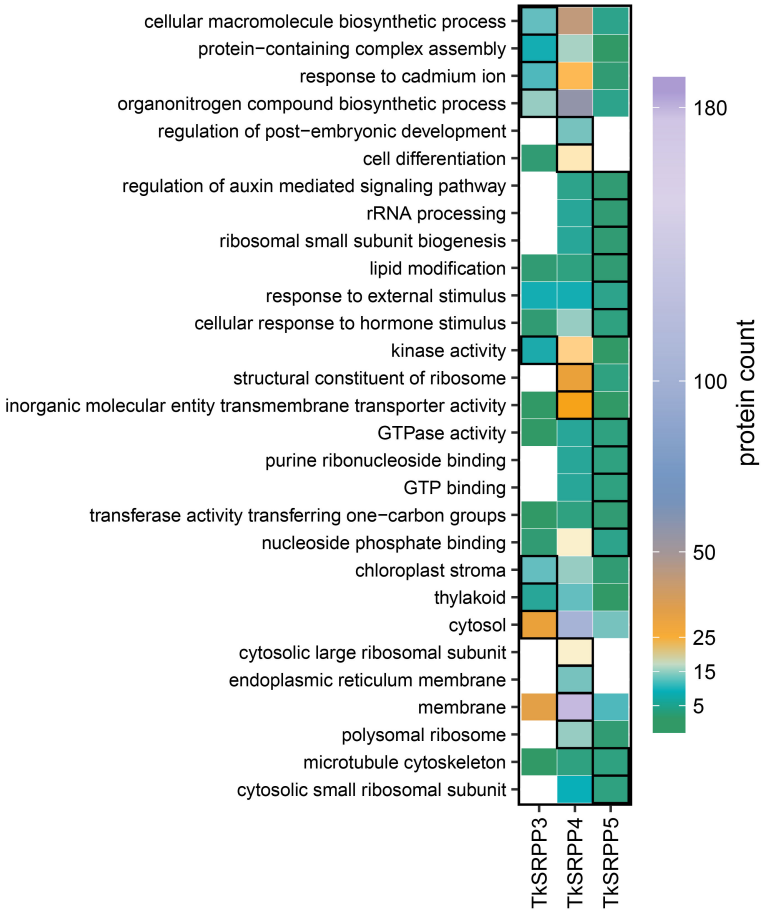


FIGURE 4
Exclusive TkSRPP3/4/5 interactors differ in their assigned GO terms. Heat map showing the number of interactors exclusive to one TkSRPP paralog assigned to a GO term. Framed boxes highlight GO terms significantly enriched (p-value < 0.05) among the exclusive interactors of one TkSRPP paralog compared to all TkSRPP3/4/5 interactors. Significance levels were calculated using a weighted Fisher's exact test.

TkSRPP4 and TkSRPP5 are associated with cytosolic ribosomal processes. It is possible that the enrichment of ribosomal proteins was favored because of the assumed connection between TkSRPPs and the ER, the likely origin of rubber particles. But their detection may also reflect the high affinity of ribosomal proteins for the agarose beads (Keilhauer et al., 2015). For TkSRPP5 interactors, we further observed the enrichment of 'microtubule cytoskeleton' (3) in the CC category, as well as proteins related to responses to hormone and external stimuli (3, 4), 'lipid modification' (2) and 'regulation of auxin mediated signaling pathway' (2) in the BP category. For the TkSRPP3 interactors, the BP category terms 'cellular macromolecule biosynthetic process' (12), 'protein-containing complex assembly' (8), 'response to cadmium ion' (11) and 'organonitrogen compound biosynthetic process' (15) were enriched. The TkSRPP4 interactors were enriched for two more general BP terms relating to 'regulation of post-embryonic development' (13) and 'cell differentiation' (20). In the MF category, 'kinase activity' (7) was the only enriched term among the TkSRPP3 interactors, whereas the TkSRPP5 interactors were enriched for terms related to GTP, purine ribonucleoside and nucleoside phosphate binding (3, 3, 4), 'GTPase activity' (3) and 'transferase activity' (2).

Protein class analysis based on annotated UniProt IDs revealed similar distributions for all three interactomes, partially matching the enriched GO terms (Figure 4; Supplementary Figure S3). Prominent protein classes among the interactors included metabolite interconversion enzymes such as oxidoreductases, transferases and hydrolases, protein-modifying enzymes such as proteases, and protein-binding affinity modulators such as protease inhibitors and G-proteins.

We extracted data relating to the abundance of each interactor from our dataset in wild-type *T. koksaghyz* roots at 8, 12 and 24 weeks (Benninghaus et al., 2020), and clustered the interactors and corresponding bait TkSRPP according to their temporal accumulation profiles (Supplementary Figure S4). This identified interacting proteins with similar abundance profiles as their TkSRPP partners, which supports their status as interaction partners because gene co-expression is more likely for interacting proteins than random protein pairs (Ge et al., 2001; Jansen et al., 2002; von Mering et al., 2002). TkSRPPs 3/4/5 were all assigned to clusters with increasing protein levels over time. All three are already highly abundant after 8 weeks of growth, so that interactions with proteins from the same cluster may represent

TABLE 2 Interaction partners of TkSRPP3/4/5 differentially accumulated in the latex of *TkCPTL1-RNAi* plants compared to wild-type controls.

	ID	NCBI (non-redundant) Protein names (identity)	UniProtKB/Swiss-Prot Protein names (identity)	Log_2FC				
				TkCPTL1-RNAi - WT	Latex	RP	IP	PP
TkSRPP3	evm.model.utg11341.6	cis-prenyltransferase CPT2 [*Taraxacum brevicorniculatum*] (98.05%)	Dehydrodolichyl diphosphate synthase 6 (Dedol-PP synthase 6) (EC 2.5.1.-) (52.80%)	-3.54	2.28		1.25	
TkSRPP4	evm.model.utg10104.22	3-hydroxy-3-methylglutaryl-CoA reductase 2, partial [*Taraxacum kok-saghyz*] (98.20%)	3-hydroxy-3-methylglutaryl coenzyme A reductase 2-A (HMG-CoA reductase 2) (Hydroxymethylglutaryl-CoA reductase) (PgHMGR2) (EC 1.1.1.34) (74.46%)	-1.96			2.68	
	evm.model.utg7969.4	hypothetical protein LSAT_6X38201 [*Lactuca sativa*] (83.07%)	Ricin B-like lectin R40G3 (Osr40g3) (57.52%)	-1.77			1.73	
	evm.model.utg2280.9	probable isoprenylcysteine alpha-carbonyl methylesterase ICMEL2 [*Lactuca sativa*] (77.08%)	Probable isoprenylcysteine alpha-carbonyl methylesterase ICMEL2 (EC 3.1.1.n2) (Isoprenylcysteine methylesterase-like protein 2) (61.10%)	-1.46			1.10	2.01
	evm.model.utg3903.5	CRAL-TRIO domain-containing protein YKL091C-like [*Lactuca sativa*] (81.48%)	Sec14 cytosolic factor (Phosphatidylinositol/ phosphatidyl-choline transfer protein) (PI/PC TP) (Sporulation-specific protein 20) (31.12%)	-1.23			1.80	
	evm.model.utg29345.1	SEC14 cytosolic factor-like [*Lactuca sativa*] (87.05%)	CRAL-TRIO domain-containing protein YKL091C (26.29%)	-1.68			5.12	
	evm.model.utg29792.19	myo-inositol oxygenase 4 *[Artemisia annua*] (89.84%)	Inositol oxygenase 4 (EC 1.13.99.1) (Myo-inositol oxygenase 4) (AtMIOX4) (MI oxygenase 4) (74.13%)	-1.83				3.01
	evm.model.utg17642.9	probable glycerol-3-phosphate acyltransferase 8 isoform X2 *[Lactuca sativa*] (80.28%)	Glycerol-3-phosphate 2-O-acyltransferase 4 (AtGPAT4) (EC 2.3.1.198) (Glycerol-3-phosphate acyltransferase 4) (61.22%)	-5.65				3.86
	evm.model.utg24682.2	germacrene A oxidase [*Lactuca sativa*] (95.49%)	Germacrene A hydroxylase (EC 1.14.14.95) (Germacrene A oxidase) (LsGAO) (95.29%)	1.16			4.05	
	evm.model.utg16440.3	squalene epoxidase 1 [*Taraxacum kok-saghyz*] (98.12%)	Squalene monooxygenase SE1 (EC 1.14.14.17) (Squalene epoxidase 1) (PgSQE1) (SE) (SE1) (gse) (76.24%)	3.11			1.43	
	evm.model.utg8052.7	plastidial pyruvate kinase 2 isoform X2 [*Lactuca sativa*] (93.21%)	Plastidial pyruvate kinase 2 (PKp2) (EC 2.7.1.40) (Plastidial pyruvate kinase 1) (PKP1) (Pyruvate kinase III) (Pyruvate kinase isozyme B1, chloroplastic) (PKP-BETA1) (Plastidic pyruvate kinase beta subunit 1) (95.43%)	1.12			5.63	
TkSRPP5	evm.model.utg1886.1	putative methyltransferase DDB_G0268948 [*Lactuca sativa*] (91.83%)	Putative methyltransferase DDB_G0268948 (EC 2.1.1.-) (32.14%)	-1.30				5.34

The mean log_2FC between transgenic and wild-type plants reported in an earlier study (Niephaus et al., 2019) are shown with the log_2FC of AE-MS experiments and the corresponding latex fractions as determined in the current study.

basal interactions rather than conditional interactions in response to particular stimuli.

To identify protein interactions related to NR synthesis, the interactors were correlated with proteins that are enriched or depleted when *TkCPT-like 1* (*TkCPTL1*) is downregulated in the latex of *T. koksaghyz* plants by RNA interference (RNAi) (data obtained by Niephaus et al., 2019) (Table 2). These *TkCPTL1*-RNAi plants produce significantly less NR than wild-type plants because TkCPTL1 is thought to form heterodimers with TkCPT1 and/or TkCPT2 and thus assemble into a *cis*PT complex that catalyzes the synthesis of poly(*cis*-1,4-isoprene) on the surface of rubber particles (Niephaus et al., 2019). We identified TkCPT1 as a TkSRPP3 interactor and TkCPT1 was significantly less abundant in the *TkCPTL1*-RNAi plants. An interaction between TkSRPP3 and TkCPT1 highlights the connection between TkSRPP3 and NR synthesis on the rubber particle surface, and is consistent with the reported interaction between *Hevea brasiliensis* SRPP and CPT6 (Brown et al., 2017). TkSRPP4 interacted with 10 proteins whose abundance changed in the NR-depleted transgenic plants, including the rate-limiting enzyme of the mevalonate (MVA) pathway: 3–hydroxy-3-methylglutaryl-CoA reductase (HMGR). The MVA pathway provides the C_5 building block isopentenyl diphosphate (IPP) for NR synthesis, and was downregulated in the RNAi lines. Other interactors involved in isoprenoid metabolism, including squalene epoxidase 1 (SQE1) and germacrene oxidase (GAO), were more abundant in the *TkCPTL1*-RNAi lines. Another downregulated interactor was a homolog of a ricin B-like lectin, and additional lectin homologs were found in the TkSRPP3 and TkSRPP4 interactomes. One TkSRPP5 interactor similar to a lettuce (*Lactuca sativa*) putative methyltransferase was also downregulated in the RNAi plants.

3.4 Confirmation of selected TkSRPP3/4/5 interactors identified by AE-MS

We selected two of the 10 candidate interactors shared by TkSRPP3/4/5 for confirmation using a second method and characterized them in more detail. The first candidate (annotated as TkSRPP6) was selected because *TkSRPP6* is the only complete *TkSRPP* gene found outside the main cluster, and its function has not been studied thus far. The second candidate, annotated as a sterol 3–β-glucosyltransferase/UDP-glycosyltransferase (UGT) 80B1 family member, was designated TkUGT80B1. It was selected because glycosides are known to be involved in plant defense and stress responses, but the role of UGTs in latex has not been investigated. The proteins were enriched to different levels in different phases in the AE-MS datasets for TkSRPP3, 4 and 5 (Table 3).

For SUY2H, TkSRPP3/4/5 baits were N-terminally fused to a modified N-terminal ubiquitin fragment (N_{UbA}) with lower affinity for the ubiquitin C-terminus (C_{Ub}), thus minimizing false positive results (Johnsson and Varshavsky, 1994). The N_{UbA} fusions were co-expressed with TkSRPP6 C-terminally fused to C_{Ub}. Interactions reconstitute functional ubiquitin, leading to the cleavage and degradation of the URA3 reporter, thus conferring uracil auxotrophy and resistance to 5-FOA (Johnsson and Varshavsky, 1994; Reichel and Johnsson, 2005). We used the monomeric fluorescent protein mEmerald combined with the TkSRPPs as negative controls. Using this system, we were able to confirm that TkSRPP6 interacts with TkSRPP4 and TkSRPP5. For TkSRPP3 the growth pattern was indistinct but suggested weak interaction (Figure 5A). This supports the initial screens, but indicates that protein interactions are dependent on the experimental conditions and highlights the importance of independent confirmation. For interactions between TkSRPP3/4/5 and TkUGT80B1, we were able to pull down TkUGT80B1-3×HA with Cerulean-tagged TkSRPP3 and TkSRPP5 by Co-IP, but we could not confirm the interaction with TkSRPP4 in this experimental setup (Figures 6A; Supplementary Figure S5). In summary, additional methods confirmed four of six pairwise interactions indicated by AE-MS, two for each candidate, suggesting the unconfirmed interactions are restricted to specific native conditions or part of bigger complexes.

3.5 Sequence analysis of TkSRPP6 and TkUGT80B1

TkSRPP6 *in silico* analysis identified the REF domain common to all known dandelion SRPPs and REF proteins, as well as three potential phosphorylation sites (Figure 5B). Phylogenetic comparisons showed that TkSRPP6 has diverged from other TkSRPPs and is more closely related to other REF proteins (including those involved in stress responses in plants that do not produce NR) than to the tightly clustered TkSRPP1/2/3/4/5 and 7 (Figure 5C). This was supported by protein identities of ~60% between TkSRPPs 3/4/5, but only 37–49% when TkSRPP6 was compared to the other paralogs (Table 1). We therefore screened a 1-kb region of the *TkSRPP6* promoter for stress-responsive elements (Figure 5D; Supplementary Table S5), revealing E–box

TABLE 3 Enrichment of TkSRPP6 and TkUGT80B1 by TkSRPP3/4/5 based on AE-MS data.

Genome ID	Given name	Log_2FC											
		TkSRPP3				TkSRPP4				TkSRPP5			
		L	RP	IP	PP	L	RP	IP	PP	L	RP	IP	PP
evm.model.utg2059.21	TkSRPP6		4.7	5.0		1.1						2.2	
evm.model.utg4564.7	TkUGT80B1				2.5			6.9				3.5	

Log_2FC values are provided for each bait TkSRPP and latex fraction in which TkSRPP6 and TkUGT80B1 were significantly enriched. L, latex; RP, rubber phase; IP, interphase; PP, pellet phase.

10.3389/fpls.2024.1498737

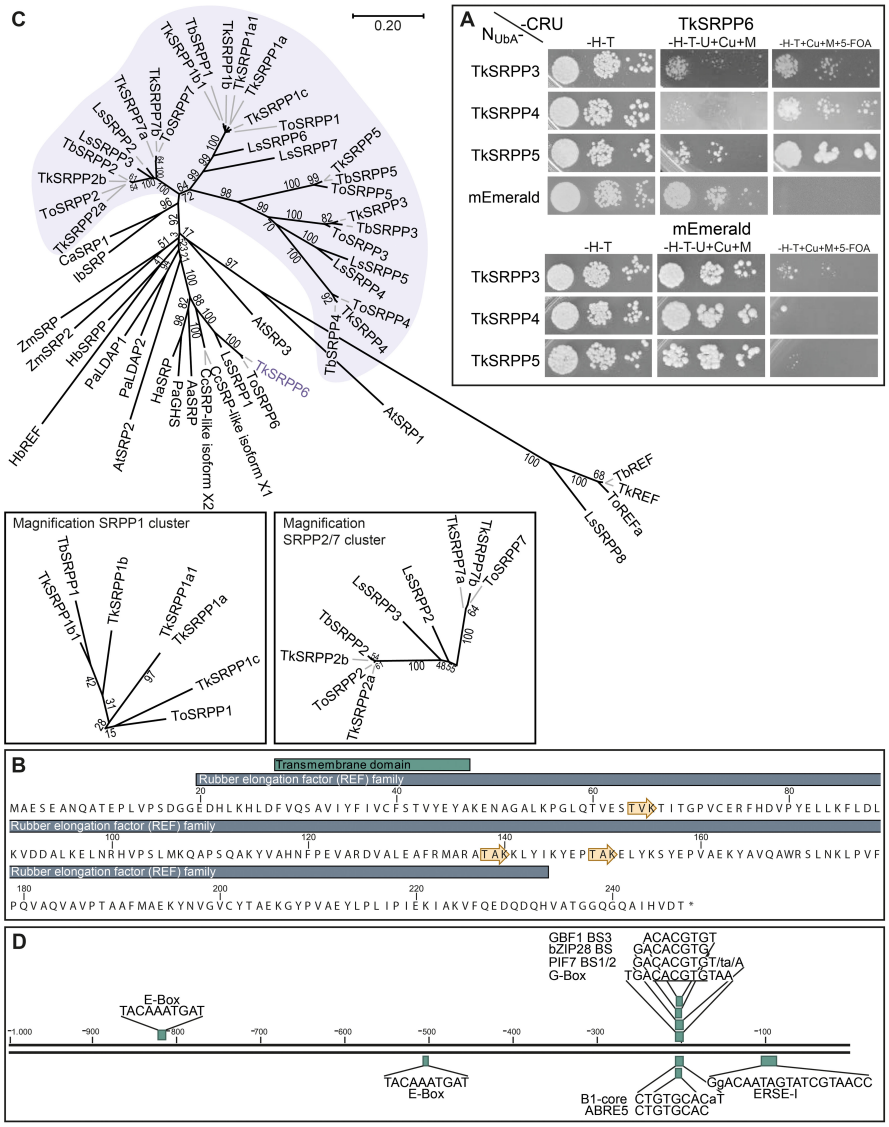


FIGURE 5
General characterization of TkSRPP6. **(A)** Split-ubiquitin yeast-two hybrid (SUY2H) indicating protein interactions between TkSRPP6 and TkSRPP4/TkSRPP5. Yeast expressing TkSRPP6 C-terminally fused to the C-terminal part of ubiquitin and URA3 as a reporter (CRU), and TkSRPP3/4/5 N-terminally fused to the N-terminal part of ubiquitin (N_{UbA}), were dropped in three different dilutions on selective media and grown for 2–3 days. Medium lacking histidine and tryptophan (–H–T) was used as a control medium to select for the plasmid encoding the proteins of interest. Medium additionally lacking uracil but containing 50 µM $CuSO_4$ and 300 µM methionine (–H–T–U+Cu+M) was used to select for URA3 activity. Medium containing uracil and 1 g/L 5-FOA (–H–T +Cu+M+5-FOA) was used to select for URA3 inactivity reflecting bait/prey interactions. The monomeric mEmerald fluorophore was used as a negative control. **(B)** TkSRPP6 protein sequence containing a REF domain and a short N-terminal transmembrane domain predicted by InterPro. Yellow arrows represent predicted phosphorylation sites. **(C)** Phylogenetic analysis reveals clustering of TkSRPP6 with stress-related REF family proteins from non-rubber plants, separated from the other TkSRPPs. Multiple sequences were aligned using CLUSTALW and the phylogenetic tree was constructed using the neighbor-joining algorithm and a bootstrap of 500. Values at branches indicate bootstrap values. The phylogenetic distance is indicated by the scale bar. Accession numbers: AaSRP, *Artemisia annua* stress-related protein (PWA88416.1); AtSRP1, *Arabidopsis thaliana* REF/SRPP-like protein At1g67360 (NP_176904.1); AtSRP2, *A. thaliana* REF/SRPP-like protein At2g47780 (NP_182299.1); AtSRP3, *A. thaliana* REF/SRPP-like protein At3g05500 (NP_187201.1); CaSRP1, ADI60300.1; CcSRPP-like isoform X1, *Cynara cardunculus* var. *scolymus* stress-related protein-like isoform X1 (XP_024981582.1); CcSRPP-like isoform X2, *C. cardunculus* var. *scolymus* stress-related protein-like isoform X2 (XP_024981583.1); HaSRP, *Helianthus annuus* putative stress-related protein (A0A251TGA8); HbREF, *Hevea brasiliensis* REF (P15252); HbSRPP, *H. brasiliensis* SRPP (O82803); IbSRP, *Ipomoea batatas* stress-related protein (ABP35522.1); LsSRPP1, *Lactuca sativa* SRPP1 (XP_023771881.1); LsSRPP2, (AJC97799.1); LsSRPP3, (AJC97800.1); LsSRPP4, (AJC97801.1); LsSPP5, (AJC97802.1); LsSRPP6, (AJC97803.1); LsSPP7, (AJC97804.1); LsSRPP8, (AJC97805.1); PaGHS, *Parthenium argentatum* rubber synthesis protein (AAQ11374.1); PaLDAP1, *Persea americana* lipid droplet-associated protein 1 (AGQ04593.1); PaLDAP2, *P. americana* lipid droplet-associated protein 2 (AGQ04594.1); TbREF, *Taraxacum brevicorniculatum* REF (A0A291LM03); TbSRPP1, *T. brevicorniculatum* SRPP1 (M9PNN1); TbSRPP2, (AGE89407.1); TbSRPP3, (M9PNQ7); TbSRPP4, (M9PNN3); TbSRPP5, (M9PNM8); TkREF, *Taraxacum koksaghyz* REF (GWHPBCHF036022); TkSRPP1a, GWHPAAAA010568; TkSRPP1a1, GWHPBCHF033216; TkSRPP1b, GWHPAAAA043688; TkSRPP1b1, GWHPBCHF033215; TkSRPP1c, GWHPAAAA010568; TkSRPP2: deduced from identified gene locus (Supplementary Table S2); TkSRPP2a, GWHPAAAA010566; TkSRPP3, GWHPAAAA015362; TkSRPP4, GWHPAAAA015361; TkSRPP5, GWHPAAAA015359; TkSRPP6, GWHPAAAA016929; TkSRPP7, GWHPBCHF033106; TkSRPP7a, GWHPBCHF033213; ZmSRP, *Zea mays* stress-related protein (ACG39345.1); ZmSRP2, *Z. mays* REF/SRPP-like protein (NP_001149834.1). TkSRPP sequence IDs originate from published genome data (Lin et al., 2018, 2022). **(D)** 1 kb promotor region of *TkSRPP6* containing different *cis*-acting regulatory elements connected to plant stress responses. Promotor region was extracted from the published *T. koksaghyz* genome (Lin et al., 2022) and regulatory elements were determined using NSITE-PL.

10.3389/fpls.2024.1498737

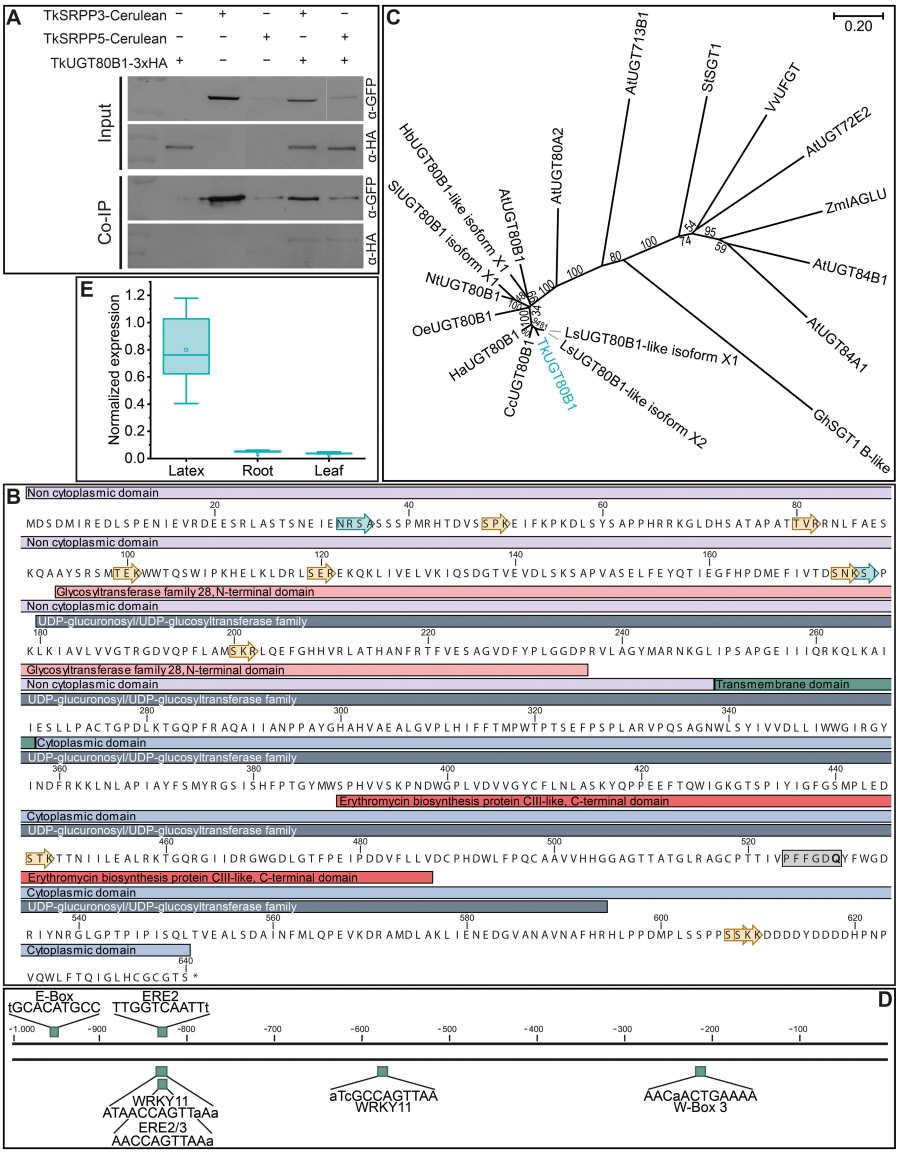


FIGURE 6
General characterization of TkUGT80B1. **(A)** Co-Immunoprecipitation (Co-IP) assay showing the interaction of TkUGT80B1-3xHA with TkSRPP3/5-Cerulean. The top two panels show the input samples and the bottom two panels protein detection after immunoprecipitation with an α-GFP antibody. Fusion proteins were extracted from yeast cells. **(B)** TkUGT80B1 protein sequence with assigned domains. The gray box highlights the UDPGT motif with glutamine in the last position, characteristic of UDP-glucosyltransferases. Yellow arrows represent phosphorylation and cyan arrows *N*-glycosylation sites predicted using CLC Main Workbench. **(C)** Phylogenetic analysis reveals clustering of TkUGT80B1 with UGT80B1 proteins from other Asteraceae. Multiple sequences were aligned using CLUSTALW and the phylogenetic tree was constructed using the neighbor-joining algorithm and a bootstrap of 500. Values at branches indicate bootstrap values. The phylogenetic distance is indicated by the scale bar. Accession numbers: AtUGT72E2, *Arabidopsis thaliana* UDP-glycosyltransferase superfamily protein UGT72E2 (NP_201470.1); AtUGT80A2, *A. thaliana* sterol 3-β -glucosyltransferase UGT80A2 (NP_566297); AtUGT80B1, *A. thaliana* sterol 3-β -glucosyltransferase UGT80B1 (NP_175027); AtUGT84A1, *A. thaliana* UDP-glycosyltransferase 84A1 (NP_193283.2); AtUGT713B1, *A. thaliana* glycosyltransferase UGT713B1 (NP_568452); CcUGT80B1, *Cynara cardunculus* var. *scolymus* sterol 3-β-glucosyltransferase UGT80B1 (XP_024976598.1); GhSGT1 B-like, *Gossypium hirsutum* sterol glucosyltransferase 1 homolog B-like (JN004107); HaUGT80B1, *Helianthus annuus* sterol 3-β-glucosyltransferase UGT80B1 (XP_035834958.1); HbUGT80B1-like isoform X1, *H. brasiliensis* sterol 3-β-glucosyltransferase UGT80B1-like isoform X1 (XP_021673215.1); LsUGT80B1-like isoform X1, *Lactuca sativa* sterol 3-β-glucosyltransferase UGT80B1-like isoform X1 (XP_023742443.1); LsUGT80B1-like isoform X2, *L. sativa* sterol 3-β-glucosyltransferase UGT80B1-like isoform X2 (XP_023742444.1); NtUGT80B1, *Nicotiana tomentosiformis* sterol 3-β-glucosyltransferase UGT80B1 (XP_009595972.1); OeUGT80B1, *Olea europaea* subsp. *europaea* sterol 3-β-glucosyltransferase UGT80B1 (CAA2989377.1); SlUGT80B1 isoform X1, *Solanum lycopersicum* sterol 3-β-glucosyltransferase UGT80B1 isoform X1 (XP_004237799.1); StSGT1, *Solanum tuberosum* UDP-galactose: solanidine galactosyltransferase (AB48444.2); TkUGT80B1, *Taraxacum koksaghyz* UDP-glycosyltransferase 80B1 (GWHPAAAA034502); VvUFGT, *Vitis vinifera*, UDP glucose:flavonoid 3-*O*-glucosyltransferase (AAB81683.1); ZmIAGLU, *Zea mays* indole-3-acetate β-glucosyltransferase (Q41819). *T. officinale* sequences were obtained from unpublished data. **(D)** *TkUGT80B1* 1-kb promoter region containing different *cis*-acting regulatory elements associated with plant stress responses. Promoter region was extracted from the published *T. koksaghyz* genome (Lin et al., 2022) and regulatory elements were determined using NSITE-PL. **(E)** *TkUGT80B1* is predominantly expressed in latex. Normalized gene expression levels in different tissues of 10-week-old wild-type *T. koksaghyz* plants. Box plots represent data from five individual plants. Expression levels were normalized against *elongation factor-1 α* (*TkEF1α*) and *ribosomal protein L27* (*TkRP*).

elements (CANNTG) at positions –200 and –500 bp relative to the start codon, and the core sequence of a G-box type E-box (CACGTG) and extended G-box elements at –800 bp (Figure 5D) (Galvão et al., 2019; Nagao et al., 1993; Shahmuradov and Solovyev, 2015). The G-box recruits G-box binding factors (GBFs), which include bZIP and bHLH proteins such as MYC2 (Heim et al., 2003; Menkens et al., 1995; Sibéril et al., 2001; Williams et al., 1992; Zhang et al., 2019). G-box elements mediate the effects of hormones, light and temperature (Eyal et al., 1993; Guiltinan et al., 1990; Hong et al., 1995; Mason et al., 1993; Shaikhali et al., 2012; Toledo-Ortiz et al., 2014), whereas E-box elements regulate temperature-dependent and circadian expression in stress-responsive genes (Liu et al., 2015; Seitz et al., 2010). These findings suggest that *TkSRPP6* is transcriptionally regulated by different stress factors, in agreement with other data for *SRPP* genes (Cao et al., 2017; Dong et al., 2023a; Fricke et al., 2013; He et al., 2024; Hillebrand et al., 2012).

TkUGT80B1 was found to contain a UDP-glucuronosyltransferase/UDP-glucosyltransferase domain, an N-terminal domain similar to glycosyltransferase family 28, and a C-terminal domain resembling that of CIII-like, another glycosyltransferase, including the nucleotide diphosphate sugar binding site (Figure 6B) (Moncrieffe et al., 2012). The last amino acid in the so–called UDPGT motif differs between UDP-glucosyltransferases (where it is glutamine) and UDP-galactosyltransferases (histidine), so the presence of glutamine in TkUGT80B1 suggests it has UDP-glucose transferase activity (Figure 6B, gray box) (Kubo et al., 2004). Sequence analysis also predicted that the C-terminal domain is cytosolic, separated from the N-terminal part by a transmembrane domain of 19 amino acids. The protein contains nine putative phosphorylation sites and two *N*-glycosylation sites. Phylogenetic analysis supported the relationship between TkUGT80B1 and UGT80B1 enzymes from the family Asteraceae and other plants, as well as more distant relationships with other UGT families (Figure 6C). The *TkUGT80B1* promoter (Figure 6D) contains three elicitor response elements (ERE1-3), which contribute to fungal elicitor-mediated gene expression (Yang et al., 1998). Additionally, we found an E-box, two WRKY11-binding sites (one overlapping with ERE2/3) and one WRKY40-binding site (W-box). WRKY transcription factors are involved in plant defense (Javed and Gao, 2023), suggesting stress-responsive transcriptional regulation, which ties in with the role of glycosylated secondary metabolites in the plant defense system (Hussain et al., 2019; Louveau and Osbourn, 2019).

3.6 Gene expression profiles of TkSRPP3/4/5 and their interaction partners TkSRPP6 and TkUGT80B1

We had already determined the spatial expression profile of *TkSRPP6* when comparing SRPP paralogs (Figure 1). Applying the same approach to *TkUGT80B1* in 10-week-old wild-type *T. koksaghyz* plants, we observed strong expression in the latex (consistent with the AE-MS experiments) but low expression in roots and leaves (Figure 6E), similar to the expression profiles of *TkSRPP3/4/5*. The previous detection of TkUGT80B1 protein in roots may reflect the large amount of latex in this tissue (Benninghaus et al., 2020).

Temporal expression profiling in latex revealed a steady increase in *TkSRPP3/4/5* mRNA levels during weeks 6–14 (Figure 7), as shown for root protein levels before (Benninghaus et al., 2020). However, transcript levels stayed constant or decreased slightly between weeks 14 and 16 (Figure 7A). *TkSRPP3* and *TkSRPP4* expression declined after 12 weeks but increased again after 14 weeks. Similarly, *TkUGT80B1* expression increased over time, declined slightly after 12 weeks, and stayed constant between weeks 14 and 16 (Figure 7B). *TkSRPP6* expression was constant at low levels throughout the experiment (Figure 7B). The expression data reflected the high level of heterogeneity between individuals reported earlier (McAssey et al., 2016; Nowicki et al., 2019; Panara et al., 2018; Wieghaus et al., 2022). Our data demonstrated comparable temporal expression patterns for *TkUGT80B1* and *TkSRPP3/4/5*, but not *TkSRPP6*.

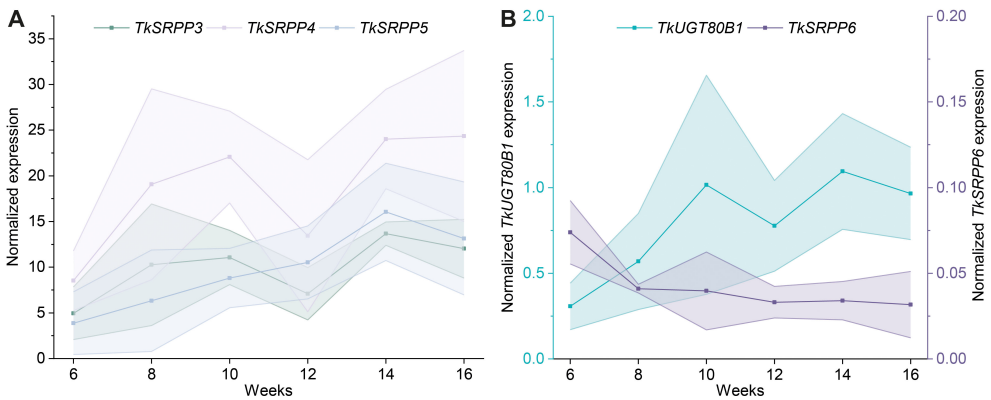


FIGURE 7
TkSRPP4-5 and *TkUGT80B1* show similar temporal expression patterns in latex. Normalized gene expression levels of **(A)** *TkSRPP3/4/5* and **(B)** *TkUGT80B1* and *TkSRPP6* in *T. koksaghyz* wild-type latex over time. Data points are means of 4–7 individual plants. The shaded areas represent the areas within in the standard deviations. Expression levels were normalized against *elongation factor-1α* (*TkEF1α*) and *ribosomal protein L27* (*TkRP*).

10.3389/fpls.2024.1498737

3.7 Cellular localization of TkSRPP6 and TkUGT80B1

The analysis of different latex phases by AE-MS provided crude data concerning the potential localization of TkSRPP6 and TkUGT80B1. For more detailed analysis, we expressed fusion proteins in *N. benthamiana* along with subcellular markers. We prepared constructs in which TkSRPP6 and TkUGT80B1 were C-terminally fused to the fluorescent reporter Cerulean, and transiently co-expressed them with ER and tonoplast markers. TkSRPP6-Cerulean fluorescence and the ER marker CYP51G1-mRFP (Bassard et al., 2012) overlapped almost completely (Figure 8A), whereas TkUGT80B1-Cerulean fluorescence largely coincided with the tonoplast marker TPC1-OFP (Batistič et al., 2010) (Figure 8C). Tk/TbSRPPs 3/4/5 were previously shown to be associated with rubber particles (Collins-Silva et al., 2012; Hillebrand et al., 2012), which are related to LDs, thus explaining the LD localization of TbSRPPs in *N. benthamiana* (Laibach et al., 2018). We therefore determined whether TkSRPP6 and TkUGT80B1 also associate with LDs by co-expressing the Cerulean fusion constructs with *AtLEC2*, encoding a transcription factor that promotes LD formation in leaves (Santos Mendoza et al., 2005). We then stained the LDs with the lipophilic fluorescent dye Nile red. We found that the Cerulean fluorescence profiles of TkSRPP6 and TkUGT80B1 described above included additional punctuate fluorescence that overlapped with the Nile red signal (Figures 8B, D). The affinity of these candidates for LDs, despite the absence of enrichment in the RP fraction in AE-MS experiments, suggests they interact with TkSRPP3/4/5 on the surface of rubber particles but in a conditional manner.

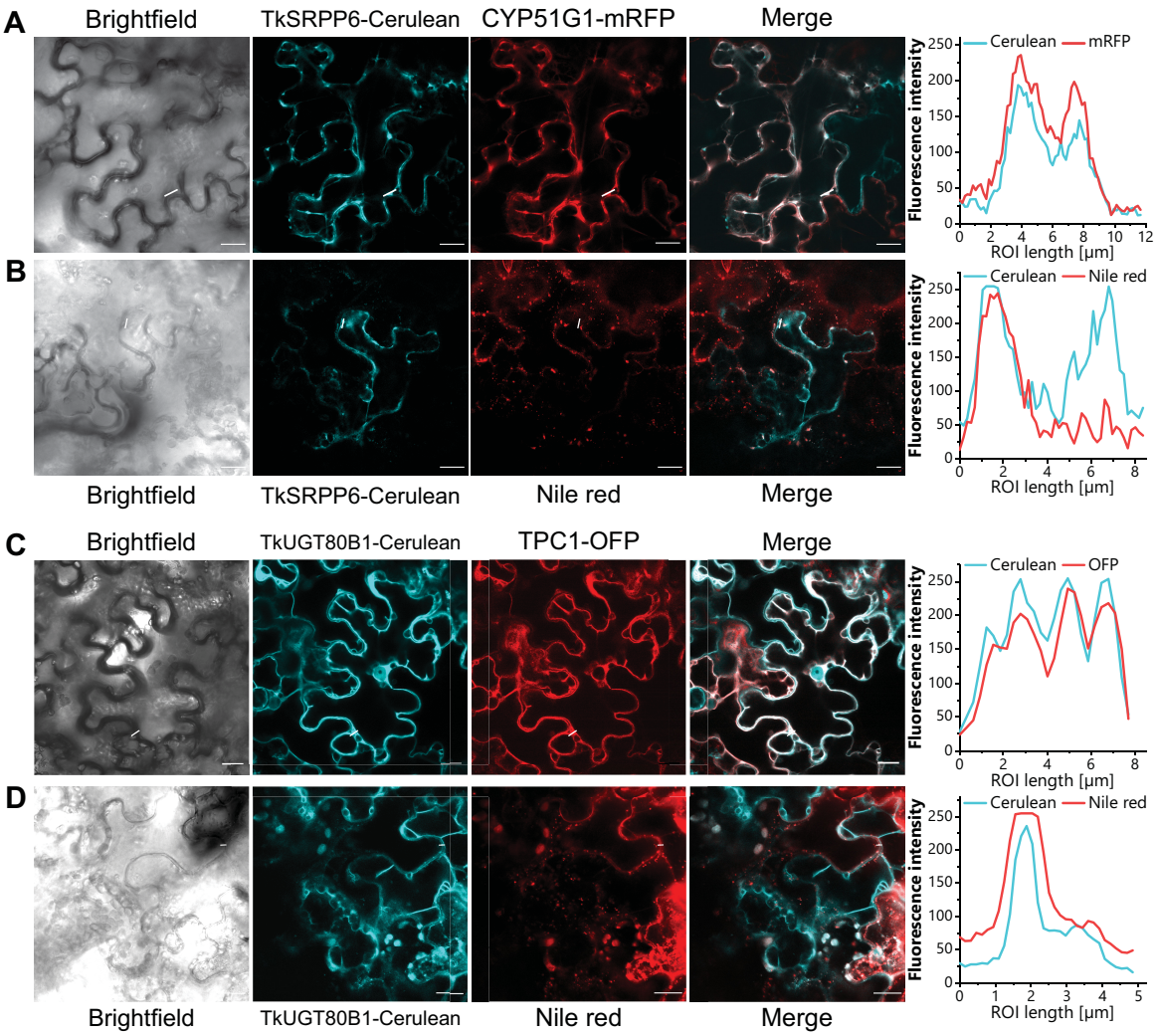


FIGURE 8
TkSRPP6 and TkUGT80B1 localize to the ER and tonoplast, respectively, and show affinity to LDs. *N. benthamiana* leaf epidermal cells expressing N-terminal Cerulean fusion constructs (cyan) and mRFP or OFP fusion subcellular markers are shown. **(A)** TkSRPP6-Cerulean and ER marker CYP51G1-mRFP. **(B)** TkSRPP6-Cerulean and Nile red signal representing LDs. **(C)** TkUGT80B1-Cerulean and tonoplast marker TPC1-OFP. **(D)** TkUGT80B1-Cerulean and Nile red signal representing LDs. For LD formation, Cerulean fusion constructs were co-expressed with *AtLEC2* and LDs were stained with the lipophilic fluorescent dye Nile red. Fluorescence intensities in regions of interest are depicted on the right. Scale bar = 20 µm.

3.8 Glycosyltransferase activity of TkUGT80B1

Finally, we tested the predicted UGT activity of TkUGT80B1 in a yeast strain engineered for optimized pentacyclic triterpenoid synthesis and harboring a *T. koksaghyz* lupeol synthase gene (*TkLup*) (Bröker et al., 2018). Lupeol is a pentacyclic triterpenoid present in *T. koksaghyz* roots and NR, and is therefore a potential native substrate for TkUGT80B1 (Benninghaus et al., 2020; Pütter et al., 2019). Isoprenoid metabolites were extracted from yeast cultures and LC-MS chromatograms were compared to control strains either expressing *TkLUP* together with an empty vector or *TkUGT80B1* without *TkLup* (Figure 9; Supplementary Figure S6). We observed an additional peak (*m/z* +606.5) for yeast cells expressing *TkUGT80B1* and *TkLup* (Figure 9B). The mass corresponds to a positively charged lupeol hexose ammonium ion adduct, and thus indicates TkUGT80B1 has lupeol glycosylating activity. Based on the molecular structure of lupeol, we deduce that TkUGT80B1 is a C_3-glycosyltransferase.

4 Discussion

The *T. koksaghyz* genome encodes 13 homology-based, full-length TkSRPPs (Lin et al., 2018, 2022), and their diverse sequences and expression profiles suggest non-redundant specialized functions in different tissues (Figure 1; Table 1). *TkSRPPs 3/4/5* are strongly expressed in the latex, so we sought interacting proteins that may contribute to NR biosynthesis and stress responses. The high constitutive levels of TkSRPP3/4/5 in latex indicate their requirement for basic processes without external stimuli, including rubber particle biogenesis, coating and stabilization. However, the presumably higher levels of TkSRPP3/4/5 protein following stress-induced transcriptional upregulation (Dong et al., 2023a; He et al., 2024; Laibach et al., 2018) indicate that the constitutive pool is insufficient to fulfil the extended functions needed in response to environmental changes, necessitating *de novo* protein synthesis. The presence of *N*-glycosylation and phosphorylation sites in TkSRPP3/4/5 indicates the proteins can be covalently modified, which may result in conformational and functional changes (Ha and Loh, 2012; Volkman et al., 2001). The different numbers of potential post-translational modification sites and distinct protein charges resulting from *TkSRPP* sequence divergence likely contribute to TkSRPP3/4/5 functional divergence represented by their separate interactomes.

4.1 TkSRPP3/4/5 interact with proteins related to isoprenoid and NR biosynthesis

Our AE-MS experiments revealed distinct but overlapping interactomes for TkSRPP3/4/5 in whole latex and its three fractions. TkSRPP4 interacted with more proteins than the others and may function as a hub. TkSRPP3 and TkSRPP5 also interacted with each other, although enrichment was only observed from whole latex and IP, not from the RP fraction (Supplementary Table S7). SRPP heterodimers have also been reported for *T. brevicorniculatum* (Laibach et al., 2018). These findings suggest TkSRPPs can act cooperatively, in agreement with the additive effect of TbSRPPs 3/4/5 on artificial poly(*cis*-1,4-isoprene) body size and dispersity (Laibach et al., 2018). Lipid–protein interactions influence membrane composition (Harayama and Riezman, 2018) so the TkSRPP3/TkSRPP5 interaction may induce specific rearrangements in the lipid monolayer of rubber particles that promote the most stable lipid distribution, and/or enhance the steric repulsion assumed to be caused by SRPPs on the rubber particle surface (Figure 10A) (Hillebrand et al., 2012). TkSRPP3 and TkSRPP5 may also form complexes with their common interactors, including a REF family protein distantly related to a perilipin-4-like protein from the tobacco hawkmoth *Manduca sexta*, which was significantly less abundant in NR–depleted *T. koksaghyz* roots (Benninghaus et al., 2020). Perilipins are LD-associated proteins in animals that promote the formation and stability of LDs by

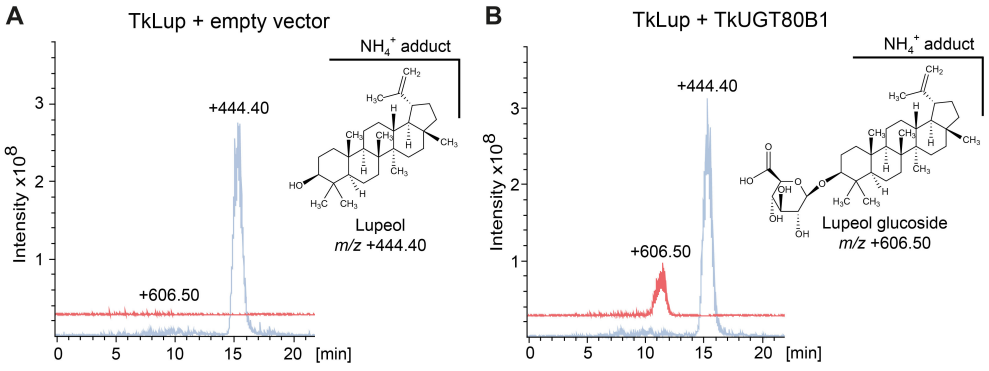


FIGURE 9
TkUGT80B1 glycosylates the triterpenoid lupeol in yeast. LC-MS chromatograms of extracts from yeast metabolically engineered for increased triterpenoid production (Bröker et al., 2018) expressing additionally **(A)** *lupeol synthase* (*TkLup*) and *TkUGT80B1* and **(B)** only *TkUGT80B1*. The signal at *m/z* 444.40 corresponds to the lupeol ammonium ion and *m/z* 606.50 to the ammonium ion of glucosylated lupeol that is only detectable when *TkLup* and *TkUGT80B1* are co-expressed. The chemical structure of lupeol and the predicted structure of the C_3 glucosylated lupeol corresponding to *m/z* 606.50 are shown next to the chromatograms. Chromatograms of an additional control strain containing *TkLup* and an empty vector control are shown in Supplementary Figure S5.

10.3389/fpls.2024.1498737

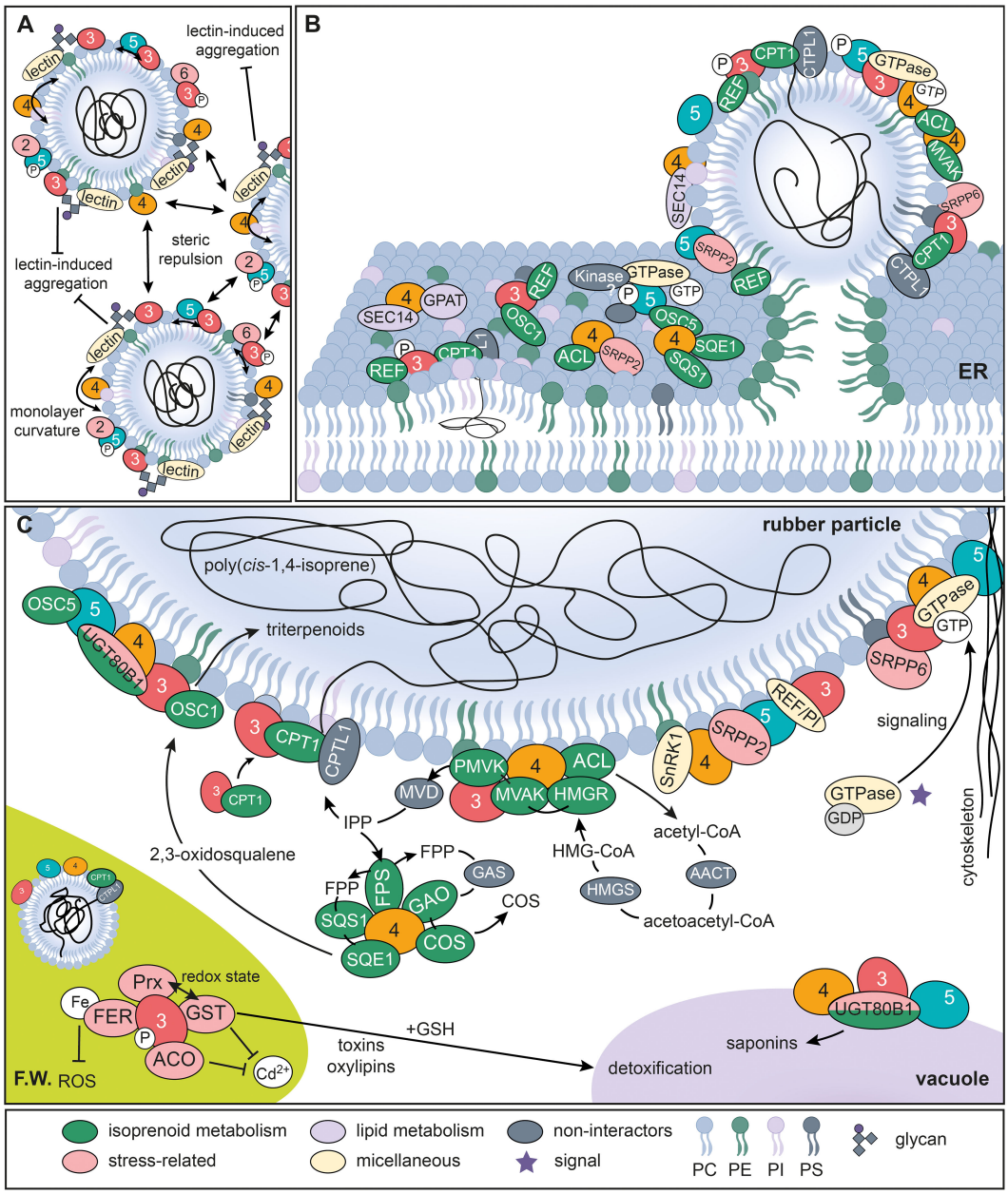


FIGURE 10
Models of TkSRPP3/4/5 potential protein interaction networks. **(A)** Model of processes involved in rubber particle stabilization and dispersity. **(B)** Illustration of potential TkSRPP3/4/5 protein interactions at the ER membrane initiating rubber particle formation. **(C)** Proposed protein interactions of TkSRPP3/4/5 at the surface of rubber particles and other cellular components. Black lines and arrows indicate molecule movement. T-shaped arrows indicate inhibitory effects. Two-sided arrows depict repulsion. TkSRPP3/4/5 names are shortened to their respective numbers. AACT, acetoacetyl-CoA thiolase; ACL, ATP-citrate synthase; COS, costunolid/costunolid synthase; CPT, *cis*-prenyltransferase; CPTL, *cis*-prenyltransferase-like; FER, ferritin; FPP, farnesyl diphosphate; FPS, farnesyl diphosphate synthase; FW, Frey-Wyssling complex; GAO, germacrene A oxidase; GDP, guanosine diphosphate; GPAT, glycerol-3-phosphate acyltransferase; GSH, glutathione sulfhydryl form; GST, glutathione S-transferase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl-coenzyme A synthase; MVD, mevalonate diphosphate decarboxylase; MVK, mevalonate kinase; OSC, oxidosqualene cyclase; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; Prx, peroxiredoxin; PS, phosphatidylserine; REF/PI, REF domain containing/perilipin-like related protein; PMVK, phosphomevalonate kinase; ROS, reactive oxygen species; SnRK1, Snf1-related protein kinase 1; SQS1, squalene synthase 1; SQE1, squalene epoxidase 1; UGT80B1, UDP-glycosyltransferase 80B1.

regulating lipolysis (Griseti et al., 2024). Although this protein was not enriched from the RP, further analysis to determine its impact on rubber particles would be interesting, especially given its lack of interaction with TkSRPP4.

To understand the role of TkSRPP3/4/5 in NR biosynthesis beyond rubber particle stability, we screened the interactomes for further proteins related to NR and other isoprenoids (Supplementary Table S7). Many candidate TkSRPP interactors

were found to differ in abundance when comparing proteins in the latex of *TkCPTL1*-RNAi plants and wild-type controls, suggesting a contribution to NR biosynthesis (Niephaus et al., 2019) (Table 2). Interestingly, TkSRPP3 interacted with TkCPT1 and TkREF, two well-known components of the NR biosynthesis machinery. The interaction with TkCPT1 supports the hypothesis that TkSRPPs affect TkCPT1 activity, causing the low NR content in *Tk/TbSRPP* RNAi plants (Collins-Silva et al., 2012; Hillebrand et al., 2012). Based on the proposed role of TbREF in rubber particle biogenesis (Laibach et al., 2015) and the recruitment of HbCPT6 from the cytosol to the ER by HbSRPP (Brown et al., 2017), TkSRPP3 may recruit TkCPT1 to ER sites where TkSRPP3 is in contact with other TkSRPPs and TkREF, thus modifying the lipid composition, helping TkCPT1 to channel nascent poly(*cis*-1,4-isoprene) chains between ER leaflets and favoring the formation of rubber particles (Figure 10B). This is supported by the reported interaction between HbSRPP and HbREF (Yamashita et al., 2016). The TkSRPP3/TkCPT1 interaction was detected from whole latex and the IP. Although NR-producing CPTs have mostly been identified in rubber particles (Dai et al., 2013; Schmidt et al., 2010b), the recruitment of TkCPT1 from the cytosol to the ER by TkSRPP3 may be the mechanism by which TkCPT1 becomes localized to this compartment.

TkSRPP3 and TkSRPP4 interacted with MVA pathway enzymes that provide the C_5 building block IPP for NR polymerization (Pu et al., 2021; Salehi et al., 2021). This group comprised ATP-citrate synthase (TkACL1), TkHMGR, and mevalonate kinase (TkMVAK1) (all interacting with TkSRPP4), as well as TkMVAK10 (interacting with TkSRPP3) and phosphomevalonate kinase 3 (TkPMVK3, interacting with TkSRPPs 3 and 4). TkSRPP4 also interacted with an FPP synthase (TkFPS1), which provides the most likely starter molecule for NR polymerization *in vivo* (Puskas et al., 2006; Tanaka et al., 1996; Xie et al., 2008). TkSRPP3 and TkSRPP4 may therefore be required to ensure an efficient supply of metabolic precursors to the *cis*PT complex by forming the structural components of a NR metabolon (Figure 10C), although metabolic channeling experiments would be required for confirmation (Srere, 1972, 1987; Zhang and Fernie, 2021). This model is supported by the enrichment of TkMVAK1, a cytosolic enzyme (Cho et al., 2022; Niu et al., 2021; Simkin et al., 2011), from the RP by TkSRPP4 (Supplementary Data S2 in Supplementary Data Sheet 1). TkSRPP3 may also stabilize the complex by interacting with TkSRPP5, which in turn interacts with multiple proteins related to the 'microtubule cytoskeleton' (Figure 4). TkHMGR was enriched from the IP, whereas HMGRs are usually found in the ER (Hampton et al., 1996; Leivar et al., 2005). However, the presence of eight *HMGR* paralogs in the *T. koksaghyz* genome (Lin et al., 2022) and the differential expression of *TbHMGRs* (Van Deenen et al., 2012) suggests there is scope for functional specialization, with at least one HMGR associated with NR synthesis (Campos and Boronat, 1995; Chappell, 1995; Leivar et al., 2005; Lin et al., 2022; McCaskill and Croteau, 1998). The involvement of this TkHMGR paralog in the supply of precursors for NR synthesis is supported by its depletion in *TkCPTL1*-RNAi plants (Table 2) (Niephaus et al., 2019).

Intriguingly, TkSRPP4 and TkSRPP5 also interacted with plastidial methylerythritol (MEP) pathway enzymes, an alternative route to IPP (Vranová et al., 2012). Specifically, TkSRPP4 interacted with 1-deoxy-D-xylulose-5-phosphate synthase (TkDXS8) whereas both TkSRPP4 and TkSRPP5 interacted with 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase (TkHDS1). Although latex does not contain genuine chloroplasts, MEP pathway enzymes and low levels of corresponding mRNAs have been detected in *T. koksaghyz* latex before (Lin et al., 2018; Niephaus et al., 2019). Additionally, specialized plastids known as Frey-Wyssling (F.W.) complexes (Frey-Wyssling, 1929) have been described in *T. koksaghyz* and *H. brasiliensis* latex (Abdul Ghaffar, 2017; Dickenson, 1969; Gomez and Hamzah, 1989; Moir, 1959), and these compartments may comprise the MEP pathway in latex. However, TkDXS8 and TkHDS1 were enriched from the RP, IP and whole latex but not the PP where F.W. complexes would be presumed. This may reflect the disruption of F.W. complexes during processing or the liberation of MEP pathway enzymes by another mechanism, although we would also expect interactions with TkSRPP3 in this scenario due to the significant number of chloroplast-related TkSRPP3 interactors. Feeding experiments in *H. brasiliensis* showed that in this species the MEP pathway contributes to carotenoid rather than NR biosynthesis in latex (Sando et al., 2008). Still, the interactions of TkSRPP3/4/5 with different isoprenoid precursor pathways raise the possibility that they drive IPP flux towards NR synthesis and provide more evidence that TkSRPP4 is a hub protein whereas TkSRPP3 and TkSRPP5 are more specialized.

TkSRPP4 also appears to engage with isoprenoid pathways downstream of FPP by interacting with enzymes involved in sesquiterpene lactone and triterpenoid biosynthesis (González-Coloma et al., 2011; Huber et al., 2015, 2016; Padilla-Gonzalez et al., 2016), the latter including TkSQS1, TkSQE1 and oxidosqualene cyclase 5 (TkOSC5). TkSRPP3 and TkSRPP5 interacted with TkOSC1 and TkOSC5, respectively. Both are latex-specific enzymes and TkOSC1 produces at least four different triterpenoids from 2,3-oxidosqualene, most likely provided by TkSQE1 (Pütter et al., 2019; Unland et al., 2018), whereas TkOSC5 did not produce any triterpenoids in *N. benthamiana* (Pütter et al., 2019). The transcriptional co–regulation of *TkSQS1*, *TkSQE1* and *TkOSC1* facilitates tight metabolic coupling, and TkSQS1 colocalizes with TkSQE1 in the ER of *N. benthamiana* (Unland et al., 2018). Given the presence of transmembrane domains in both proteins, their interaction with TkSRPP4 in the IP may reflect their translocation caused by the phase separation procedure. TkOSC1 has yet to be detected in the ER following heterologous expression (Pütter, 2017). Therefore, TkSRPP3 and TkSRPP4 may cooperatively mediate the assembly of these three consecutive enzymes in the cytosol, ER or on rubber particles for the efficient synthesis of bioactive triterpenoids, given that other OSCs were shown to localize to LDs in yeast (Milla et al., 2003) and triterpenoids are the most abundant non-polyisoprenoid component in separated NR from *T. koksaghyz* (Pütter et al., 2019). TkOSC5 may need to interact with TkSRPPs to maintain stability or activity. TkSRPP3/4/5 could thus affect the quality of NR as an

industrial raw material because the triterpenoid content is proposed to influence the physical properties of the polymer (Xu et al., 2017).

4.2 TkSRPP3/4/5 interactomes suggest their involvement in rubber particle biogenesis, integrity and dispersity

4.2.1 TkSRPP4 protein interactions may contribute to rubber particle biogenesis from the ER

So far, TkSRPP3/4/5 have only been found associated with rubber particles in latex (Collins-Silva et al., 2012) but their presence in different latex phases and Tb/TkSRPP localization in *N. benthamiana* suggest they may also be present in the ER and cytosol (He et al., 2024; Laibach et al., 2018). In support of that, exclusive TkSRPP4 interactors were related to the ER membrane (Figure 4). TkSRPP4 is therefore likely to be the most important component of ER-related SRPP functions, whereas TkSRPP3 and TkSRPP5 cooperate with other TkSRPPs in the ER. These processes could include the transmembrane transport of inorganic molecules because a related GO term was also enriched among TkSRPP4 interactors (Figure 4). Glycerol-3-phosphate acyltransferase (GPAT) was enriched with TkSRPP4 from the PP and was downregulated in *TkCPTL1*-RNAi plants (Niephaus et al., 2019). GPATs catalyze the transfer of an acyl group to the *sn*-1 position of glycerol-3-phosphate leading to the formation of lysophosphatidic acid, and this can be acylated further to phosphatidic acid, the common precursor of other phospholipids and TAGs (Jayawardhane et al., 2018; Testerink and Munnik, 2011). Phosphatidic acid is also important for ROS signaling under biotic stress (Gong et al., 2024). The conversion of lysophosphatidic to phosphatidic acid may cause negative monolayer curvature that could play a role in rubber particle biogenesis (Kooijman et al., 2003). GPAT influences TAG biosynthesis (Gidda et al., 2009; Shockey et al., 2015) and confers tolerance against freezing stress, which often affects *T. koksaghyz* (Kasapoğlu et al., 2024; Sui et al., 2007a, 2007). Some GPATs also contain a phosphatase domain, and *sn*-2-monoacylglycerol was the major product of Arabidopsis GPATs (W. Yang et al., 2010). They have been found in different cellular compartments (Fernández-Santos et al., 2020; Gidda et al., 2009; Kasapoğlu et al., 2024; Sun et al., 2021) and contribute to LD formation, which was associated with their role in TAG biosynthesis (Gao et al., 2013; Wilfling et al., 2013). Rubber particles are not known to store TAGs, but GPAT may contribute to rubber particle budding from the ER via its interaction with TkSRPP4 (Figure 10B). A phosphatase activity and supply of phosphatidic acid could drive phospholipid synthesis and incorporation into the rubber particle or ER membrane. GPAT is therefore an interesting candidate for further analysis of lipid modifications that contribute to rubber particle formation and stress tolerance. Two SEC14 cytosolic factors containing CRAL-TRIOL domains prevalent in lipid-binding proteins (Panagabko et al., 2003) were also identified as TkSRPP4 interactors and were among the proteins downregulated in *TkCPTL1*-RNAi plants (Niephaus et al., 2019). SEC14 proteins are PI/phosphatidylserine transfer proteins that modulate membrane identity, including lipid raft formation (Curwin et al., 2013; Montag et al., 2023). TkSRPP4 may cooperate with these proteins in the ER to accumulate the proteins and lipids needed for rubber particle formation (Figure 10B).

4.2.2 TkSRPP3 and TkSRPP4 protein interactions at the rubber particle may contribute to its integrity and dispersity

The identification of several lipid-modifying enzymes among the TkSRPP3 RP interactors indicates they act on monolayer lipids and the rubber particle lipid composition may be continuously modified and rearranged. The fact that all TkSRPP3 RP interactors are exclusive to this paralog demonstrates TkSRPP functional divergence and potential functional specialization of TkSRPP3 at the rubber particle, which could be mediated by its specific physicochemical properties. TkSRPP4 interactors enriched from the RP comprised several proteins associated with ubiquitination/de-ubiquitination and ubiquitin-dependent proteasomal degradation, suggesting a role in rubber particle protein homeostasis, which is probably required to maintain particle integrity and efficient NR biosynthesis. Several other TkSRPP4 RP interactors were related to sucrose non-fermenting 1 (Snf1) and its plant homolog Snf1-related protein kinase 1 (SnRK1), a major regulator of developmental plasticity including lipid biosynthesis (Jamsheer K et al., 2021). These proteins inactivate HMGR (Robertlee et al., 2017; Sugden et al., 1999) and a key enzyme in PC biosynthesis (Caldo et al., 2019), and also regulate TAG biosynthesis (Zhai et al., 2017). Therefore, the presence of these kinases on the rubber particle surface and their interactions with TkSRPP4 may also influence the lipid composition of the monolayer and the stored NR and triterpenoids.

Further, the interaction of TkSRPP4 with a lectin downregulated in NR-depleted *TkCPTL1*-RNAi plants could play a role in rubber particle dispersity (Table 2) (Niephaus et al., 2019). A latex lectin in *H. brasiliensis* that induces rubber particle aggregation is inhibited by binding to a glycosylated SRPP and the *N*-acetylglucosamine residue of the SRPP was necessary for binding, which is typical for lectins (Rüdiger and Gabius, 2002; Wititsuwannakul et al., 2008). The TkSRPP4 and TkSRPP3 interactomes also featured additional lectins, and the single *N*-glycosylation site found in TkSRPP3 and TkSRPP4 suggests a similar role in rubber particle dispersity that could synergize with the induced steric repulsion (Figure 10A).

The interaction of TkSRPP4 with a homolog of isoprenylcysteine α–carbonyl methylesterase-like 2 (ICMEL2) that was also downregulated in *TkCPTL1*-RNAi plants (Table 2) (Niephaus et al., 2019) may also play a role in rubber particle biogenesis and integrity. Proteins can be C-terminally prenylated by the addition of farnesyl or geranylgeranyl groups to a cysteine to increase their membrane affinity, usually followed by methylation in the ER, which can be reversed by ICMEs (Clarke, 1992; Crowell, 2000; Lan et al., 2010; Zhang and Casey, 1996). The methylation status can affect protein–lipid interactions and prenylation can affect protein–protein interactions (Crowell, 2000; Hancock et al., 1991; Kuroda et al., 1993; Sapperstein et al., 1994; Zhang and Casey,

1996). Such modifications may therefore be important for protein recruitment to the ER or rubber particles mediated by TkSRPP4/ICMEL2 complexes.

4.2.3 TkSRPP5 interaction with GTPases may promote rubber particle formation

Proteins that undergo prenylation for membrane targeting include GTPases (Chavrier et al., 1991; Hancock et al., 1991; Kuroda et al., 1993), which were enriched among the exclusive TkSRPP5 interactors and part of TkSRPP3/4/5 RP interactomes (Figure 4; Supplementary Data S1–S3 in Supplementary Data Sheet 1). GTPases regulate multiple cellular processes, especially signal transduction and vesicle transport, but also play a major role in immune responses (Kawano et al., 2014; Nielsen, 2020). They switch between inactive GDP-bound and active GTP-bound states, in which they are prenylated and associate with membranes, allowing them to engage with effector proteins (Grosshans et al., 2006; Kawano et al., 2014). TkSRPP5 interactions with GTPases may regulate GTP/GDP binding or membrane interactions, or the GTPases may recruit TkSRPP5 as an effector protein to change membrane lipid distribution. GTPases also recruit proteins that induce vesicle formation (Huang et al., 2001; Kawasaki et al., 2005; Spang, 2008; Springer et al., 1999) and the recruitment of TkSRPP5 and other proteins could thus induce rubber particle budding from the ER (Figure 10B), aligning with the identification of GTPases on rubber particles from *H. brasiliensis* (Nielsen et al., 2008; Yamashita et al., 2016). The interaction of TkSRPP5 with a putative methyltransferase that was downregulated in *TkCPTL1*-RNAi lines (Niephaus et al., 2019) could indicate a role in the methylation of prenylated proteins such as GTPases. Therefore, TkSRPP4 and TkSRPP5 may regulate the methylation/demethylation of prenylated latex proteins that affect membrane association and protein interactions. Further, TkSRPP5 interactions with lipid-modifying proteins (Figure 4) could help to establish the membrane conditions needed for rubber particle biogenesis. Beyond that, GO enrichment analysis revealed that exclusive TkSRPP5 interactors were enriched for the term 'microtubule cytoskeleton' (Figure 4). Such proteins are present on rubber particles in *H. brasiliensis* (Dai et al., 2013) and the interaction of TkSRPP5 with those proteins may facilitate the transport of rubber particles along the cytoskeleton to the vacuole, which contains many cytoskeleton-related proteins in the tonoplast (Carter et al., 2004) and stores rubber particles in *T. koksaghyz* (Abdul Ghaffar, 2017).

4.3 Further evidence for the involvement of TkSRPPs and TkSRPP heterodimers in stress responses

Common TkSRPP3/4/5 interactors were related to membrane and vesicular trafficking, lipid metabolism and stress responses (Supplementary Table S6). TkSRPP3/4/5 may therefore promote stress tolerance by affecting lipid modification, the proteolytic cleavage of pathogen-derived proteins, the inhibition of pathogen-derived proteases and/or cellular adaptations by membrane trafficking. Rab7 proteins, for example, are involved in vacuolar trafficking and improve abiotic stress tolerance when overexpressed (Mazel et al., 2004; Rodriguez-Furlan et al., 2023; Tripathy et al., 2017).

As the protein TkSRPP7 was enriched from whole latex and the IP by TkSRPP3/4/5, despite the relatively low *TkSRPP7* transcript levels in the latex of 10-week-old plants (Figure 1B) and the lack of evidence showing its association with rubber particles, this paralog may interact with TkSRPP3/4/5 elsewhere in the laticifers as part of stress-related processes. This may also explain why TkSRPP2 interacted with TkSRPP4 and TkSRPP5 in all latex fractions, and why TkSRPP1 interacted with TkSRPP3 in whole latex (Supplementary Table S7). The roles of these TkSRPPs in stress responses are supported by their transcriptional upregulation following treatment with MeJA or the overexpression of *TkMYC2* (He et al., 2024; Wu et al., 2024). TkSRPPs may also form multimers to fulfil their functions, as suggested for HbSRPP (Wititsuwannakul et al., 2008).

The overrepresentation of proteins associated with the chloroplast stroma and thylakoids among exclusive TkSRPP3 interactors was striking (Figure 4). TkSRPP3 and its interactors related to the chloroplast stroma and thylakoids may associate with the plastid-like F.W. complexes, explaining why most of these proteins were enriched from the PP. Notably, rubber particles were also observed within plastid-like structures in *T. koksaghyz* laticifers (Abdul Ghaffar, 2017), so TkSRPP3 and its chloroplast-related latex interactors may also be connected with such rubber particles. The presence of ferritins, peroxiredoxins, glutathione S-transferase (GST) and aconitate hydratase among those TkSRPP3 interactors indicate roles in antioxidant defense, redox regulation and detoxification (Briat, 1996; Dietz, 2003; Kumar et al., 2024; Kumar and Trivedi, 2018; Moeder et al., 2007; Pascual et al., 2021). Ferritin inhibits the formation of reactive oxygen species (ROS) by Fe (Halliwell and Gutteridge, 1984; Kroh and Pilon, 2020; Ravet et al., 2009), peroxiredoxins are antioxidants that mediate redox-dependent signaling (Liebthal et al., 2018; Sevilla et al., 2015), and GSTs counter oxidative stress by conjugating glutathione (Cozza et al., 2017; Dixon and Edwards, 2009; Labrou et al., 2015; Nianiou-Obeidat et al., 2017; Wagner et al., 2002). Overexpression of these genes confers abiotic and biotic stress tolerance (Deák et al., 1999; Lo Cicero et al., 2017; Roxas et al., 2000; Wang et al., 2023; Xiao et al., 2023; Zang et al., 2017), and their endogenous promoters are responsive to stress and phytohormones (Dellagi et al., 2005; García Mata et al., 2001; Horling et al., 2002, 2003; Marrs, 1996; Tiwari et al., 2016). GST and aconitate hydratase are also associated with cadmium stress tolerance, a significantly enriched process among TkSRPP3 interactors (Figure 4) that has also been experimentally linked to REF proteins (Dixit et al., 2011; Kim et al., 2011; Liu et al., 2013, 2016; Zhou et al., 2019). These interactors may therefore form a TkSRPP3-dependent network of stress tolerance effectors within F.W. complexes (Cerveau et al., 2016; Manevich et al., 2004), as illustrated in Figure 10C. The large number of kinases among the TkSRPP3 interactors suggests the effector network is regulated by kinase cascades, supported by multiple phosphorylation sites on TkSRPP3. Finally, the glutathione peroxidase activity of GSTs prevents lipid oxidation (Dixon et al., 2009; Dixon and Edwards, 2009; Ohkama-Ohtsu et al., 2011) and

could be particularly valuable in latex that contains diverse lipids, many associated with bioactive properties (Bae et al., 2020; González-Coloma et al., 2011; Pütter et al., 2019).

The involvement of TkSRPP5 in stress response is supported by the enrichment of interactors associated with responses to different stimuli (Figure 4). The activation of TkSRPP5 in response to such stimuli could be achieved by phosphorylation at the six predicted phosphorylation sites or by *N*-glycosylation (Figure 1E).

4.4 TkSRPP6 forms heteromeric complexes with TkSRPP4 and TkSRPP5

TkSRPP6 was one of two candidate interactors that we chose for further analysis, due to its isolated genomic locus and phylogenetic clustering with stress-related proteins from non-rubber-producing plants rather than other TkSRPPs (Figure 5C). The *TkSRPP6* gene was also expressed at lower levels than *TkSRPP3/4/5* in all tissues (and in latex over time) (Figures 1, 7). SUY2H results confirmed that TkSRPP6 interacted with TkSRPP4 and TkSRPP5, but not TkSRPP3 (Figure 5A). The interaction of TkSRPP6 with the abundant latex proteins TkSRPP4 and TkSRPP5, despite their different molecular characteristics, was striking. The identification of MYC2-binding sites in the *TkSRPP6* promoter suggested inducible expression, supported by the transcriptional induction observed after MeJA treatment and *TkMYC2* overexpression (Figures 1, 5D) (*TkSRPP1* in He et al., 2024; *TkSRPP7* in Wu et al., 2024). This induction might be transient, as shown for MeJA (He et al., 2024), and would therefore not show up in our qPCR data for wild-type plants. Stress-induced expression is a common feature of the REF family and is supported by the homology of TkSRPP6 to stress-related proteins (Figure 5C). The heterologous expression of TkSRPP6 in *N. benthamiana* resulted in localization to the ER and LDs, matching its enrichment from the RP by TkSRPP3 (Figures 8A, B; Table 3). A recent study (He et al., 2024) suggested TkSRPP6 (named TkSRPP1 therein) was localized to the cytosol, plasma membrane and chloroplast, but the authors did not induce LD formation nor did they use plasma membrane and ER markers to confirm their assumptions. We observed no plastid signals for our TkSRPP6-Cerulean fusion protein (Figures 8A, B). TbSRRPs were found to be localized to the cytosol (Laibach et al., 2018) and TkSRPP6 enrichment from the IP supports a cytosolic localization. We conclude that TkSRPP6 may have affinities for different cellular compartments that may change depending on specific conditions.

The localization of TbSRPPs 4 and 5 and TkSRPP6 together with indications of inducible gene expression suggest that TkSRPP6 may interact at basal levels with TkSRPP4 and TkSRPP5 on the surface of the ER and rubber particles, but predominantly after its short-term transcriptional induction in response to various stimuli. These interactions most likely play a role in stress tolerance, and TkSRPP6 could engage with established complexes formed by TkSRPP4/TkSRPP5. TkSRPP6 may respond to stress not only in the latex but also in green tissues as its transcript levels were comparable low in all tissues.

4.5 TkSRPP3 and TkSRPP5 interact with TkUGT80B1 potentially contributing to plant stress tolerance by producing triterpenoid saponins

TkUGT80B1 was the second candidate selected for in depth characterization because glycosyltransferases have not been described in dandelion latex before and glycosides are involved in stress responses (Augustin et al., 2011; Rogowska and Szakiel, 2020). An independent co-IP assay confirmed that TkUGT80B1 interacts with TkSRPP3 and TkSRPP5 (Figure 6A). The interactions were further corroborated by the similar temporal expression patterns of *TkUGT80B1* and *TkSRPP3/4/5* in latex (Figure 7) (Ge et al., 2001; Jansen et al., 2002; von Mering et al., 2002).

The predicted glycosyltransferase activity of TkUGT80B1 was confirmed for the C_3 position of the triterpenoid lupeol in yeast, with UDP-glucose as the most likely sugar donor (Figure 9). This was supported by the presence of glutamine in the C-terminal UDPGT motif, which is conserved among glucosyltransferases but not galactosyltransferases (Figure 6D) (Kubo et al., 2004). Accordingly, we have identified the first enzyme from the latex of *T. koksaghyz* that produces a triterpenoid saponin and have provided first evidence for the presence of these compounds in dandelions. Given the low substrate specificities of UGTs (Vogt and Jones, 2000), the glycosylation of additional, structurally similar triterpenoids in *T. koksaghyz* latex is likely. The heterologous expression of TkUGT80B1 did not result in the glycosylation of yeast sterols when sterol synthesis was repressed (Figure 9). It is unclear whether TkUGT80B1 can also utilize phytosterol substrates, as described for its homologs (Stucky et al., 2014), or other lipids in latex. The analysis of Arabidopsis UGTs indicated that AtUGT80B1 is not required for the synthesis of major steryl glucosides but rather for the production of minor glucosides (Stucky et al., 2014). Accordingly, TkUGT80B1 may glycosylate triterpenoids rather than major membrane sterols in accordance with its predominant expression in latex (Figure 6E). Adaptation to freezing stress was inhibited in *atugt80b1* knockout plants, and was potentially related to low levels of sterol glycosides, the products of AtUGT80B1 (Mishra et al., 2015). It would be interesting to determine whether triterpenoid saponins have a similar positive effect on freezing tolerance in *T. koksaghyz* because the lipid–rich latex in its roots has already been proposed to act as an anti-freezing protectant during extremely cold winters, which are common in its native habitat.

The presence of several stress-related *cis*-acting regulatory elements in the *TkUGT80B1* promoter, including binding sites for WRKY transcription factors, suggests the gene is transcriptionally regulated in response to biotic stress (Figure 6D). WRKY transcription factors can affect defense response positively as well as negatively (Javed and Gao, 2023; Wani et al., 2021), but given the reported positive correlation between SRPPs and stress tolerance (Kim et al., 2016, 2012; Laibach et al., 2018; Seo et al., 2010), TkUGT80B1 is also likely to improve stress tolerance, although this should be investigated in more detail.

Saponins are stored in the vacuole (Kesselmeier and Urban, 1983; Mylona et al., 2008; Urban et al., 1983). *TkUGT80B1* expression in *N. benthamiana* suggested tonoplast localization

(Figure 8C) and the enrichment of TkUGT80B1 from the PP by TkSRPP3 supports this finding. Enrichment from the IP by TkSRPP4 and TkSRPP5 may reflect the rupture of vacuoles during phase separation. Tonoplast localization may allow the immediate storage of nascent saponins in the vacuole. TkSRPP3 and TkSRPP5 could mediate the transfer of saponins from their biosynthetic enzyme to a transporter by linking both proteins. The transport of saponins to the vacuole may be mediated by ABC-type transporters (Kato et al., 2022; Ramilowski et al., 2013) and two proteins with homology to ABC-type transporters were identified as TkSRPP4 interactors, one of which also interacted with TkSRPP5. Interactions between TkSRPP4/TkSRPP5 and these transporters could also promote saponin efflux from the vacuole to promote stress tolerance. Localization studies in *N. benthamiana* showed that TkUGT80B1 can also accumulate in LDs, suggesting it might be located on the surface of rubber particles and interact with TkSRPPs there. The interaction between TkSRPPs and TkUGT80B1 may also promote the metabolic flux towards triterpenoid saponin synthesis by linking the triterpenoid synthesizing TkOSC1/5 with the glycosylating TkUGT80B1. TkSRPP3 and TkSRPP5 could further recruit either TkUGT80B1 or specific lipid substrates to direct glycoside synthesis and ultimately modify the composition of membranes in response to environmental conditions.

5 Conclusion

Our study sheds light on the *SRPP* gene family in *T. koksaghyz* and presents a comprehensive analysis of the protein interaction partners of the major latex proteins TkSRPP3/4/5. We identified protein interactions that suggest TkSRPP3/4/5 contribute directly to increased stress tolerance as well as rubber particle biogenesis and integrity. Two candidates were characterized at the molecular level, revealing the first evidence for saponin synthesis in *T. koksaghyz* latex and linking it with TkSRPP3 and TkSRPP5. Our data contribute to the functional differentiation between TkSRPP paralogs and demonstrate unexpected interactions that will help to further identify the network of proteins linking TkSRPPs, stress responses and NR biosynthesis. These new insights into the complexity of latex will eventually help to establish commercially feasible rubber crops.

Data availability statement

MS raw data are available for review at the following URL (https://repository.jpostdb.org/preview/48789265366a7c7a166d94) and will be made publicly available under the identifier JPST003234 upon publication. All other data presented in this report are available either in the supplemental tables, figures or data, or in publicly available databases.

Author contributions

SMW: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. NL: Investigation, Methodology, Writing – review & editing. JR: Investigation, Writing – review & editing. K-UR: Data Curation, Writing – review & editing. BM: Conceptualization, Investigation, Methodology, Resources, Writing – review & editing. JE: Investigation, Data Curation, Writing – review & editing. RMT: Writing – review & editing. IF: Resources, Writing – review & editing. DP: Funding acquisition, Resources, Writing – review & editing. CSG: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, 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 DFG grants GZ: INST 211/1037-1 LAGG and INST 211/744-1 FUGG. Part of the work was funded by the Federal Ministry for Education and Research (grant number: 031B0059C).

Acknowledgments

We thank Denise Weinberg for her great and continuous technical support and Jos Cox (both Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Münster) for many helpful discussions and critical thoughts, which considerably contributed to the development of this work. We would also like to thank Paulina Heinkow (Institute for Plant Biology and Biotechnology, Münster) for her technical assistance.

Conflict of interest

Author RMT was employed by the company TRM Ltd.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1498737/full#supplementary-material

References

Abdul Ghaffar, M. A. B. (2017). *Rubber Particle Ontogeny in Taraxacum kok-saghyz* (USA: Ohio State University). Available at: http://rave.ohiolink.edu/etdc/view?acc_num=osu1512031318000982.

Agatep, R., Kirkpatrick, R. D., Parchaliuk, D. L., Woods, R. A., and Gietz, R. D. (1998). Transformation of Saccharomyces cerevisiae by the lithium acetate/single-stranded carrier DNA/polyethylene glycol protocol. *Tech. Tips Online* 3, 133–137. doi: 10.1016/s1366-2120(08)70121-1

Alexa, A., and Rahnenfuhrer, J. (2022). topGO: Enrichment Analysis for Gene Ontology *(R package version 2.50.0)*.

Augustin, J. M., Kuzina, V., Andersen, S. B., and Bak, S. (2011). Molecular activities, biosynthesis and evolution of triterpenoid saponins. *Phytochemistry* 72, 435–457. doi: 10.1016/J.PHYTOCHEM.2011.01.015

Bae, S. W., Jung, S., Choi, S. C., Kim, M. Y., and Ryu, S. B. (2020). Lipid Composition of Latex and Rubber Particles in Hevea brasiliensis and Taraxacum kok-saghyz. *Mol. (Basel Switzerland)* 25, 5110. doi: 10.3390/molecules25215110

Bassard, J. E., Mutterer, J., Duval, F., and Werck-Reichhart, D. (2012). A novel method for monitoring the localization of cytochromes P450 and other endoplasmic reticulum membrane associated proteins: A tool for investigating the formation of metabolons. *FEBS J.* 279, 1576–1583. doi: 10.1111/j.1742-4658.2011.08312.x

Batistič, O., Waadt, R., Steinhorst, L., Held, K., and Kudla, J. (2010). CBL-mediated targeting of CIPKs facilitates the decoding of calcium signals emanating from distinct cellular stores. *Plant J.* 61, 211–222. doi: 10.1111/J.1365-313X.2009.04045.X

Benninghaus, V. A., Van Deenen, N., Müller, B., Roelfs, K. U., Lassowskat, I., Finkemeier, I., et al. (2020). Comparative proteome and metabolome analyses of latex-exuding and non-exuding Taraxacum koksaghyz roots provide insights into laticifer biology. *J. Exp. Bot.* 71, 1278. doi: 10.1093/JXB/ERZ512

Berthelot, K., Lecomte, S., Estevez, Y., and Peruch, F. (2014). Hevea brasiliensis REF (Hev b 1) and SRPP (Hev b 3): An overview on rubber particle proteins. *Biochimie* 106, 1–9. doi: 10.1016/J.BIOCHI.2014.07.002

Briat, J. F. (1996). Roles of ferritin in plants. *J. Plant Nutr.* 19, 1331–1342. doi: 10.1080/01904169609365202

Bröker, J. N., Müller, B., van Deenen, N., Prüfer, D., and Schulze Gronover, C. (2018). Upregulating the mevalonate pathway and repressing sterol synthesis in Saccharomyces cerevisiae enhances the production of triterpenes. *Appl. Microbiol. Biotechnol.* 102, 6923–6934. doi: 10.1007/s00253-018-9154-7

Brown, D., Feeney, M., Ahmadi, M., Lonoce, C., Sajari, R., Di Cola, A., et al. (2017). Subcellular localization and interactions among rubber particle proteins from Hevea brasiliensis. *J. Exp. Bot.* 68, 5045–5055. doi: 10.1093/JXB/ERX331

Caldo, K. M. P., Xu, Y., Falarz, L., Jayawardhane, K., Acedo, J. Z., and Chen, G. (2019). Arabidopsis CTP:phosphocholine cytidylyltransferase 1 is phosphorylated and inhibited by sucrose nonfermenting 1–related protein kinase 1 (SnRK1). *J. Biol. Chem.* 294, 15862. doi: 10.1074/JBC.RA119.008047

Campos, N., and Boronat, A. (1995). Targeting and topology in the membrane of plant 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Plant Cell* 7, 2163–2174. doi: 10.1105/TPC.7.12.2163

Cantalapiedra, C. P., Herņandez-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol. Biol. Evol.* 38, 5825–5829. doi: 10.1093/MOLBEV/MSAB293

Cao, X. W., Yan, J., Lei, J. L., Li, J., Zhu, J. B., and Zhang, H. Y. (2017). *De novo* Transcriptome Sequencing of MeJA-Induced Taraxacum koksaghyz Rodin to Identify Genes Related to Rubber Formation. *Sci. Rep.* 7, 1–13. doi: 10.1038/s41598-017-14890-z

Carter, C., Pan, S., Zouhar, J., Avila, E. L., Girke, T., and Raikhel, N. V. (2004). The vegetative vacuole proteome of Arabidopsis thaliana reveals predicted and unexpected proteins. *Plant Cell* 16, 3285. doi: 10.1105/TPC.104.027078

Cerveau, D., Kraut, A., Stotz, H. U., Mueller, M. J., Couté, Y., and Rey, P. (2016). Characterization of the Arabidopsis thaliana 2-Cys peroxiredoxin interactome. *Plant Sci.* 252, 30–41. doi: 10.1016/J.PLANTSCI.2016.07.003

Chappell, J. (1995). Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46, 521–547. doi: 10.1146/annurev.pp.46.060195.002513

Chavrier, P., Gorvel, J. P., Stelzer, E., Simons, K., Gruenberg, J., and Zerial, M. (1991). Hypervariable C-termmal domain of rab proteins acts as a targeting signal. *Nature* 353, 769–772. doi: 10.1038/353769a0

Cho, S. H., Tóth, K., Kim, D., Vo, P. H., Lin, C. H., Handakumbura, P. P., et al. (2022). Activation of the plant mevalonate pathway by extracellular ATP. *Nat. Commun.* 13, 450. doi: 10.1038/S41467-022-28150-W

Clarke, S. (1992). Protein isoprenylation and methylation at carboxyl-terminal cysteine residues. *Annu. Rev. Biochem.* 61, 355–386. doi: 10.1146/ANNUREV.BI.61.070192.002035/CITE/REFWORKS

Collins-Silva, J., Nural, A. T., Skaggs, A., Scott, D., Hathwaik, U., Woolsey, R., et al. (2012). Altered levels of the Taraxacum kok-saghyz (Russian dandelion) small rubber particle protein, TkSRPP3, result in qualitative and quantitative changes in rubber metabolism. *Phytochemistry* 79, 46–56. doi: 10.1016/j.phytochem.2012.04.015

Cornish, K., Wood, D. F., and Windle, J. J. (1999). Rubber particles from four different species, examined by transmission electron microscopy and electron-paramagnetic-resonance spin labeling, are found to consist of a homogeneous rubber core enclosed by a contiguous, monolayer biomembrane. *Planta* 210, 85–96. doi: 10.1007/s004250050657

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372. doi: 10.1038/nbt.1511

Cozza, G., Rossetto, M., Bosello-Travain, V., Maiorino, M., Roveri, A., Toppo, S., et al. (2017). Glutathione peroxidase 4-catalyzed reduction of lipid hydroperoxides in membranes: The polar head of membrane phospholipids binds the enzyme and addresses the fatty acid hydroperoxide group toward the redox center. *Free Radical Biol. Med.* 112, 1–11. doi: 10.1016/J.FREERADBIOMED.2017.07.010

Crowell, D. N. (2000). Functional implications of protein isoprenylation in plants. *Prog. Lipid Res.* 39, 393–408. doi: 10.1016/S0163-7827(00)00010-2

Curwin, A. J., LeBlanc, M. A., Fairn, G. D., and McMaster, C. R. (2013). Localization of lipid raft proteins to the plasma membrane is a major function of the phospholipid transfer protein Sec14. *PloS One* 8, e55388. doi: 10.1371/JOURNAL.PONE.0055388

Dai, L., Kang, G., Li, Y., Nie, Z., Duan, C., and Zeng, R. (2013). In-depth proteome analysis of the rubber particle of Hevea brasiliensis (para rubber tree). *Plant Mol. Biol.* 82, 155–168. doi: 10.1007/s11103-013-0047-y

Deák, M., Horváth, G. V., Davletova, S., Török, K., Sass, L., Vass, I., et al. (1999). Plants ectopically expressing the ironbinding protein, ferritin, are tolerant to oxidative damage and pathogens. *Nat. Biotechnol.* 17, 192–196. doi: 10.1038/6198

Dellagi, A., Rigault, M., Segond, D., Roux, C., Kraepiel, Y., Cellier, F., et al. (2005). Siderophore-mediated upregulation of Arabidopsis ferritin expression in response to Erwinia chrysanthemi infection. *Plant J.* 43, 262–272. doi: 10.1016/S0021-9258(18)51510-6

Dennis, M. S., and Light, D. R. (1989). Rubber elongation factor from Hevea brasiliensis. Identification, characterization, and role in rubber biosynthesis. *J. Biol. Chem.* 264, 18608–18617. doi: 10.1111/J.1365-313X.2005.02451.X

Dickenson, P. B. (1969). Electron microscopial studies of latex vessel system of Hevea brasiliensis. *J. Rubber Res.* 21, 543–559.

Dietz, K. J. (2003). Plant peroxiredoxins. *Annu. Rev. Plant Biol.* 54, 93–107. doi: 10.1146/ANNUREV.ARPLANT.54.031902.134934/CITE/REFWORKS

Dixit, P., Mukherjee, P. K., Ramachandran, V., and Eapen, S. (2011). Glutathione Transferase from Trichoderma virens Enhances Cadmium Tolerance without Enhancing Its Accumulation in Transgenic Nicotiana tabacum. *PloS One* 6, e16360. doi: 10.1371/JOURNAL.PONE.0016360

Dixon, D. P., and Edwards, R. (2009). Selective binding of glutathione conjugates of fatty acid derivatives by plant glutathione transferases. *J. Biol. Chem.* 284, 21249–21256. doi: 10.1074/jbc.M109.020107

Dixon, D. P., Hawkins, T., Hussey, P. J., and Edwards, R. (2009). Enzyme activities and subcellular localization of members of the Arabidopsis glutathione transferase superfamily. *J. Exp. Bot.* 60, 1207–1218. doi: 10.1093/JXB/ERN365

Dong, G., Fan, M., Wang, H., Leng, Y., Sun, J., Huang, J., et al. (2023a). Functional characterization of TkSRPP promoter in response to hormones and wounding stress in transgenic tobacco. *Plants* 12, 252. doi: 10.3390/PLANTS12020252

Dong, G., Wang, H., Qi, J., Leng, Y., Huang, J., Zhang, H., et al. (2023b). Transcriptome analysis of Taraxacum kok-saghyz reveals the role of exogenous methyl jasmonate in regulating rubber biosynthesis and drought tolerance. *Gene* 867, 147346. doi: 10.1016/J.GENE.2023.147346

Epping, J., van Deenen, N., Niephaus, E., Stolze, A., Fricke, J., Huber, C., et al. (2015). A rubber transferase activator is necessary for natural rubber biosynthesis in dandelion. *Nat. Plants* 1. doi: 10.1038/nplants.2015.48

Eyal, Y., Meller, Y., Lev-Yadun, S., and Fluhr, R. (1993). A basic-type PR-1 promoter directs ethylene responsiveness, vascular and abscission zone-specific expression. *Plant J.* 4, 225–234. doi: 10.1046/J.1365-313X.1993.04020225.X

Fernández-Santos, R., Izquierdo, Y., López, A., Muñiz, L., Martínez, M., Cascón, T., et al. (2020). Protein Profiles of Lipid Droplets during the Hypersensitive Defense Response of Arabidopsis against Pseudomonas Infection. *Plant Cell Physiol.* 61, 1144–1157. doi: 10.1093/PCP/PCAA041

Frey-Wyssling, A. (1929). Microscopic investigations on the occurrence of resins in Hevea latex. *Arch. Rubbercult* 13, 392–294.

Fricke, J., Hillebrand, A., Twyman, R. M., Prüfer, D., and Schulze Gronover, C. (2013). Abscisic acid-dependent regulation of small rubber particle protein gene expression in Taraxacum brevicorniculatum is mediated by TbbZIP1. *Plant Cell Physiol.* 54, 448–464. doi: 10.1093/pcp/pcs182

Galvão, V. C., Fiorucci, A. S., Trevisan, M., Franco-Zorilla, J. M., Goyal, A., Schmid-Siegert, E., et al. (2019). PIF transcription factors link a neighbor threat cue to accelerated reproduction in Arabidopsis. *Nat. Commun.* 10, 1–10. doi: 10.1038/s41467-019-11882-7

Gao, Q., Shang, Y., Huang, W., and Wang, C. (2013). Glycerol-3-phosphate acyltransferase contributes to triacylglycerol biosynthesis, lipid droplet formation,

and host invasion in Metarhizium robertsii. *Appl. Environ. Microbiol.* 79, 7646–7653. doi: 10.1128/AEM.02905-13/SUPPL_FILE/ZAM999104940SO1.PDF

García Mata, C., Lamattina, L., and Cassia, R. O. (2001). Involvement of iron and ferritin in the potato-Phytophthora infestans interaction. *Eur. J. Plant Pathol.* 107, 557–562. doi: 10.1023/A:1011228317709/METRICS

Ge, H., Liu, Z., Church, G. M., and Vidal, M. (2001). Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. *Nat. Genet.* 29, 482–486. doi: 10.1038/ng776

Gidda, S. K., Park, S., Pyc, M., Yurchenko, O., Cai, Y., Wu, P., et al. (2016). Lipid droplet-associated proteins (LDAPs) are required for the dynamic regulation of neutral lipid compartmentation in plant cells. *Plant Physiol.* 170, 2052–2071. doi: 10.1104/pp.15.01977

Gidda, S. K., Shockey, J. M., Rothstein, S. J., Dyer, J. M., and Mullen, R. T. (2009). Arabidopsis thaliana GPAT8 and GPAT9 are localized to the ER and possess distinct ER retrieval signals: Functional divergence of the dilysine ER retrieval motif in plant cells. *Plant Physiol. Biochem.* 47, 867–879. doi: 10.1016/J.PLAPHY.2009.05.008

Gidda, S. K., Watt, S. C., Collins-Silva, J., Kilaru, A., Arondel, V., Yurchenko, O., et al. (2013). Lipid droplet-associated proteins (LDAPs) are involved in the compartmentalization of lipophilic compounds in plant cells. *Plant Signal. Behav.* 8, e27141-1-e27141-4. doi: 10.4161/psb.27141

Gomez, J. B., and Hamzah, S. (1989). Frey-Wyssling complex in hevea latex - uniqueness of the organelle. *J. Natural Prod.* 4, 75–85. doi: 10.1007/978-3-319-59379-1%0

Gong, Q., Yao, S., Wang, X., and Li, G. (2024). Fine-tuning phosphatidic acid production for optimal plant stress responses. *Trends Biochem. Sci* 49 (8), 663–666. doi: 10.1016/j.tibs.2024.05.008

González-Coloma, A., López-Balboa, C., Santana, O., Reina, M., and Fraga, B. M. (2011). Triterpene-based plant defenses. *Phytochem. Rev.* 10, 245–260. doi: 10.1007/S11101-010-9187-8/TABLES/4

Griseti, E., Bello, A. A., Bieth, E., Sabbagh, B., Iacovoni, J. S., Bigay, J., et al. (2024). Molecular mechanisms of perilipin protein function in lipid droplet metabolism. *FEBS Lett.* 598, 1170–1198. doi: 10.1002/1873-3468.14792

Grosshans, B. L., Ortiz, D., and Novick, P. (2006). Rabs and their effectors: Achieving specificity in membrane traffic. *Proc. Natl. Acad. Sci. United States America* 103, 11821–11827. doi: 10.1073/PNAS.0601617103/SUPPL_FILE/01617FIG8.PDF

Guiltinan, M. J., Marcotte, W. R., and Quatrano, R. S. (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. *Sci. (New York N.Y.)* 250, 267–271. doi: 10.1126/SCIENCE.2145628

Guo, D., Li, H. L., Tang, X., and Peng, S. Q. (2014). Molecular and functional characterization of the HbSRPP promoter in response to hormones and abiotic stresses. *Transgenic Res.* 23, 331–340. doi: 10.1007/S11248-013-9753-0/FIGURES/4

Ha, J. H., and Loh, S. N. (2012). Protein conformational switches: from nature to design. *Chem. (Weinheim an Der Bergstrasse Germany)* 18, 7984. doi: 10.1002/CHEM.201200348

Halliwell, B., and Gutteridge, J. M. C. (1984). Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochem. J.* 219, 1. doi: 10.1042/BJ2190001

Hampton, R. Y., Koning, A., Wright, R., and Rine, J. (1996). *In vivo* examination of membrane protein localization and degradation with green fluorescent protein. *Proc. Natl. Acad. Sci. United States America* 93, 828. doi: 10.1073/PNAS.93.2.828

Hancock, J. F., Cadwallader, K., and Marshall1, C. J. (1991). Methylation and proteolysis are essential for efficient membrane binding of prenylated p21K-ras(B). *EMBO J.* 10, 641–646. doi: 10.1002/J.1460-2075.1991.TB07992.X

Harayama, T., and Riezman, H. (2018). Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* 19, 281–296. doi: 10.1038/nrm.2017.138

He, H., Wang, J., Meng, Z., Dijkwel, P. P., Du, P., Shi, S., et al. (2024). Genome-Wide Analysis of the SRPP/REF Gene Family in Taraxacum kok-saghyz Provides Insights into Its Expression Patterns in Response to Ethylene and Methyl Jasmonate Treatments. *Int. J. Mol. Sci.* 25, 6864. doi: 10.3390/IJMS25136864

Heberle, H., Meirelles, V. G., da Silva, F. R., Telles, G. P., and Minghim, R. (2015). InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinf.* 16, 1–7. doi: 10.1186/s12859-015-0611-3

Heim, M. A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., and Bailey, P. C. (2003). The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* 20, 735–747. doi: 10.1093/MOLBEV/MSG088

Herman, E. M. (2008). Endoplasmic reticulum bodies: solving the insoluble. *Curr. Opin. Plant Biol.* 11, 672–679. doi: 10.1016/j.pbi.2008.08.004

Hillebrand, A., Post, J. J., Wurbs, D., Wahler, D., Lenders, M., Krzyzanek, V., et al. (2012). Down-regulation of small rubber particle protein expression affects integrity of rubber particles and rubber content in Taraxacum brevicorniculatum. *PloS One* 7, e41874. doi: 10.1371/journal.pone.0041874

Hong, J. C., Cheong, Y. H., Nagao, R. T., Bahk, J. D., Key, J. L., and Cho, M. J. (1995). Isolation of two soybean G-box binding factors which interact with a G-box sequence of an auxin-responsive gene. *Plant J.* 8, 199–211. doi: 10.1046/J.1365-313X.1995.08020199.X

Horling, F., König, J., and Dietz, K. J. (2002). Type II peroxiredoxin C, a member of the peroxiredoxin family of Arabidopsis thaliana: its expression and activity in comparison with other peroxiredoxins. *Plant Physiol. Biochem.* 40, 491–499. doi: 10.1016/S0981-9428(02)01396-7

Horling, F., Lamkemeyer, P., König, J., Finkemeier, I., Kandlbinder, A., Baier, M., et al. (2003). Divergent light-, ascorbate-, and oxidative stress-dependent regulation of expression of the peroxiredoxin gene family in Arabidopsis. *Plant Physiol.* 131, 317–325. doi: 10.1104/PP.010017

Horn, P. J., James, C. N., Gidda, S. K., Kilaru, A., Dyer, J. M., Mullen, R. T., et al. (2013). Identification of a new class of lipid droplet-associated proteins in plants. *Plant Physiol.* 162, 1926–1936. doi: 10.1104/pp.113.222455

Huang, A. H. C. (2018). Plant lipid droplets and their associated proteins: Potential for rapid advances. *Plant Physiol.* 176, 1894–1918. doi: 10.1104/pp.17.01677

Huang, M., Weissman, J. T., Béraud-Dufour, S., Luan, P., Wang, C., Chen, W., et al. (2001). Crystal structure of Sar1-GDP at 1.7 A resolution and the role of the NH2 terminus in ER export. *J. Cell Biol.* 155, 937–948. doi: 10.1083/JCB.200106039

Huber, M., Epping, J., Schulze Gronover, C., Fricke, J., Aziz, Z., Brillatz, T., et al. (2016). A latex metabolite benefits plant fitness under root herbivore attack. *PloS Biol.* 14, e1002332. doi: 10.1371/journal.pbio.1002332

Huber, M., Triebwasser-Freese, D., Reichelt, M., Heiling, S., Paetz, C., Chandran, J. N., et al. (2015). Identification, quantification, spatiotemporal distribution and genetic variation of major latex secondary metabolites in the common dandelion (Taraxacum officinale agg.). *Phytochemistry* 115, 89–98. doi: 10.1016/J.PHYTOCHEM.2015.01.003

Hussain, M., Debnath, B., Qasim, M., Bamisile, B. S., Islam, W., Hameed, M. S., et al. (2019). Role of saponins in plant defense against specialist herbivores. *Molecules* 24, 2067. doi: 10.3390/molecules24112067

Jach, G., Pesch, M., Richter, K., Frings, S., and Uhrig, J. F. (2006). An improved mRFP1 adds red to bimolecular fluorescence complementation. *Nat. Methods* 3, 597–600. doi: 10.1038/nmeth901

Jamsheer K, M., Kumar, M., and Srivastava, V. (2021). SNF1-related protein kinase 1: the many-faced signaling hub regulating developmental plasticity in plants. *J. Exp. Bot.* 72, 6042–6065. doi: 10.1093/JXB/ERAB079

Jansen, R., Greenbaum, D., and Gerstein, M. (2002). Relating whole-genome expression data with protein-protein interactions. *Genome Res.* 12, 37–46. doi: 10.1101/GR.205602

Javed, T., and Gao, S. J. (2023). WRKY transcription factors in plant defense. *Trends Genet.* 39, 787–801. doi: 10.1016/J.TIG.2023.07.001

Jayawardhane, K. N., Singer, S. D., Weselake, R. J., and Chen, G. (2018). Plant sn-glycerol-3-phosphate acyltransferases: biocatalysts involved in the biosynthesis of intracellular and extracellular lipids. *Lipids* 53, 469–480. doi: 10.1002/LIPD.12049

Jekat, S. B., Ernst, A. M., von Bohl, A., Zielonka, S., Twyman, R. M., Noll, G. A., et al. (2013). P-proteins in Arabidopsis are heteromeric structures involved in rapid sieve tube sealing. *Front. Plant Sci.* 4. doi: 10.3389/FPLS.2013.00225

Johnsson, N., and Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions *in vivo*. *Proc. Natl. Acad. Sci. U S A.* 91, 10340–10344. doi: 10.1073/pnas.91.22.10340

Kasapoğlu, A. G., Muslu, S., Aygören, A. S., Öner, B. M., Güneş, E., İlhan, E., et al. (2024). Genome-wide characterization of the GPAT gene family in bean (Phaseolus vulgaris L.) and expression analysis under abiotic stress and melatonin. *Genet. Resour. Crop Evol.* 71, 4549–4569. doi: 10.1007/S10722-024-01899-3

Kato, K., Horiba, A., Hayashi, H., Mizukami, H., and Terasaka, K. (2022). Characterization of triterpene saponin glycyrrhizin transport by Glycyrrhiza glabra. *Plants* 11, 1250. doi: 10.3390/PLANTS11091250/S1

Kawano, Y., Kaneko-Kawano, T., and Shimamoto, K. (2014). Rho family GTPase-dependent immunity in plants and animals. *Front. Plant Sci.* 5. doi: 10.3389/FPLS.2014.00522

Kawasaki, M., Nakayama, K., and Wakatsuki, S. (2005). Membrane recruitment of effector proteins by Arf and Rab GTPases. *Curr. Opin. Struct. Biol.* 15, 681–689. doi: 10.1016/J.SBI.2005.10.015

Keilhauer, E. C., Hein, M. Y., and Mann, M. (2015). Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol. Cell. Proteomics* 14, 120–135. doi: 10.1074/MCP.M114.041012

Kesselmeier, J., and Urban, B. (1983). Subcellular localization of saponins in green and etiolated leaves and green protoplasts of oat (Avena sativa L.). *Protoplasma* 114, 133–140. doi: 10.1007/BF01279877

Kim, E. Y., Park, K. Y., Seo, Y. S., and Kim, W. T. (2016). Arabidopsis small rubber particle protein homolog SRPs play dual roles as positive factors for tissue growth and development and in drought stress responses. *Plant Physiol.* 170, 2494–2510. doi: 10.1104/pp.16.00165

Kim, E. Y., Seo, Y. S., Lee, H., and Kim, W. T. (2010). Constitutive expression of CaSRP1, a hot pepper small rubber particle protein homolog, resulted in fast growth and improved drought tolerance in transgenic Arabidopsis plants. *Planta* 232, 71–83. doi: 10.1007/s00425-010-1149-2

Kim, I. J., Ryu, S. B., Kwak, Y. S., and Kang, H. (2004). A novel cDNA from Parthenium argentatum Gray enhances the rubber biosynthetic activity *in vitro**. *J. Exp. Bot.* 55, 377–385. doi: 10.1093/jxb/erh039

Kim, I. S., Shin, S. Y., Kim, S. H., and Yoon, H. S. (2012). Ectopic expression of sweet potato MuS1 increases acquired stress tolerance and fermentation yield in Saccharomyces cerevisiae. *J. Microbiol.* 50, 544–546. doi: 10.1007/s12275-012-2043-3

Kim, Y. N., Kim, J. S., Seo, S. G., Lee, Y., Baek, S. W., Kim, I. S., et al. (2011). Cadmium resistance in tobacco plants expressing the MuSI gene. *Plant Biotechnol. Rep.* 5, 323–329. doi: 10.1007/S11816-011-0186-Z/FIGURES/3

Kooijman, E. E., Chupin, V., de Kruijff, B., and Burger, K. N. J. (2003). Modulation of membrane curvature by phosphatidic acid and lysophosphatidic acid. *Traffic* 4, 162–174. doi: 10.1034/J.1600-0854.2003.00086.X

Kretzschmar, F. K., Doner, N. M., Krawczyk, H. E., Scholz, P., Schmitt, K., Valerius, O., et al. (2020). Identification of low-abundance lipid droplet proteins in seeds and seedlings. *Plant Physiol.* 182, 1326. doi: 10.1104/PP.19.01255

Kroh, G. E., and Pilon, M. (2020). Regulation of iron homeostasis and use in chloroplasts. *Int. J. Mol. Sci.* 21, 3395. doi: 10.3390/IJMS21093395

Kubo, A., Arai, Y., Nagashima, S., and Yoshikawa, T. (2004). Alteration of sugar donor specificities of plant glycosyltransferases by a single point mutation. *Arch. Biochem. Biophys.* 429, 198–203. doi: 10.1016/J.ABB.2004.06.021

Kumar, M., Kesawat, M. S., Du, X., Siddique, K. H. M., Kant, S., and Chung, S. M. (2024). In silico analysis and expression profiling reveal the presence of abiotic stress and developmental stage specific Aconitase genes in rice (Oryza sativa L.). *Plant Stress* 11, 100416. doi: 10.1016/J.STRESS.2024.100416

Kumar, S., and Trivedi, P. K. (2018). Glutathione S-transferases: role in combating abiotic stresses including arsenic detoxification in plants. *Front. Plant Sci.* 9. doi: 10.3389/FPLS.2018.00751

Kuroda, Y., Suzuki, N., and Kataoka, T. (1993). The effect of posttranslational modifications on the interaction of Ras2 with adenylyl cyclase. *Science* 259, 683–686. doi: 10.1126/SCIENCE.8430318

Labrou, N. E., Papageorgiou, A. C., Pavli, O., and Flemetakis, E. (2015). Plant GSTome: structure and functional role in xenome network and plant stress response. *Curr. Opin. Biotechnol.* 32, 186–194. doi: 10.1016/J.COPBIO.2014.12.024

Laibach, N., Hillebrand, A., Twyman, R. M., Prüfer, D., and Schulze Gronover, C. (2015). Identification of a *Taraxacum brevicorniculatum* rubber elongation factor protein that is localized on rubber particles and promotes rubber biosynthesis. *Plant J.* 82, 609–620. doi: 10.1111/tpj.12836

Laibach, N., Schmidl, S., Müller, B., Bergmann, M., Prüfer, D., and Schulze Gronover, C. (2018). Small rubber particle proteins from *Taraxacum brevicorniculatum* promote stress tolerance and influence the size and distribution of lipid droplets and artificial poly(*cis* -1,4-isoprene) bodies. *Plant J.* 93, 1045–1061. doi: 10.1111/tpj.13829

Lan, P., Li, W., Wang, H., and Ma, W. (2010). Characterization, sub-cellular localization and expression profiling of the isoprenylcysteine methylesterase gene family in Arabidopsis thaliana. *BMC Plant Biol.* 10, 212–212. doi: 10.1186/1471-2229-10-212

Lassowskat, I., Hartl, M., Hosp, F., Boersema, P. J., Mann, M., and Finkemeier, I. (2017). Dimethyl-labeling-based quantification of the lysine acetylome and proteome of plants. *Methods Mol. Biol. (Clifton N.J.)* 1653, 65–81. doi: 10.1007/978-1-4939-7225-8_5

Lazar, C., and Burger, T. (2022). imputeLCMD: A Collection of Methods for Left-Censored Missing Data Imputation *(R package version 2.1)*. Available online at: https://cran.r-project.org/package=imputeLCMD.

Leivar, P., González, V. M., Castel, S., Trelease, R. N., López-Iglesias, C., Arró, M., et al. (2005). Subcellular localization of Arabidopsis 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Plant Physiol.* 137, 57–69. doi: 10.1104/PP.104.050245

Liebthal, M., Maynard, D., and Dietz, K. J. (2018). Peroxiredoxins and redox signaling in plants. *Antioxid. Redox Signaling* 28, 609–624. doi: 10.1089/ARS.2017.7164/ASSET/IMAGES/LARGE/FIGURE7.JPEG

Lin, T., Xu, X., Du, H., Fan, X., Chen, Q., Hai, C., et al. (2022). Extensive sequence divergence between the reference genomes of Taraxacum kok-saghyz and Taraxacum mongolicum. *Sci. China Life Sci.* 65, 515–528. doi: 10.1007/S11427-021-2033-2/METRICS

Lin, T., Xu, X., Ruan, J., Liu, S., Wu, S., Shao, X., et al. (2018). Genome analysis of Taraxacum kok-saghyz Rodin provides new insights into rubber biosynthesis. *Natl. Sci. Rev.* 5, 78–87. doi: 10.1093/nsr/nwx101

Liu, D., Liu, Y., Rao, J., Wang, G., Li, H., Ge, F., et al. (2013). Overexpression of the glutathione S-transferase gene from Pyrus pyrifolia fruit improves tolerance to abiotic stress in transgenic tobacco plants. *Mol. Biol.* 47, 515–523. doi: 10.1134/S0026893313040109

Liu, M., Qiu, W., He, X., Zheng, L., Song, X., Han, X., et al. (2016). Functional characterization of a gene in Sedum alfredii hance resembling rubber elongation factor endowed with functions associated with cadmium tolerance. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00965

Liu, Y., Ji, X., Nie, X., Qu, M., Zheng, L., Tan, Z., et al. (2015). Arabidopsis AtbHLH112 regulates the expression of genes involved in abiotic stress tolerance by binding to their E-box and GCG-box motifs. *New Phytol.* 207, 692–709. doi: 10.1111/NPH.13387

Lo Cicero, L., Catara, V., Strano, C. P., Bella, P., Madesis, P., and Lo Piero, A. R. (2017). Over-expression of CsGSTU promotes tolerance to the herbicide alachlor and resistance to Pseudomonas syringe pv. tabaci in transgenic tobacco. *Biol. Plant.* 61, 169–177. doi: 10.1007/S10535-016-0659-6/METRICS

Louveau, T., and Osbourn, A. (2019). The sweet side of plant-specialized metabolism. *Cold Spring Harbor Perspect. Biol.* 11, a034744. doi: 10.1101/CSHPERSPECT.A034744

Manevich, Y., Feinstein, S. I., and Fisher, A. B. (2004). Activation of the antioxidant enzyme 1-CYS peroxiredoxin requires glutathionylation mediated by heterodimerization with πGST. *Proc. Natl. Acad. Sci. U S A.* 101, 3780–3785. doi: 10.1073/PNAS.0400181101

Marrs, K. A. (1996). The functions and regulation of glutathione s-transferases in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 127–158. doi: 10.1146/ANNUREV.ARPLANT.47.1.127

Mason, H. S., DeWald, D. B., and Mullet, J. E. (1993). Identification of a methyl jasmonate-responsive domain in the soybean vspB promoter. *Plant Cell* 5, 241–251. doi: 10.1105/TPC.5.3.241

Mazel, A., Leshem, Y., Tiwari, B. S., and Levine, A. (2004). Induction of salt and osmotic stress tolerance by overexpression of an intracellular vesicle trafficking protein AtRab7 (AtRabG3e). *Plant Physiol.* 134, 118–128. doi: 10.1104/PP.103.025379

McAssey, E. V., Gudger, E. G., Zuellig, M. P., and Burke, J. M. (2016). Population genetics of the rubber-producing Russian dandelion (Taraxacum kok-saghyz). *PloS One* 11, e0146417. doi: 10.1371/JOURNAL.PONE.0146417

McCaskill, D., and Croteau, R. (1998). Some caveats for bioengineering terpenoid metabolism in plants. *Trends Biotechnol.* 16, 349–355. doi: 10.1016/S0167-7799(98)01231-1

Menkens, A. E., Schindler, U., and Cashmore, A. R. (1995). The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. *Trends Biochem. Sci.* 20, 506–510. doi: 10.1016/S0968-0004(00)89118-5

Milla, P., Viola, F., Oliaro-Bosso, S., Rocco, F., Cattel, L., Joubert, B. M., et al. (2003). Subcellular localization of oxidosqualene cyclases from Arabidopsis thaliana, Trypanosoma cruzi, and Pneumocystis carinii expressed in yeast. *Lipids* 37, 1171–1176. doi: 10.1007/S11745-002-1017-9/METRICS

Mishra, M. K., Singh, G., Tiwari, S., Singh, R., Kumari, N., and Misra, P. (2015). Characterization of Arabidopsis sterol glycosyltransferase TTG15/UGT80B1 role during freeze and heat stress. *Plant Signal. Behav.* 10, e1075682. doi: 10.1080/15592324.2015.1075682

Moeder, W., Del Pozo, O., Navarre, D. A., Martin, G. B., and Klessig, D. F. (2007). Aconitase plays a role in regulating resistance to oxidative stress and cell death in Arabidopsis and Nicotiana benthamiana. *Plant Mol. Biol.* 63, 273–287. doi: 10.1007/S11103-006-9087-X/TABLES/1

Mofidi, S. S. H., Naghavi, M. R., Sabokdast, M., Jariani, P., Zargar, M., and Cornish, K. (2024). Effect of drought stress on natural rubber biosynthesis and quality in Taraxacum kok-saghyz roots. *PloS One* 19, e0295694. doi: 10.1371/JOURNAL.PONE.0295694

Moir, G. F. J. (1959). Ultracentrifugation and staining of hevea latex. *Nature* 184, 1626–1628. doi: 10.1038/1841626a0

Moncrieffe, M. C., Fernandez, M. J., Spiteller, D., Matsumura, H., Gay, N. J., Luisi, B. F., et al. (2012). Structure of the glycosyltransferase EryCIII in complex with its activating P450 homologue EryCII. *J. Mol. Biol.* 415, 92–101. doi: 10.1016/J.JMB.2011.10.036

Montag, K., Ivanov, R., and Bauer, P. (2023). Role of SEC14-like phosphatidylinositol transfer proteins in membrane identity and dynamics. *Front. Plant Sci.* 14. doi: 10.3389/FPLS.2023.1181031

Müller, B., Noll, G. A., Ernst, A. M., Rüping, B., Groscurth, S., Twyman, R. M., et al. (2010). Recombinant artificial forisomes provide ample quantities of smart biomaterials for use in technical devices. *Appl. Microbiol. Biotechnol.* 88, 689–698. doi: 10.1007/s00253-010-2771-4

Murphy, D. J. (2011). The dynamic roles of intracellular lipid droplets: from archaea to mammals. *Protoplasma* 249, 541–585. doi: 10.1007/S00709-011-0329-7

Mylona, P., Owatworakit, A., Papadopoulou, K., Jenner, H., Qin, B., Findlay, K., et al. (2008). Sad3 and Sad4 are required for saponin biosynthesis and root development in oat. *Plant Cell* 20, 201–212. doi: 10.1105/TPC.107.056531

Nagao, R. T., Goekjian, V. H., Hong, J. C., and Key, J. L. (1993). Identification of protein-binding DNA sequences in an auxin-regulated gene of soybean. *Plant Mol. Biol.* 21, 1147–1162. doi: 10.1007/BF00023610

Nianiou-Obeidat, I., Madesis, P., Kissoudis, C., Voulgari, G., Chronopoulou, E., Tsaftaris, A., et al. (2017). Plant glutathione transferase-mediated stress tolerance: functions and biotechnological applications. *Plant Cell Rep.* 36, 791–805. doi: 10.1007/S00299-017-2139-7

Nielsen, E. (2020). The small GTPase superfamily in plants: A conserved regulatory module with novel functions. *Annu. Rev. Plant Biol.* 71, 247–272. doi: 10.1146/ANNUREV-ARPLANT-112619-025827/1

Nielsen, E., Cheung, A. Y., and Ueda, T. (2008). The regulatory RAB and ARF GTPases for vesicular trafficking. *Plant Physiol.* 147, 1516–1526. doi: 10.1104/PP.108.121798

Niephaus, E., Müller, B., van Deenen, N., Lassowskat, I., Bonin, M., Finkemeier, I., et al. (2019). Uncovering mechanisms of rubber biosynthesis in Taraxacum koksaghyz – role of cis-prenyltransferase-like 1 protein. *Plant J.* 100, 591–609. doi: 10.1111/tpj.14471

Niu, M., Xiong, Y., Yan, H., Zhang, X., Li, Y., da Silva, J. A. T., et al. (2021). Cloning and expression analysis of mevalonate kinase and phosphomevalonate kinase genes associated with the MVA pathway in Santalum album. *Sci. Rep.* 11, 1–13. doi: 10.1038/s41598-021-96511-4

Nowicki, M., Zhao, Y., Boggess, S. L., Fluess, H., Payá-Milans, M., Staton, M. E., et al. (2019). Taraxacum kok-saghyz (rubber dandelion) genomic microsatellite loci reveal modest genetic diversity and cross-amplify broadly to related species. *Sci. Rep.* 9, 1–17. doi: 10.1038/s41598-019-38532-8

Oh, S. K., Kang, H., Shin, D. H., Yang, J., Chow, K.-S., Yeang, H. Y., et al. (1999). Isolation, Characterization, and Functional Analysis of a Novel cDNA Clone Encoding a Small Rubber Particle Protein from Hevea brasiliensis. *J. Biol. Chem.* 274, 17132–17138. doi: 10.1074/JBC.274.24.17132

Ohkama-Ohtsu, N., Sasaki-Sekimoto, Y., Oikawa, A., Jikumaru, Y., Shinoda, S., Inoue, E., et al. (2011). 12-oxo-phytodienoic acid–glutathione conjugate is transported into the vacuole in Arabidopsis. *Plant Cell Physiol.* 52, 205–209. doi: 10.1093/PCP/PCQ181

Padilla-Gonzalez, G. F., dos Santos, F. A., and Da Costa, F. B. (2016). Sesquiterpene lactones: more than protective plant compounds with high toxicity. *Crit. Rev. Plant Sci.* 35, 18–37. doi: 10.1080/07352689.2016.1145956

Panagabko, C., Morley, S., Hernandez, M., Cassolato, P., Gordon, H., Parsons, R., et al. (2003). Ligand specificity in the CRAL-TRIO protein family. *Biochemistry* 42, 6467–6474. doi: 10.1021/BI034086V

Panara, F., Lopez, L., Daddiego, L., Fantini, E., Facella, P., and Perrotta, G. (2018). Comparative transcriptomics between high and low rubber producing Taraxacum koksaghyz R. plants. *BMC Genomics* 19, 1–14. doi: 10.1186/S12864-018-5287-4

Pascual, J., Rahikainen, M., Angeleri, M., Alegre, S., Gossens, R., Shapiguzov, A., et al. (2021). ACONITASE 3 is part of theANAC017 transcription factor-dependent mitochondrial dysfunction response. *Plant Physiol.* 186, 1859. doi: 10.1093/PLPHYS/KIAB225

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., et al. (2023). InterPro in 2022. *Nucleic Acids Res.* 51, D418–D427. doi: 10.1093/NAR/GKAC993

Post, J., van Deenen, N., Fricke, J., Kowalski, N., Wurbs, D., Schaller, H., et al. (2012). Laticifer-specific cis-prenyltransferase silencing affects the rubber, triterpene, and inulin content of Taraxacum brevicorniculatum. *Plant Physiol.* 158, 1406–1417. doi: 10.1104/pp.111.187880

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/SUPPINFO

Puskas, J. E., Gautriaud, E., Deffieux, A., and Kennedy, J. P. (2006). Natural rubber biosynthesis—A living carbocationic polymerization? *Prog. Polymer Sci.* 31, 533–548. doi: 10.1016/J.PROGPOLYMSCI.2006.05.002

Pütter, K. M. (2017). *Isoprenoid pathway engineering in the rubber-producing genus Taraxacum* (Germany: Westphalian Wilhelms-University Münster).

Pütter, K. M., van Deenen, N., Müller, B., Fuchs, L., Vorwerk, K., Unland, K., et al. (2019). The enzymes OSC1 and CYP716A263 produce a high variety of triterpenoids in the latex of Taraxacum koksaghyz. *Sci. Rep.* 9, 1–13. doi: 10.1038/s41598-019-42381-w

Ramilowski, J. A., Sawai, S., Seki, H., Mochida, K., Yoshida, T., Sakurai, T., et al. (2013). Glycyrrhiza uralensis transcriptome landscape and study of phytochemicals. *Plant Cell Physiol.* 54, 697–710. doi: 10.1093/pcp/pct057

Ravet, K., Touraine, B., Boucherez, J., Briat, J. F., Gaymard, F., and Cellier, F. (2009). Ferritins control interaction between iron homeostasis and oxidative stress in Arabidopsis. *Plant J.* 57, 400–412. doi: 10.1111/J.1365-313X.2008.03698.X

Reichel, C., and Johnsson, N. (2005). The split-ubiquitin sensor: measuring interactions and conformational alterations of proteins *in vivo*. *Methods Enzymol.* 399, 757–776. doi: 10.1016/S0076-6879(05)99050-2

Robertlee, J., Kobayashi, K., Suzuki, M., and Muranaka, T. (2017). AKIN10, a representative Arabidopsis SNF1-related protein kinase 1 (SnRK1), phosphorylates and downregulates plant HMG-CoA reductase. *FEBS Lett.* 591, 1159–1166. doi: 10.1002/1873-3468.12618

Rodriguez-Furlan, C., Borna, R., and Betz, O. (2023). RAB7 GTPases as coordinators of plant endomembrane traffic. *Front. Plant Sci.* 14. doi: 10.3389/FPLS.2023.1240973

Rogowska, A., and Szakiel, A. (2020). The role of sterols in plant response to abiotic stress. *Phytochem. Rev.* 19, 1525–1538. doi: 10.1007/S11101-020-09708-2

Roxas, V. P., Lodhi, S. A., Garrett, D. K., Mahan, J. R., and Allen, R. D. (2000). Stress tolerance in transgenic tobacco seedlings that overexpress glutathione S-transferase/glutathione peroxidase. *Plant Cell Physiol.* 41, 1229–1234. doi: 10.1093/PCP/PCD051

Rüdiger, H., and Gabius, H. J. (2002). Plant lectins: Occurrence, biochemistry, functions and applications. *Glycoconjugate J.* 18, 589–613. doi: 10.1023/A:1020687518999

Salehi, M., Cornish, K., Bahmankar, M., and Naghavi, M. R. (2021). Natural rubber-producing sources, systems, and perspectives for breeding and biotechnology studies of Taraxacum kok-saghyz. *Ind. Crops Prod.* 170, 113667. doi: 10.1016/J.INDCROP.2021.113667

Sando, T., Takeno, S., Watanabe, N., Okumoto, H., Kuzuyama, T., Yamashita, A., et al. (2008). Cloning and characterization of the 2- C -methyl- D -erythritol 4-phosphate (MEP) pathway genes of a natural-rubber producing plant, *Hevea brasiliensis*. *Biosci. Biotechnol. Biochem.* 72, 2903–2917. doi: 10.1271/bbb.80387

Santos Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M., and Lepiniec, L. (2005). LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in *Arabidopsis* leaves. *FEBS Lett.* 579, 4666–4670. doi: 10.1016/j.febslet.2005.07.037

Sapperstein, S., Berkower, C., and Michaelis, S. (1994). Nucleotide sequence of the yeast STE14 gene, which encodes farnesylcysteine carboxyl methyltransferase, and demonstration of its essential role in a-factor export. *Mol. Cell. Biol.* 14, 1438–1449. doi: 10.1128/MCB.14.2.1438

Schmidt, T., Hillebrand, A., Wurbs, D., Wahler, D., Lenders, M., Schulze Gronover, C., et al. (2010a). Molecular cloning and characterization of rubber biosynthetic genes from Taraxacum koksaghyz. *Plant Mol. Biol. Rep.* 28, 277–284. doi: 10.1007/s11105-009-0145-9

Schmidt, T., Lenders, M., Hillebrand, A., van Deenen, N., Munt, O., Reichelt, R., et al. (2010b). Characterization of rubber particles and rubber chain elongation in Taraxacum koksaghyz. *BMC Biochem.* 11. doi: 10.1186/1471-2091-11-11

Seitz, S. B., Weisheit, W., and Mittag, M. (2010). Multiple roles and interaction factors of an E-Box element in Chlamydomonas reinhardtii. *Plant Physiol.* 152, 2243. doi: 10.1104/PP.109.149195

Seo, S. G., Kim, J. S., Yang, Y. S., Jun, B. K., Kang, S. W., Lee, G. P., et al. (2010). Cloning and characterization of the new multiple stress responsible gene i (MuSI) from sweet potato. *Genes Genomics* 32, 544–552. doi: 10.1007/s13258-010-0093-7

Sevilla, F., Camejo, D., Ortiz-Espín, A., Calderón, A., Lázaro, J. J., and Jiménez, A. (2015). The thioredoxin/peroxiredoxin/sulfiredoxin system: current overview on its redox function in plants and regulation by reactive oxygen and nitrogen species. *J. Exp. Bot.* 66, 2945–2955. doi: 10.1093/JXB/ERV146

Shahmuradov, I. A., and Solovyev, V. V. (2015). Nsite, NsiteH and NsiteM computer tools for studying transcription regulatory elements. *Bioinformatics* 31, 3544. doi: 10.1093/BIOINFORMATICS/BTV404

Shaikhali, J., Norén, L., De Dios Barajas-López, J., Srivastava, V., König, J., Sauer, U. H., et al. (2012). Redox-mediated mechanisms regulate DNA binding activity of the G-group of basic region leucine zipper (bZIP) transcription factors in Arabidopsis. *J. Biol. Chem.* 287, 27510–27525. doi: 10.1074/jbc.M112.361394

Shockey, J., Regmi, A., Cotton, K., Adhikari, N., Browse, J., and Bates, P. D. (2015). Identification of Arabidopsis GPAT9 (At5g60620) as an essential gene involved in triacylglycerol biosynthesis. *Plant Physiol.* 170, 163–179. doi: 10.1104/PP.15.01563

Sibéril, Y., Doireau, P., and Gantet, P. (2001). Plant bZIP G-box binding factors. *Eur. J. Biochem.* 268, 5655–5666. doi: 10.1046/J.0014-2956.2001.02552.X

Sikorski, R. S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces ceratisiae. *Genetics* 122, 19–27. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1203683/.

Simkin, A. J., Guirimand, G., Papon, N., Courdavault, V., Thabet, I., Ginis, O., et al. (2011). Peroxisomal localization of the final steps of the mevalonic acid pathway in planta. *Planta* 234, 903–914. doi: 10.1007/S00425-011-1444-6/FIGURES/5

Slocombe, S. P., Cornah, J., Pinfield-Wells, H., Soady, K., Zhang, Q., Gilday, A., et al. (2009). Oil accumulation in leaves directed by modification of fatty acid breakdown and lipid synthesis pathways. *Plant Biotechnol. J.* 7, 694–703. doi: 10.1111/J.1467-7652.2009.00435.X

Spang, A. (2008). Membrane traffic in the secretory pathway: The life cycle of a transport vesicle. *Cell. Mol. Life Sci.* 65, 2781–2789. doi: 10.1007/S00018-008-8349-Y/METRICS

Springer, S., Spang, A., and Schekman, R. (1999). A primer on vesicle budding. *Cell* 97, 145–148. doi: 10.1016/S0092-8674(00)80722-9

Srere, P. A. (1972). "Is there an organization of Krebs cycle enzymes in the mitochondrial matrix?" In *Energy metabolism and the regulation of metabolic processes in mitochondria*. M. A. Mehlman and R. W. Hanson (Eds.), (Academic Press) 79–91. doi: 10.1016/B978-0-12-487850-1.50011-7

Srere, P. A. (1987). Complexes of sequential metabolic enzymes. *Annu. Rev. Biochem.* 56, 89–124. doi: 10.1146/ANNUREV.BI.56.070187.000513

Stucky, D. F., Arpin, J. C., and Schrick, K. (2014). Functional diversification of two UGT80 enzymes required for steryl glucoside synthesis in Arabidopsis. *J. Exp. Bot.* 66, 189–201. doi: 10.1093/JXB/ERU410

Sugden, C., Donaghy, P. G., Halford, N. G., and Hardie, D. G. (1999). Two SNF1-related protein kinases from spinach leaf phosphorylate and inactivate 3-hydroxy-3-methylglutaryl-coenzyme a reductase, nitrate reductase, and sucrose phosphate synthase *in vitro*. *Plant Physiol.* 120, 257–274. doi: 10.1104/PP.120.1.257

Sui, N., Li, M., Liu, X. Y., Wang, N., Fang, W., and Meng, Q. W. (2007a). Response of xanthophyll cycle and chloroplastic antioxidant enzymes to chilling stress in tomato over-expressing glycerol-3-phosphate acyltransferase gene. *Photosynthetica* 45, 447–454. doi: 10.1007/S11099-007-0074-5

Sui, N., Li, M., Zhao, S. J., Li, F., Liang, H., and Meng, Q. W. (2007b). Overexpression of glycerol-3-phosphate acyltransferase gene improves chilling tolerance in tomato. *Planta* 226, 1097–1108. doi: 10.1007/S00425-007-0554-7/FIGURES/6

Sun, L. P., Ouyang, L. L., Bao, H., Liu, J. G., Sun, Z., and Zhou, Z. G. (2021). Comparison between two isoforms of glycerol-3-phosphate acyltransferase in microalga Myrmecia incisa: Subcellular localization and role in triacylglycerol synthesis. *Algal Res.* 54, 102172. doi: 10.1016/J.ALGAL.2020.102172

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

Tanaka, Y., Aik-Hwee, E., Ohya, N., Nishiyama, N., Tangpakdee, J., Kawahara, S., et al. (1996). Initiation of rubber biosynthesis in Hevea brasiliensis: characterization of initiating species by structural analysis. *Phytochemistry* 41, 1501–1505. doi: 10.1016/0031-9422(95)00817-9

Team, R. C. (2021). R: A language and environment for statistical computing (RStudio 2022.12.0). *R Foundation for Statistical Computing*. Available at: https://www.r-project.org/ Under the given URL one can download R-.

Testerink, C., and Munnik, T. (2011). Molecular, cellular, and physiological responses to phosphatidic acid formation in plants. *J. Exp. Bot.* 62, 2349–2361. doi: 10.1093/JXB/ERR079

Thomas, P. D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L. P., and Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* 31, 8–22. doi: 10.1002/pro.4218

Tiwari, V., Patel, M. K., Chaturvedi, A. K., Mishra, A., and Jha, B. (2016). Functional Characterization of the Tau Class Glutathione-S-Transferases Gene (SbGSTU) Promoter of Salicornia brachiata under Salinity and Osmotic Stress. *PloS One* 11, e0148494. doi: 10.1371/JOURNAL.PONE.0148494

Toledo-Ortiz, G., Johansson, H., Lee, K. P., Bou-Torrent, J., Stewart, K., Steel, G., et al. (2014). The HY5-PIF regulatory module coordinates light and temperature control of photosynthetic gene transcription. *PloS Genet.* 10, e1004416. doi: 10.1371/JOURNAL.PGEN.1004416

Tripathy, M. K., Tiwari, B. S., Reddy, M. K., Deswal, R., and Sopory, S. K. (2017). Ectopic expression of PgRab7 in rice plants (Oryza sativa L.) results in differential tolerance at the vegetative and seed setting stage during salinity and drought stress. *Protoplasma* 254, 109–124. doi: 10.1007/S00709-015-0914-2/FIGURES/11

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* 11, 2301–2319. doi: 10.1038/nprot.2016.136

Unland, K., Pütter, K. M., Vorwerk, K., van Deenen, N., Twyman, R. M., Prüfer, D., et al. (2018). Functional characterization of squalene synthase and squalene epoxidase in Taraxacum koksaghyz. *Plant Direct* 2, 1–15. doi: 10.1002/pld3.63

Urban, B., Laudenbach, U., and Kesselmeier, J. (1983). Saponin distribution in the etiolated leaf tissue and subcellular localization of steroidal saponins in etiolated protoplasts of oat (Avena sativa L.). *Protoplasma* 118, 121–123. doi: 10.1007/BF01293068/METRICS

Van Deenen, N., Bachmann, A. L., Schmidt, T., Schaller, H., Sand, J., Prü fer, D., et al. (2012). Molecular cloning of mevalonate pathway genes from Taraxacum brevicorniculatum and functional characterization of the key enzyme 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Mol. Biol. Rep.* 39, 4337–4339. doi: 10.1007/S11033-011-1221-4/TABLES/3

Vogt, T., and Jones, P. (2000). Glycosyltransferases in plant-natural product synthesis: Characterization of a supergene family. *Trends Plant Sci.* 5, 380–386. doi: 10.1016/S1360-1385(00)01720-9

Volkman, B. F., Lipson, D., Wemmer, D. E., and Kern, D. (2001). Two-state allosteric behavior in a single-domain signaling protein. *Science* 291, 2429–2433. doi: 10.1126/SCIENCE.291.5512.2429

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., et al. (2002). Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* 417, 399–403. doi: 10.1038/nature750

Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/JHERED/93.1.77

Vranová, E., Coman, D., and Gruissem, W. (2012). Structure and dynamics of the isoprenoid pathway network. *Mol. Plant* 5, 318–333. doi: 10.1093/mp/sss015

Wagner, U., Edwards, R., Dixon, D. P., and Mauch, F. (2002). Probing the diversity of the Arabidopsis glutathione S-transferase gene family. *Plant Mol. Biol.* 49, 515–532. doi: 10.1023/A:1015557300450/METRICS

Wahler, D., Colby, T., Kowalski, N. A., Harzen, A., Wotzka, S. Y., Hillebrand, A., et al. (2012). Proteomic analysis of latex from the rubber-producing plant Taraxacum brevicorniculatum. *Proteomics* 12, 901–905. doi: 10.1002/pmic.201000778

Wang, J., Song, J., Qi, H., Zhang, H., Wang, L., Zhang, H., et al. (2023). Overexpression of 2-Cys Peroxiredoxin alleviates the NaHCO3 stress-induced photoinhibition and reactive oxygen species damage of tobacco. *Plant Physiol. Biochem.* 201, 107876. doi: 10.1016/J.PLAPHY.2023.107876

Wani, S. H., Anand, S., Singh, B., Bohra, A., and Joshi, R. (2021). WRKY transcription factors and plant defense responses: latest discoveries and future prospects. *Plant Cell Rep.* 40, 1071–1085. doi: 10.1007/S00299-021-02691-8

Wei, R., Wang, J., Su, M., Jia, E., Chen, S., Chen, T., et al. (2018). Missing value imputation approach for mass spectrometry-based metabolomics data. *Sci. Rep.* 8, 663. doi: 10.1038/S41598-017-19120-0

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag).

Wieghaus, A., Roelfs, K. U., Twyman, R. M., Prüfer, D., and Schulze Gronover, C. (2022). Comparative transcriptome analysis in Taraxacum koksaghyz to identify genes that determine root volume and root length. *Front. Genet.* 12. doi: 10.3389/FGENE.2021.784883/FULL

Wilfling, F., Haas, J. T., Walther, T. C., and Farese, R. V. Jr. (2014). Lipid droplet biogenesis. *Curr. Opin. Cell Biol.* 29, 39–45. doi: 10.1016/j.ceb.2014.03.008

Wilfling, F., Wang, H., Haas, J. T., Krahmer, N., Gould, T. J., Uchida, A., et al. (2013). Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocalizing from the ER to lipid droplets. *Dev. Cell* 24, 384–399. doi: 10.1016/j.devcel.2013.01.013

Williams, M. E., Foster, R., and Chua, N. H. (1992). Sequences flanking the hexametric G-box core CACGTG affect the specificity of protein binding. *Plant Cell* 4, 485–496. doi: 10.1105/TPC.4.4.485

Wititsuwannakul, R., Rukseree, K., Kanokwiroon, K., and Wititsuwannakul, D. (2008). A rubber particle protein specific for Hevea latex lectin binding involved in latex coagulation. *Phytochemistry* 69, 1111–1118. doi: 10.1016/J.PHYTOCHEM.2007.12.007

Wood, D. F., and Cornish, K. (2000). Microstructure of purified rubber particles. *Int. J. Plant Sci.* 161, 435–445. doi: 10.1086/314269

Wu, Y., Dong, G., Luo, F., Xie, H., Li, X., and Yan, J. (2024). TkJAZs-TkMYC2-TkSRPP/REF regulates the biosynthesis of natural rubber in Taraxacum kok-saghyz. *Plants* 13, 2034. doi: 10.3390/plants13152034

Xiao, G., Zhao, M., Liu, Q., Zhou, J., Cheng, Z., Wang, Q., et al. (2023). TaBAS1 encoding a typical 2-Cys peroxiredoxin enhances salt tolerance in wheat. *Front. Plant Sci.* 14. doi: 10.3389/FPLS.2023.1152375/BIBTEX

Xie, W., McMahan, C. M., DeGraw, A. J., Distefano, M. D., Cornish, K., Whalen, M. C., et al. (2008). Initiation of rubber biosynthesis: *In vitro* comparisons of benzophenone-modified diphosphate analogues in three rubber-producing species. *Phytochemistry* 69, 2539–2545. doi: 10.1016/J.PHYTOCHEM.2008.07.011

Xu, T., Jia, Z., Wu, L., Chen, Y., Luo, Y., Jia, D., et al. (2017). Effect of acetone extract from natural rubber on the structure and interface interaction in NR/CB composites. *RSC Adv.* 7, 26458–26467. doi: 10.1039/C7RA03354K

Yamashita, S., Yamaguchi, H., Waki, T., Aoki, Y., Mizuno, M., Yanbe, F., et al. (2016). Identification and reconstitution of the rubber biosynthetic machinery on rubber particles from Hevea brasiliensis. *ELife* 5, e19022. doi: 10.7554/eLife.19022

Yang, Q., Grimmig, B., and Matern, U. (1998). Anthranilate N-hydroxycinnamoyl/benzoyltransferase gene from carnation: Rapid elicitation of transcription and promoter analysis. *Plant Mol. Biol.* 38, 1201–1214. doi: 10.1023/A:1006003731919/METRICS

Yang, W., Pollard, M., Li-Beisson, Y., Beisson, F., Feig, M., and Ohlrogge, J. (2010). A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. *Proc. Natl. Acad. Sci. United States America* 107, 12040–12045. doi: 10.1073/PNAS.0914149107/SUPPL_FILE/SAPP.PDF

Zang, X., Geng, X., Wang, F., Liu, Z., Zhang, L., Zhao, Y., et al. (2017). Overexpression of wheat ferritin gene TaFER-5B enhances tolerance to heat stress and other abiotic stresses associated with the ROS scavenging. *BMC Plant Biol.* 17, 1–13. doi: 10.1186/S12870-016-0958-2/FIGURES/9

Zhai, Z., Liu, H., and Shanklin, J. (2017). Phosphorylation of WRINKLED1 by KIN10 results in its proteasomal degradation, providing a link between energy homeostasis and lipid biosynthesis. *Plant Cell* 29, 871. doi: 10.1105/TPC.17.00019

Zhang, F. L., and Casey, P. J. (1996). Protein prenylation: Molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* 65, 241–269. doi: 10.1146/ANNUREV.BI.65.070196.001325/CITE/REFWORKS

Zhang, N., McHale, L. K., and Finer, J. J. (2019). Changes to the core and flanking sequences of G-box elements lead to increases and decreases in gene expression in both native and synthetic soybean promoters. *Plant Biotechnol. J.* 17, 724. doi: 10.1111/PBI.13010

Zhang, Y., and Fernie, A. R. (2021). Metabolons, enzyme–enzyme assemblies that mediate substrate channeling, and their roles in plant metabolism. *Plant Commun.* 2, 100081. doi: 10.1016/J.XPLC.2020.100081

Zhou, Q., Yang, Y., and Yang, Z. (2019). Molecular dissection of cadmium-responsive transcriptome profile in a low-cadmium-accumulating cultivar of Brassica parachinensis. *Ecotoxicol. Environ. Saf.* 176, 85–94. doi: 10.1016/J.ECOENV.2019.03.077

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 February 2025
DOI 10.3389/fpls.2024.1425154

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Shouchuang Wang,
Hainan University, China
Carlos Henrique Meneses,
State University of Paraíba, Brazil
Ralph Dewey,
North Carolina State University, United States

*CORRESPONDENCE
Zhaoliang Geng
zhaolianggeng1104@126.com
Bin Cai
caibin_hcri@qq.com

†These authors have contributed equally to this work

RECEIVED 29 April 2024
ACCEPTED 07 October 2024
PUBLISHED 20 February 2025

CITATION
Guo D, Gao H, Yan T, Xia C, Lin B, Xiang X, Cai B and Geng Z (2025) Proteomic and metabolomic insights into the impact of topping treatment on cigar tobacco. *Front. Plant Sci.* 15:1425154. doi: 10.3389/fpls.2024.1425154

Proteomic and metabolomic insights into the impact of topping treatment on cigar tobacco

Dong Guo†, Huajun Gao†, Tongjing Yan, Changjian Xia, Beisen Lin, Xiaohua Xiang, Bin Cai* and Zhaoliang Geng*

Haikou Cigar Research Institute, Hainan Province Company, China National Tobacco Corporation, Haikou, China

Top removal is a widely utilized method in production process of tobacco, but little is known regarding the way it impacts protein and metabolic regulation. In this study, we investigated the underlying processes of alterations in cigar tobacco leaves with and without top removal, using a combined proteomic and metabolomic approach. The results revealed that: (1) Topping significantly affected superoxide anion (O_2^-) levels, superoxide dismutase (SOD) activity, and malondialdehyde (MDA) content, (2) In the cigar tobacco proteome, 385 differentially expressed proteins (DEPs) were identified, with 228 proteins upregulated and 156 downregulated. Key pathways enriched included flavonoid biosynthesis, porphyrin and chlorophyll metabolism, cysteine and methionine metabolism, and amino acid biosynthesis and metabolism. A network of 161 nodes interconnected by 102 significantly altered proteins was established, (3) In the cigar tobacco metabolome, 247 significantly different metabolites (DEMs) were identified, with 120 upregulated and 128 downregulated metabolites, mainly comprising lipids and lipid-like molecules, phenylpropanoids and polyketides, organic acids and derivatives, and organic heterocyclic compounds, (4) KEGG pathway enrichment revealed upregulation of proteins such as chalcone synthase (CHS), chalcone isomerase (CHI), naringenin 3-dioxygenase (F3H), and flavonoid 3'-monooxygenase (F3'H), along with metabolites like pinocembrin, kaempferol, trifolin, rutin, and quercetin, enhancing the pathways of 'flavonoid' and 'flavone and flavonol' biosynthesis. This study sheds light on the metabolic and proteomic responses of cigar tobacco after topping.

KEYWORDS

tobacco, top removal, flavonoid biosynthesis, proteomics, metabolomics

1 Introduction

Tobacco is an important economic crop cultivated over vast expanses (Chen et al., 2021). In China, it is categorized into four types based on drying techniques and agronomic traits: flue-cured, sun-cured, air-cured, and burley tobacco (Liu et al., 2021). Cigar tobacco, an air-cured variety, is one of the most commonly grown and traditional types of tobacco. A cigar consists of three main components: the wrapper, binder, and filler, all made from tobacco leaves (Zhang et al., 2023a). As a result, the quality and yield of tobacco leaves are exceedingly significant. In tobacco cultivation, topping—the removal of the plant's upper portion—is a key agronomic practice that plays a crucial role in enhancing both the quality and quantity of the leaves (Baldwin, 2001; Qin et al., 2020).

Topping in plants, particularly in tobacco, involves the removal of the upper flowering parts and young leaves and is an essential component of tobacco farming (Dai et al., 2022). As tobacco plants mature, they progress from the vegetative to the reproductive phase, with the apical meristem of the main stem converting into floral meristems. When the plant begins flowering, it diverts a significant amount of nutrients to the top, creating a reproductive growth zone. This process involves the movement of water and nutrients from the roots, along with the redistribution of resources from the lower and middle leaves to the upper parts (Shi et al., 2006). This nutrient redistribution becomes increasingly evident as tobacco flowers and seeds develop (Geuns et al., 2001). Nutrients generated in flue-cured tobacco leaves are primarily directed toward flowering, which limits leaf growth. This results in smaller, lighter upper leaves, diminishing the overall quality and yield (Chen et al., 2019). When carried out at the appropriate moment, it redirects the plant's energy toward the remaining leaves, improving their yield and quality (Yan et al., 2019; Lei et al., 2022). This technique shifts the plant's primary growth and nutrient distribution centers, altering the original source-sink relationship (Ogilvie and Kozumplik, 1980). Consequently, it influences several biological functions, including secondary metabolism and hormone regulation (Guo et al., 2023). Topping substantially elevates the levels of secondary metabolites like nicotine and aromatic amines (Shi et al., 2006; Banožić et al., 2020), and enhances the expression of genes linked to hormone metabolism and plant defense (Guo et al., 2011). Essentially a form of mechanical damage, topping mimics the effects of insect attacks or physical injuries, leading to an accumulation of reactive oxygen species (ROS), which cause oxidative stress (Gaquerel et al., 2009; Luo et al., 2016; Yan et al., 2019). This activates the plant's defense mechanisms, prompting the production of secondary metabolites and volatile compounds such as alkaloids and carotenoids. Jasmonic acid (JA) plays a central role in this defense response, regulating carbohydrate distribution and secondary metabolite formation (Katoh et al., 2007; Deboer et al., 2009; Ma et al., 2019), thereby increasing antimicrobial compounds. Therefore, topping is a crucial field management practice during tobacco plant maturation. However, research on cigar tobacco's proteomic and metabolomic responses to topping remains incomplete.

With the ongoing advancement and adoption of omics technologies, proteomics and metabolomics offer deeper insights into tobacco proteins and their metabolites (Wang et al., 2021). Proteomics, through protein quantification, links gene transcription to metabolic processes (Jin et al., 2015; Cao et al., 2024), and combining proteomics with transcriptomics improves the understanding of regulatory networks. Metabolomics, a vital element of systems biology, identifies a broad spectrum of endogenous metabolites (Wang et al., 2019; Yang et al., 2024). But little is known regarding the way it impacts protein and metabolic regulation, this study employs proteomic and metabolomic analyses to reveal the antioxidant response mechanisms and regulatory processes governing physiological changes in cigar tobacco after topping.

2 Materials and methods

2.1 Plant materials and treatment

The 'Haiyan103' variety of cigar tobacco, provided by the Haikou Cigar Research Institute in Haikou, China, was the subject of our investigation. Our goal was to evaluate how topping—removal of the top floral portion and upper young leaves—affects the growth of tobacco plants. We chose plants from Danzhou, one of the main tobacco-growing regions in the Hainan Province. The study included two groups: plants without topping (DDW1) and plants with topping (DDW2). Topping was performed 55 days after transplanting, and samples were collected 5 days later. The untopped plants served as the control group, with three replicates in each plot.

Plants were arranged in rows 1.2 meters apart, with 0.6 meters between individual plants. Each treatment received 180 kg of N, 270 kg of P_2O_5, and 360 kg of K_2O per hectare, and was managed according to standard cigar tobacco cultivation practices (Lei et al., 2014).

During sampling, nine plants were chosen at random from each group as a biological duplicate. On cigar tobacco plants, the eighth leaf (counted from base to tip, undamaged leaves with an area of approximately 48 × 26 cm) was chosen as the sample, and the leaf's central vein was removed. The leaves were then pooled, wrapped in foil, and quickly frozen in liquid nitrogen. For each biological replicate, we ensured sample uniformity by freezing and grinding the samples in liquid nitrogen.

2.2 Antioxidant index evaluation

To assess superoxide dismutase (SOD) activity, we prepared the enzyme solution according to the procedure outlined by Durak et al. (1993). SOD activity was quantified spectrophotometrically at 560 nm, based on its ability to inhibit the 50% reduction of nitro blue tetrazolium (NBT). The superoxide anion (O_2^-) levels were measured using the hydroxylamine oxidation method described by Zhuang et al. (2019). Malondialdehyde (MDA) levels were determined using the thiobarbituric acid (TBA) reactive substances method, as detailed by Draper et al. (1993).

2.3 Proteomic analysis

2.3.1 Protein isolation and digestion process

Protein isolation was carried out by grinding 0.5 g of tobacco leaf tissue into a powder in liquid nitrogen. The powder was then swiftly transferred to a pre-cooled centrifuge tube containing 800 µL of SDT lysis buffer (with 100 mM NaCl) and 1/100 volume of dithiothreitol (DTT). The mixture was shaken thoroughly and sonicated in an ice bath for 5 minutes to lyse the sample completely. It was then heated at 95°C for 8-15 minutes, cooled in an ice bath for 2 minutes, and centrifuged at 4°C at 12,000 g for 15 minutes. To the supernatant, 80 µL of iodoacetamide (IAM) solution was added and kept in the dark (in a closed drawer) for 1 hour. Next, 2 mL of pre-cooled acetone was added and the mixture was precipitated at -20°C for 2 hours, followed by centrifugation at 4°C at 12,000 g for 15 minutes, and the resultant precipitate was collected. The precipitate was then suspended and washed in 1 mL of pre-cooled acetone, collected, and air-dried. The resulting precipitate, which represents the total protein, was then dissolved in 1 mL of Dissolved Buffer (DB buffer).

Protein digestion was performed as outlined by Zhang et al. (2016). A total of 200 µL of DB protein dissolution buffer was added. Then, 2 µg of trypsin and 100 mM Triethylammonium bicarbonate (TEA) buffer were added, mixed thoroughly, and incubated at 37°C for 4 hours. Afterward, another 2 µg of trypsin and 200 mM $CaCl_2$ were added, followed by overnight incubation at 37°C for further enzymatic digestion. Formic acid was used to lower the pH below 3, and then the mixture was centrifuged at 12,000 g for 5 minutes at room temperature. The supernatant was slowly passed through a C18 desalting column, washed thrice with a solution of 0.1% formic acid and 3% acetonitrile, and eluted with 200 µL of elution solution (0.1% formic acid, 70% acetonitrile). The filtrate was then collected and lyophilized.

2.3.2 Proteomic analysis via UHPLC-MS/MS

For proteomic analysis, we employed an Easy-nLC™ 1200 UHPLC (Thermo Fisher, Germany) coupled with a Q Exploris™ HF-X mass spectrometer (Thermo Fisher, Germany). A total of 4 µg of each sample, combined with iRT reagent, was loaded onto a C18 Nano-Trap column. The gradient profile ranged from 5% to 95% acetonitrile in formic acid over 92 minutes, at a flow rate of 600 nL/min. Peptide analysis was performed using a Q Exactive™ HF-X mass spectrometer with specific settings for full scan range, resolution, AGC target, ion injection time, and fragmentation. The top 40 precursors were selected for MS/MS analysis, with DIA mode applied for broader peptide coverage.

2.3.3 Identification and quantification of proteins

We individually analyzed the spectra from each fraction against the *Nicotiana tabacum* protein database using Proteome Discoverer 2.2. The search parameters were carefully optimized for precision and identified proteins were required to satisfy strict criteria, including FDR and amino acid coverage.

2.3.4 Data analysis and statistical approaches

For functional annotation, we used Gene Ontology and InterPro analyses through InterProScan, comparing against an extensive protein database. Additionally, we utilized the COG and KEGG databases to further analyze protein families and pathways. Differential protein expression was examined with various tools, including Volcano plot and heat map analysis. Protein-protein interactions were predicted using the STRING-db server.

2.4 Metabolomics analysis

2.4.1 Metabolite extraction and untargeted metabolomic analysis

0.1 g of leaf tissue was ground in liquid nitrogen to form a fine powder, which was then reconstituted in pre-chilled 80% methanol. After vigorously shaking the mixture, it was placed on ice for 5 minutes and then centrifuged at 15,000 g for 20 minutes at 4°C. The clear supernatant was diluted to a 53% methanol concentration using UHPLC-MS/MS grade water. The diluted mixture was transferred to new Eppendorf tubes and centrifuged again under the same conditions. The resulting supernatant was then prepared for UHPLC-MS/MS analysis.

The samples were analyzed using a Vanquish UHPLC system (Thermo Fisher, Germany), paired with an Orbitrap Q Exactive™ HF mass spectrometer (Thermo Fisher, Germany). The samples were loaded onto a Hypersil Gold column (100 × 2.1 mm, 1.9 µm) (Thermo Fisher, USA) using a 12-minute linear gradient at a flow rate of 0.2 mL/min. For positive polarity mode, eluent A was 0.1% formic acid in water and eluent B was methanol. For negative polarity mode, eluent A was 5 mM ammonium acetate (pH 9) and eluent B was methanol. The solvent gradient was: 2% B, 1.5 minutes; 2-85% B, 3 minutes; 85-100% B, 10 minutes; 100-2% B, 10.1 minutes; 2% B, 12 minutes. The mass spectrometer parameters, including spray voltage, capillary temperature, gas flow rates, and S-lens RF level, were precisely calibrated. The MS scan range was 90 to 900 m/z, with fragmentation data acquisition in both polarity modes.

2.4.2 Processing and analysis of metabolomics data

To process the metabolomics data, we employed Compound Discoverer 3.3. The metabolite identification process involved comparing the acquired data against various databases, including mzCloud, mzVault, KEGG, and others, while adhering to a strict mass tolerance. For metabolite annotation, we relied on established databases such as KEGG, HMDB, and LIPID Maps, which provide a wealth of information for precise identification and classification.

2.4.3 Synthesis of proteomic and metabolomic data

The incorporation of the identified differentially expressed proteins (DEPs) and metabolites (DEMs) into the KEGG pathway maps enabled the visualization of changes in key metabolic pathways.

2.5 Statistical analysis

Data preparation was carried out with Microsoft Excel 2023, and all statistical analyses were conducted in SPSS V16.0 for Windows (SPSS, Chicago, Illinois, USA). A difference was deemed statistically significant if $P < 0.05$.

3 Results

3.1 Effect of topping on antioxidant enzymes

The antioxidant enzyme indices of tobacco leaves that underwent topping were compared to those that did not. As illustrated in Figure 1, topping had a significant effect on O_2^- content, SOD activity, and MDA content. Specifically, O_2^- levels in the DDW2 group were 22.95% higher than in the DDW1 group. In a similar vein, SOD activity in the DDW2 group was 19.30% higher than in the DDW1 group. Furthermore, MDA content in the DDW2 group was 34.10% higher, demonstrating a statistically significant difference.

3.2 Identification of proteins

An unsupervised, multi-factor principal component analysis (PCA) was performed to evaluate the protein profile data of cigar tobacco leaves subjected to topping and non-topping treatments. As depicted in Figure 1, the two groups of cigar tobacco leaves were significantly separated along the first principal component, with minimal dispersion among samples within each treatment group. All samples fell within the 95% confidence interval, indicating significant differences in protein profiles between the DDW1 and DDW2 groups. Analysis using the Proteome Discoverer (PD) software confirmed high sample repeatability in this study, as illustrated in Figure 2A. The proteomic analysis of cigar tobacco leaves in the DDW1 and DDW2 groups identified a total of 12,677 proteins. All proteins in this investigation were subjected to screening and analysis utilizing a threshold of VIP > 1, FC > 1.5 or FC < 0.67, and $P < 0.05$. The results demonstrated that, compared to the DDW2 group, the DDW1 group had 385 DEPs, with 229 up-regulated and 156 down-regulated (Figure 2B; Supplementary Table 2). KEGG pathway analysis, conducted using KOBAS 3.0 online software, revealed the top 20 enriched pathways among these DEPs (Figure 2C). The significantly enriched pathways included flavone and flavonol biosynthesis, flavonoid biosynthesis, porphyrin and chlorophyll metabolism, cysteine and methionine metabolism, amino acid biosynthesis, and general metabolic pathways.

3.3 Protein-protein interaction (PPI) analysis of Differentially Expressed Proteins (DEPs)

Protein-protein interaction networks were generated using the publicly accessible STRING database, revealing a network of 161 nodes interconnected by 102 interactions among significantly altered proteins. Cluster analysis within STRING showed that these networks were mainly composed of proteins involved in key metabolic pathways. These included porphyrin and chlorophyll metabolism, seleno-compound metabolism, amino sugar and nucleotide sugar metabolism, biosynthesis of phenylalanine, tyrosine, and tryptophan, flavonoid biosynthesis, glyoxylate and dicarboxylate metabolism, amino acid biosynthesis, flavone and flavonol biosynthesis, as well as steroid and secondary metabolites biosynthesis. The results are illustrated in Figure 3 and detailed in Supplementary Table 1.

3.4 Identification and analysis of metabolites

The PCA results are displayed in Figure 4A. Samples from the DDW1 and DDW2 groups were significantly separated along the first and second principal components, with little variation within each treatment group. All samples fell within the 95% confidence interval, highlighting significant differences in the metabolite

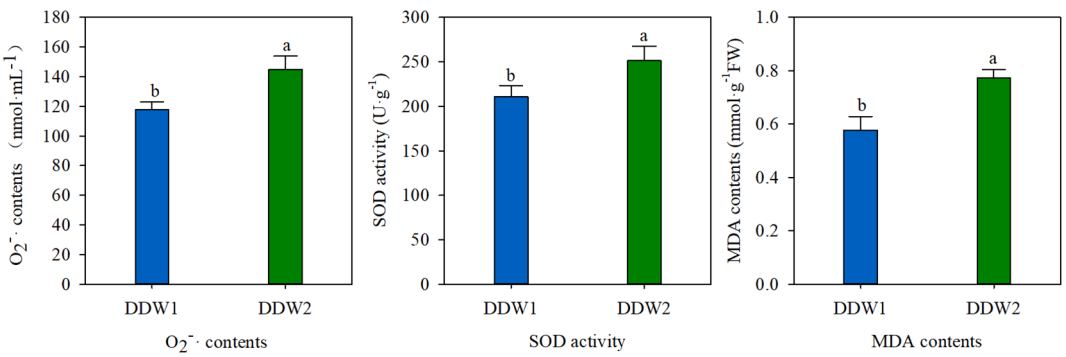


FIGURE 1
That topping has an effect on O_2^- content, SOD activity and MDA content. Different letters above the error bars indicate significant difference at the 0.05 probability level.

10.3389/fpls.2024.1425154

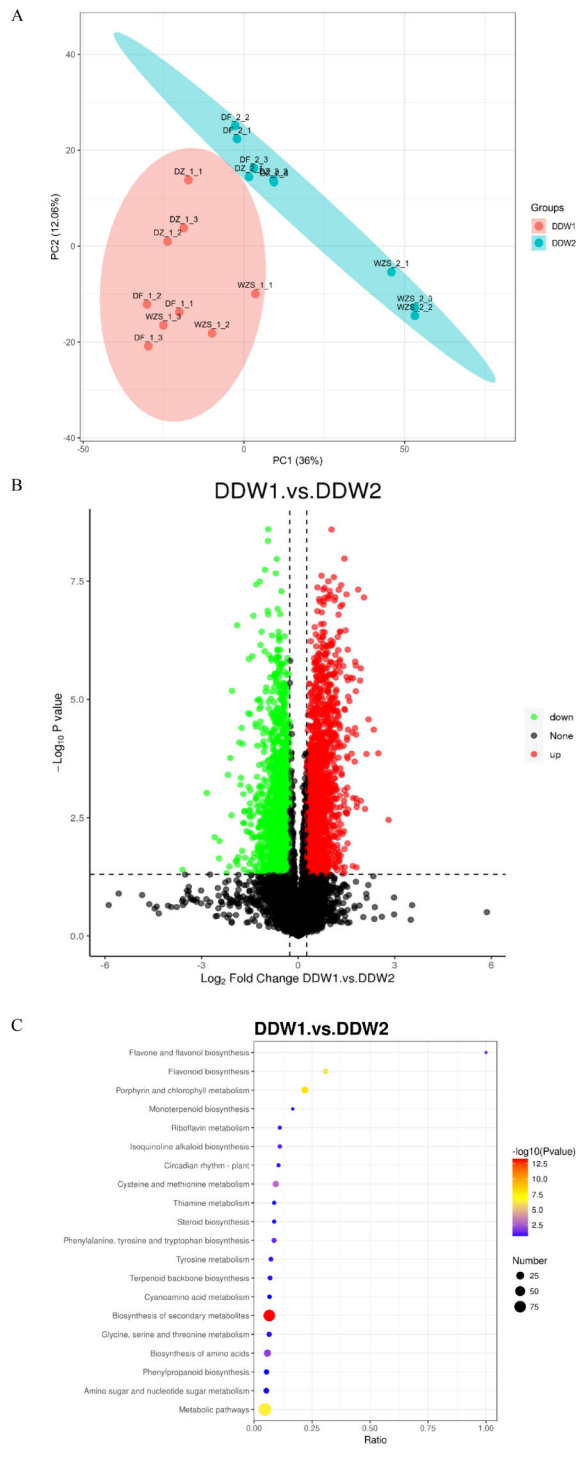


FIGURE 2
Identification of Proteins. **(A)** PCA result of proteins obtained; **(B)** Volcano map of differential proteins; **(C)** KEGG pathway enrichment result of differential proteins.

profiles of cigar tobacco leaves between the two treatment groups. Through the metabolic analysis of cigar tobacco leaves in the DDW1 and DDW2 groups, a total of 1163 metabolites were quantified. All metabolites in this investigation were screened and examined using a threshold of VIP > 1, FC > 1.2 or FC < 0.83, and $P < 0.05$. The results demonstrated that, compared to the DDW1 group, the cigar tobacco leaves in the DDW2 group had 247 DEMs, with 120 up-regulated and 128 down-regulated (Figure 4B; Supplementary Table 3). The bulk of these differential metabolites were composed of lipids and lipid-like molecules, phenylpropanoids and polyketides, organic acids and derivatives, and organoheterocyclic compounds, as depicted in Figure 4C. Pathway enrichment analysis highlighted eight key enriched pathways among these metabolites, including zeatin biosynthesis, sulfur relay system, benzoxazinoid biosynthesis, purine metabolism, flavone and flavonol biosynthesis, amino acid biosynthesis, cysteine and methionine metabolism, and flavonoid biosynthesis, shown in Figure 4D. Additionally, correlation analysis revealed that 2-hydroxyisocaproic acid had a significant positive correlation with gamma-caprolactone and octanedioic acid. Sattabacin showed a significant positive correlation with (±)-abscisic acid and methyl jasmonate, and a positive correlation was also observed between octanedioic acid and gamma-caprolactone with methyl jasmonate and (±)-abscisic acid, respectively, as detailed in Supplementary Figure 1.

3.5 Merging proteomic and metabolomic data studies

Eight primary pathways were identified through KEGG enrichment analysis of the DEPs and DEMs previously screened (Figure 5). These pathways included ‘flavonoid biosynthesis’, ‘flavone and flavonol biosynthesis’, ‘indole alkaloid biosynthesis’, ‘phenylpropanoid biosynthesis’, ‘phenylalanine metabolism’, ‘citrate cycle (TCA cycle)’, ‘monobactam biosynthesis’, and ‘pantothenate and CoA biosynthesis’.

Differential proteins and metabolites were analyzed using KEGG pathway mapper tools, with a particular focus on their role in the flavonoid biosynthesis pathway, as shown in Figure 6. Within this pathway, three notable differential metabolites were identified: pinocembrin, kaempferol, and caffeoyl shikimic acid. Pinocembrin and kaempferol were up-regulated, whereas caffeoyl shikimic acid was down-regulated. Additionally, several enzymes in this pathway, such as chalcone synthase (CHS), chalcone isomerase (CHI), naringenin 3-dioxygenase (F3H), and 5-O-(4-coumaroyl)-D-quinate 3’-monooxygenase (CYP98A), were predominantly up-regulated, suggesting an overall increase in the activity of flavonoid biosynthesis pathway.

Furthermore, four key differential metabolites-kaempferol, trifolin, rutin, and quercetin-were also up-regulated. Enzymes such as flavonoid 3’-monooxygenase (F3’H) were found to be predominantly up-regulated, indicating enhanced activity in the flavonoid biosynthesis pathway. The interactions and regulatory patterns of these enzymes and metabolites are detailed in Figure 6.

4 Discussion

Currently, most studies on topping focus primarily on tobacco leaf quality and alkaloid concentrations, often overlooking the mechanical damage it inflicts on the plant. Additionally, these

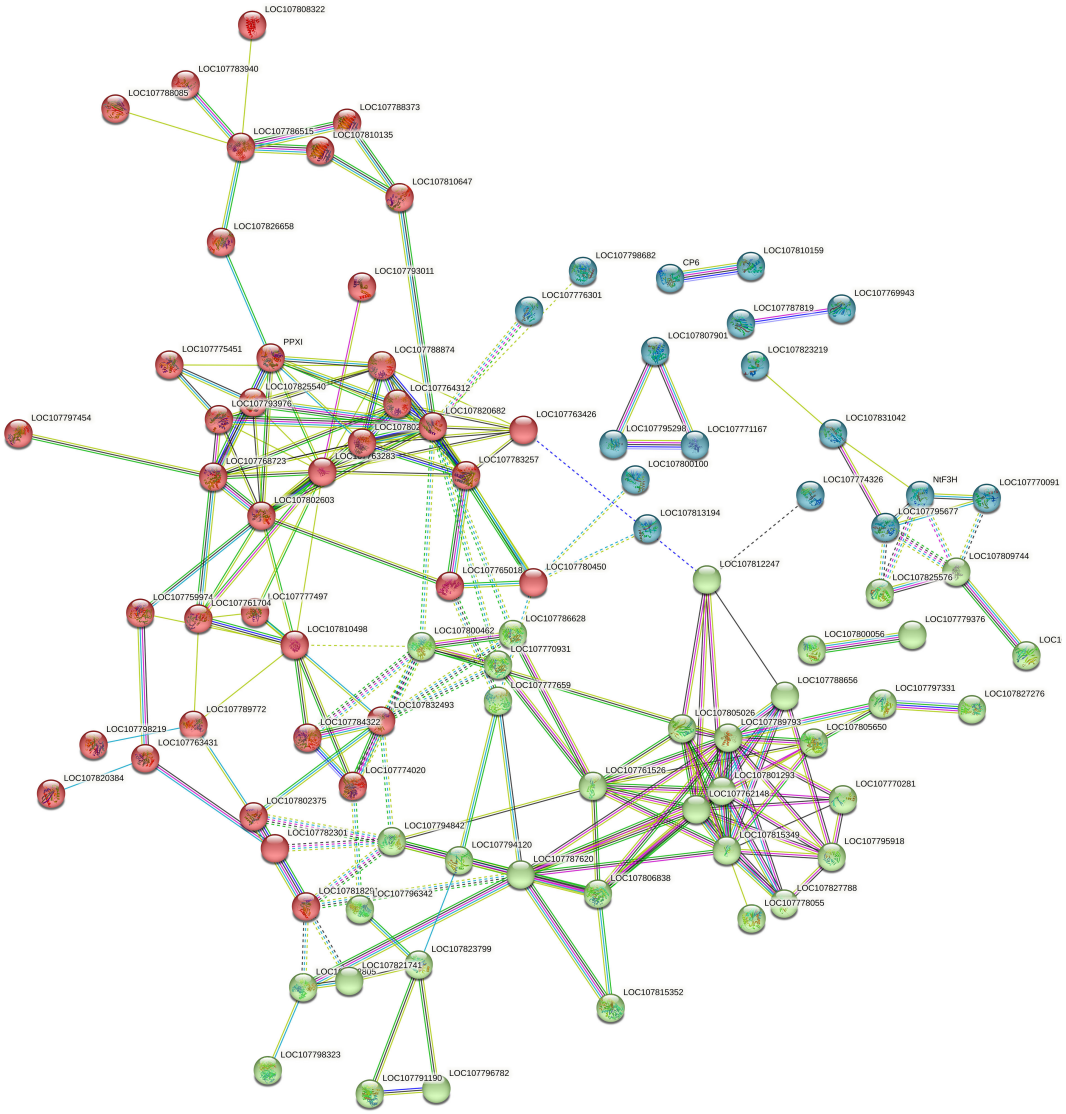


FIGURE 3
The PPI analysis of differentially expressed proteins.

studies typically employ physiological or single-omics methods to explain the process. Incorporating an integrated data approach—encompassing genetics, epigenetics, transcriptomics, and proteomics—can offer a more comprehensive view of plant responses to various abiotic stresses. Liu et al. (2016) mentioned that protein levels do not always reflect metabolite abundance, suggesting that relying solely on one method, like proteomics or metabolomics, may lead to incomplete or inaccurate interpretations. Zapalska-Sozoniuk et al. (2019) stressed the importance of integrating multiple omics approaches for accurate interpretation. By applying multi-omics studies to understand how tobacco plants respond to the abiotic stress (mechanical damage) of topping, we can gain a deeper understanding of their tolerance mechanisms.

Topping in plants results in mechanical damage that disrupts oxidative balance, leading to an increase in ROS (Das and Roychoudhury, 2014). Stress conditions exacerbate this imbalance, causing ROS accumulation that can damage proteins, lipids, and DNA (Hernandez et al., 2009). In response, plants produce various antioxidant enzymes, such as SOD and ascorbate peroxidase (APX), and small molecules like ascorbic acid (AA) and glutathione (GSH) to counteract ROS effects (Mullineaux and Rausch, 2005; Gill and Tuteja, 2010; Miller et al., 2010). Consistent with these results, our study found that the DDW2 group had higher levels of superoxide anion, SOD activity, and MDA content compared to the DDW1 group (Figures 1). GO enrichment analysis of DEPs showed a focus on oxidative stress-related terms, such as “flavonoid biosynthesis” and “flavonol and flavone biosynthesis” (Figure 2C). Similarly, DEMs were enriched in terms linked to oxidative stress, including “zeatin biosynthesis,” “flavonoid biosynthesis,” and “flavonol and flavone biosynthesis” (Figure 3). These findings denote that many metabolic pathways, especially those involved in flavonoid and flavonol/flavone biosynthesis, are altered under topping conditions, indicating a

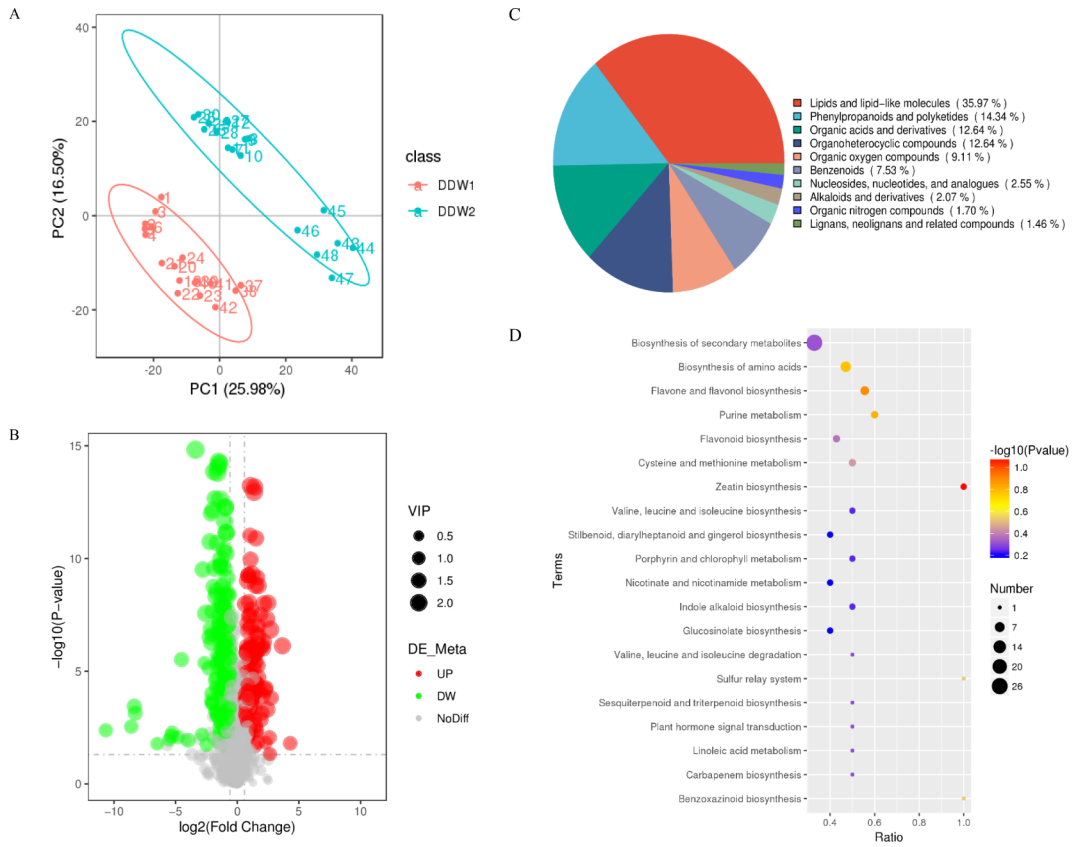


FIGURE 4
Identification and Analysis of Metabolites. **(A)** PCA result of metabolites obtained; **(B)** Volcano map of differential metabolites; **(C)** the pie chart shows the category results of the differential metabolites; **(D)** KEGG pathway enrichment result of differential proteins.

potential link between increased ROS production and the topping process.

KEGG enrichment analysis of DEPs and DEMs revealed the response pathways in tobacco under topping-induced stress. The findings showed that topping conditions significantly alter several metabolic pathways, especially those involved in flavonoid and flavonol/flavone biosynthesis (Figures 2C, 3D, 4). This implies that flavonoid compounds may play a key role in tobacco's response to topping (mechanical damage). Flavonoids, a major class of polyphenolic secondary metabolites with over 8000 identified compounds in plants, including flavone, flavonol, and isoflavone, exhibit a range of biological activities due to their unique chemical structures (Yuan et al., 2015). These activities include photoprotection (Harvaux and Kloppstech, 2001), ROS scavenging (Gayomba and Muday, 2020), regulation of auxin transport, pollinator attraction (Kellenberger et al., 2019), pathogen resistance (Yang et al., 2021), stomatal aperture regulation (Watkins et al., 2014), promotion of pollen tube growth (Muhlemann et al., 2018), and influence on root development (Silva-Navas et al., 2016). While most research on flavonoids in tobacco has concentrated on color and quality changes, their role in the plant's resistance to adverse conditions has been less explored.

In flavonoid biosynthesis, CHS is the initial key enzyme that shifts the metabolic process from phenylpropanoid to flavonoid synthesis (Wang et al., 2017), by catalyzing the reaction between coumaroyl-CoA and malonyl-CoA to generate flavonoid precursors. CHI, the second essential enzyme in this pathway, facilitates the intramolecular cyclization of chalcones to produce flavanones, which are then further modified into various flavonoid structures (Wu et al., 2018; Wang et al., 2017). Studies indicate that both CHS and CHI play significant roles in synthesizing various defensive compounds within the phenylpropanoid pathway and are directly involved in these processes. Plant responses to external stimuli, such as stress and pathogen attacks, involve the rapid activation of CHS, which boosts stress resistance (Turnbull et al., 2004; Petrussa et al., 2013). For instance, incorporating the CHS gene into poplar trees decreased their sensitivity to low temperatures. Similarly, co-expressing CHS and flavonol synthase genes in tomatoes enhanced flavonol production and antioxidant capacity (Verhoeyen et al., 2002). Additionally, overexpression of CHS, CHI, and DFR genes in potato tubers increased anthocyanin and flavonoid levels, thereby improving antioxidant properties (Lukaszewicz et al., 2004). In our study, we observed an upregulation of CHS and CHI in the biosynthesis pathway (Figure 6), leading to higher pinosylvin levels. Pinosylvin, a key flavonoid found in propolis with significant antimicrobial properties (Rasul et al., 2013), appears to be linked to the mechanical damage caused by topping, and its increased production could enhance antibacterial effects.

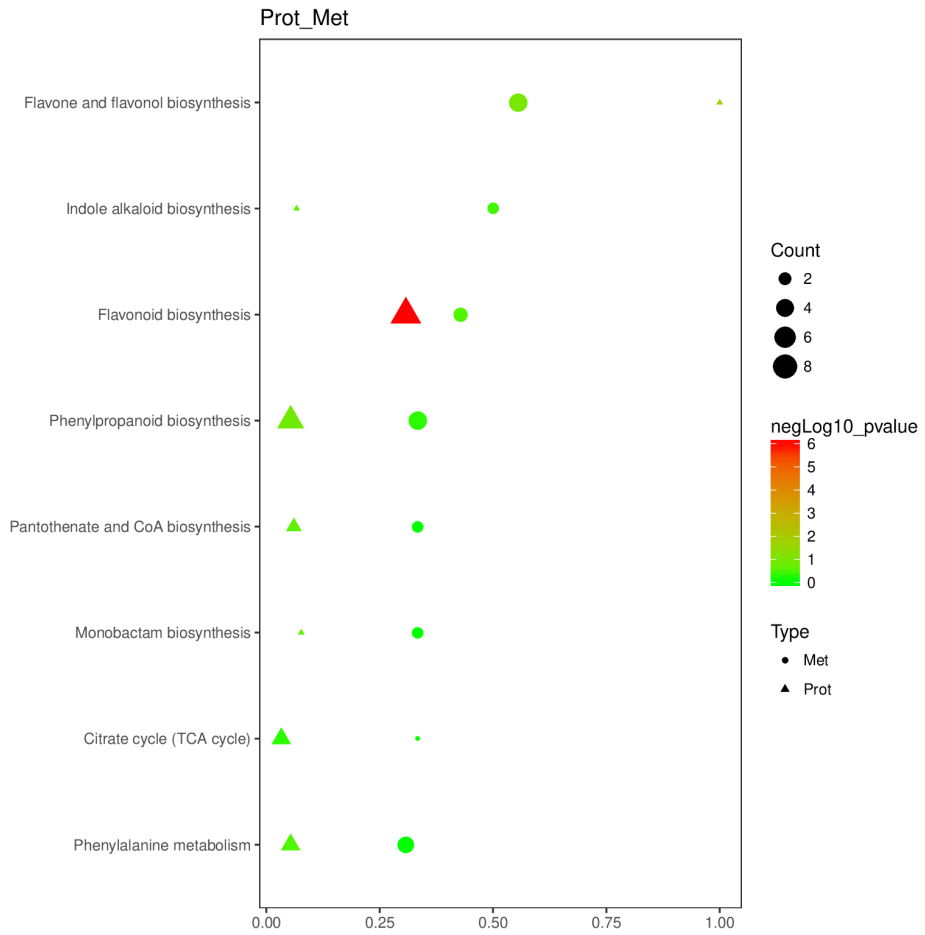


FIGURE 5
Bubble diagram of enrichment pathway of DEPs and DEMs of cigar tobacco leaves in DDW1 and DDW2 groups. Protein was represented by triangles; metabolite was represented by circles. Horizontal axis indicates the enrichment degree of DEPs and DEMs. Point color represents the P-value, and point size indicates the number of DEPs and DEMs in the corresponding pathway.

Flavanone 3-hydroxylase (F3H) is a crucial enzyme in the flavonoid biosynthetic pathway, responsible for converting flavanones into dihydroflavonols, which are precursors for flavonol and anthocyanins (Cheng et al., 2013). Regulation of the F3H gene influences plant pigment levels, which enhances plant adaptability and survival, such as increasing resistance to ultraviolet radiation (Prochazkova et al., 2011). It plays an important role in plant morphogenesis, physiological and biochemical functions, and coping with environmental stress (Prochazkova et al., 2011; Ma et al., 2020). The F3H gene has been cloned and identified in several plant species, including *Malus pumila* Mill (Davies, 1993), *Vitis vinifera* L (Sparvoli et al., 1994), *Zea mays* L (Deboer et al., 2009), *Arabidopsis thaliana* (Pelletier and Shirley, 1996), *Glycine max* (Zabala and Vodkin, 2005) and *Carthamus tinctorius* L (Tu et al., 2016). Evidence indicates that F3H is influenced by various environmental factors. For example, gibberellic acid and sucrose boost the expression of EsF3H in epimedium, enhancing the accumulation of flavonoid bioactive compounds (Zeng et al., 2013). Similarly, JA and abscisic acid (ABA) regulate the expression of AtF3H involved in flavonoid biosynthesis in *Arabidopsis* (Loreti et al., 2008; Lewis et al., 2011). Overexpression of F3H in *Arabidopsis* increases resistance to salt and oxidative stress (Li et al., 2017). Our research found that the upregulation of CHS, CHI, and F3H proteins in the flavonoid biosynthesis pathway led to higher levels of dihydroflavonols and subsequent naringenin metabolites. Naringenin, a key flavonoid with antioxidant and anti-inflammatory properties, inhibits protein kinases. Our research shows that CHS and CHI enzymes are mainly involved in synthesizing dihydroflavonols, which, together with FL3, are essential for naringenin production. Naringenin effectively prevents lipid peroxidation and scavenges superoxide anions, aligning with previous studies (Sparvoli et al., 1994). Additionally, increased sunlight and UV-B radiation significantly impact the content and ratio of naringenin and other flavonoids in plant leaves, affecting their photoprotective abilities (Ryan et al., 2002; Zhao et al., 2020; Zhang et al., 2023b).

Flavone and flavonol are classes of natural compounds produced via the phenylpropanoid pathway (Gao et al., 2021). This process begins with phenylalanine, which is first converted into cinnamic acid by phenol oxidase. Cinnamic acid is then converted into coumaric acid through a series of steps. Coumaric acid, a type of organic acid, is subsequently transformed into chalcones through the action of

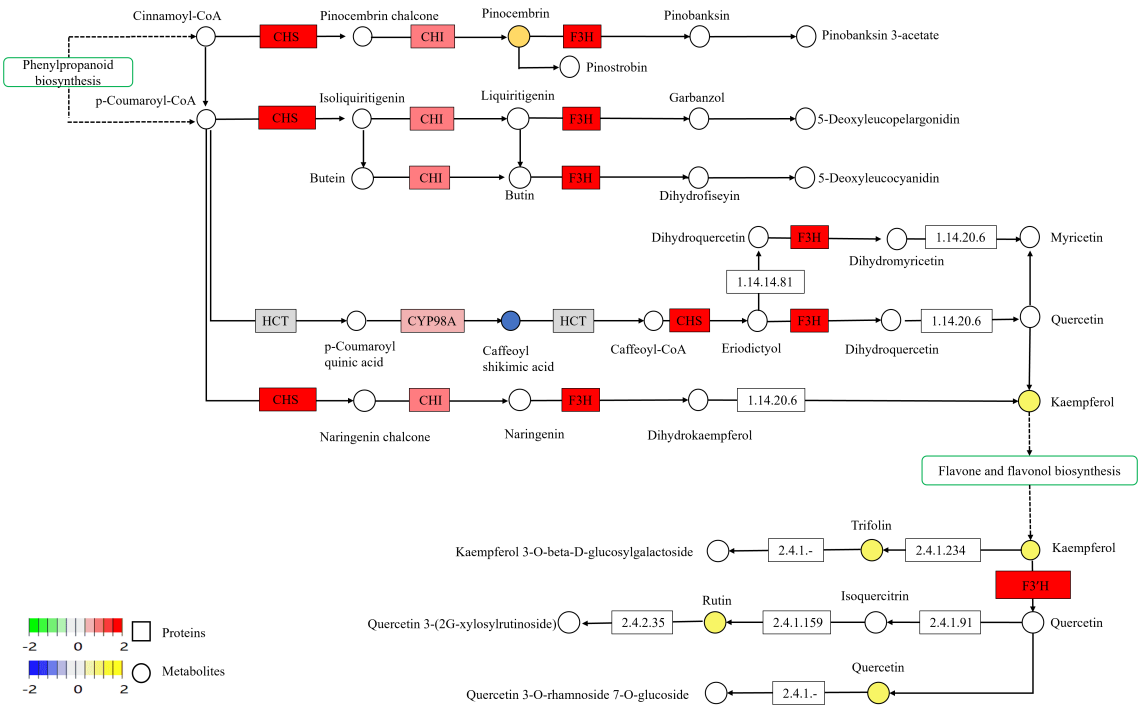


FIGURE 6
Changes in flavonoid biosynthesis proteins and metabolism. The pathways were drawn based on the KEGG database. CHS, chalcone synthase [EC:2.3.1.74]; CHI, chalcone isomerase [EC:5.5.1.6]; F3H, naringenin 3-dioxygenase [EC:1.14.11.9]; HCT, shikimate O-hydroxycinnamoyltransferase [EC:2.3.1.133]; CYP98A, 5-O-(4-coumaroyl)-D-quinate 3′-monooxygenase [EC:1.14.14.96]; F3′H flavonoid 3′-monooxygenase; flavonoid [EC:1.14.14.82].

chalcone synthase. Finally, chalcones are converted into flavone or flavonol by flavon synthase (Wu et al., 2018).

Flavonoid 3'-hydroxylase (F3'H, EC: 1.14.13.21) is a cytochrome P450 monooxygenase enzyme (Shi et al., 2006; Li et al., 2021) belonging to the CYP75 subfamily and plays a crucial role in the synthesis of flavonoid compounds. Our study found that topping increases F3'H expression. This enzyme acts on various substrates, including apigenin, naringenin, kaempferol, and dihydrokaempferol. These results align with previous findings that F3'H is upregulated under stress conditions, boosting resistance. For instance, in *Citrus sinensis*, the gene CsF3'H is significantly induced by drought stress. In transgenic *Arabidopsis* plants overexpressing CsF3'H, there are lower levels of ROS and higher levels of antioxidant flavonoids and antioxidant enzyme activity compared to wild-type plants, thereby improving drought resistance (Guo et al., 2011).

It was also found that trifolin, rutin, and quercetin experienced alterations in their biosynthesis pathways. Following topping, levels of these compounds increased, implying their antioxidant properties. Liu et al. (2021) investigated lilac (*Syringa*) responses to light stress and discovered that light can regulate the expression of key genes in the rutin synthesis pathway, such as 4CL1, CYP73A, and CYP75B1, leading to higher rutin levels. Conversely, salt stress impaired germination and seedling growth in *Apocynum venetum* L., resulting in reduced total flavonoid content. However, the content of the flavonol quercetin increased, with the upregulation of genes involved in its synthesis, including AvF3'H, AvF3H, and AvFLS (Xu et al., 2020).

Research has demonstrated that the biosynthesis of flavonoids, flavones, and flavonols can boost a plant's antioxidant defense (Prochazkova et al., 2011; Petrussa et al., 2013). This study found that topping significantly influenced the biosynthesis of these compounds, along with the activities of O_2^-, SOD, and MDA. Protein and metabolism analysis revealed that CHS, CHI, F3H, and F3'H were upregulated, and the levels of metabolites such as kaempferol, clover, rutin, and quercetin were increased. These changes support the accumulation of antioxidant enzymes, help maintain oxidative balance, and protect against oxidative stress by continuously synthesizing and accumulating these compounds. They also shield lipids and membrane proteins from oxidative damage, neutralize ROS, and reduce oxidative cell damage. While these findings shed light on the activation of antioxidant mechanisms following topping in cigar tobacco, further research is needed to gain a more comprehensive understanding of the underlying protein and metabolic processes.

5 Conclusion

In this study, we revealed that topping significantly affects the intrinsic proteins and metabolites in cigar tobacco leaves. This impairment modifies the amount of proteins associated with flavonoid biosynthesis and flavone and flavonol biosynthesis, including CHS, CHI, F3H, and F3'H, as well as the accumulation of the metabolic products kaempferol, trifolin, rutin, and quercetin. Consequently, this encourages the buildup of antioxidative enzymes.

To maintain oxidative balance and counteract oxidative stress, flavonoids are continuously produced and accumulated. This process helps protect lipids and membrane proteins from oxidative damage, neutralizes ROS, and impedes their formation, thus mitigating oxidative cell damage. Our findings offer novel insights into the metabolomic and proteomic responses of cigar tobacco to topping.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: ProteomeXchange, PXD060658.

Author contributions

DG: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. HG: Methodology, Writing – review & editing. TY: Methodology, Writing – original draft, Writing – review & editing. CX: Methodology, Project administration, Writing – review & editing. BL: Supervision, Writing – review & editing. XX: Supervision, Validation, Writing – review & editing. BC: Methodology, Writing – review & editing. ZG: Funding acquisition, Project administration, Resources, Writing – review & editing.

Funding

The author(s) declare that financial support was received for theresearch, authorship, and/or publication of this article. This studyreceived funding from Science and Technology Major Project of ChinaNational Tobacco Corporation (110202201038(XJ-09), 110202103017, 110202101012(XJ-04)),Key Science and Technology Projects of Hainan Provincial Branch ofChina National Tobacco Corporation (2020001). 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.

Conflict of interest

Authors DG, HG, TY, CX, BL, XX, BC, and ZG were employed by the company Hainan Provincial Branch of China National Tobacco Corporation.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1425154/full#supplementary-material

References

Baldwin, I. T. (2001). An ecologically motivated analysis of plant-herbivore interactions in native tobacco. *Plant Physiol.* 127, 1449–1458. doi: 10.1104/pp.010762

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

Cao, P., Yang, J., Xia, L. H., Zhang, Z. H., Wu, Z. Y., Hao, Y. C., et al. (2024). Two gene clusters and their positive regulator SlMYB13 that have undergone domestication-associated negative selection control phenolamide accumulation and drought tolerance in tomato. *Mol. Plant* 17, 579–597. doi: 10.1016/j.molp.2024.02.003

Chen, J., Li, Y., He, X., Jiao, F., Xu, M., Hu, B., et al. (2021). Influences of different curing methods on chemical compositions in different types of tobaccos. *Ind. Crops Products* 167, 113534. doi: 10.1016/j.indcrop.2021.113534

Chen, X., Sun, S., Liu, F., Shen, E., Liu, L., Ye, C., et al. (2019). A transcriptomic profile of topping responsive non-coding RNAs in tobacco roots (Nicotiana tabacum). *BMC Genomics* 20, 856. doi: 10.1186/s12864-019-6236-6

Cheng, H., Wang, J., Chu, S., Yan, H. L., and Yu, D. (2013). Diversifying selection on flavanone 3-hydroxylase and isoflavone synthase genes in cultivated soybean and its wild progenitors. *PloS One* 8, e54154. doi: 10.1371/journal.pone.0054154

Dai, J., Tian, L., Zhang, Y., Zhang, D., Xu, S., Cui, Z., et al. (2022). Plant topping effects on growth, yield, and earliness of field-grown cotton as mediated by plant density and ecological conditions. *Field Crops Res.* 1, 275. doi: 10.1016/j.fcr.2021.108337

Das, K., and Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Front. Environ. Sci.* 2, 3389. doi: 10.3389/fenvs.2014.00053

Davies, K. M. (1993). A cDNA clone for flavanone 3-hydroxylase from Malus. *Plant Physiol.* 103, 291. doi: 10.1104/pp.103.1.291

Deboer, K. D., Lye, J. C., Aitken, C. D., Su, A. K., and Hamill, J. D. (2009). The A622 gene in Nicotiana glauca (tree tobacco): evidence for a functional role in pyridine alkaloid synthesis. *Plant Mol. Biol.* 69, 299–312. doi: 10.1007/s11103-008-9425-2

Draper, H. H., Squires, E. J., Mahmoodi, H., Wu, J., Agarwal, S., and Hadley, M. (1993). A comparative evaluation of thiobarbituric acid methods for the determination of malondialdehyde in biological materials. *Free Radic. Biol. Med.* 15, 353–363. doi: 10.1016/0891-5849(93)90035-S

Durak, I., Yurtarslanl, Z., Canbolat, O., and Akyol, O. (1993). A methodological approach to superoxide dismutase (SOD) activity assay based on inhibition of nitroblue tetrazolium (NBT) reduction. *Clin. Chim. Acta* 214, 103–104. doi: 10.1016/0009-8981(93)90307-P

Gao, G., Lv, Z., Zhang, G., Li, J., Zhang, J., and He, C. (2021). An ABA-flavonoid relationship contributes to the differences in drought resistance between different sea buckthorn subspecies. *Tree Physiol.* 41, 744–755. doi: 10.1093/treephys/tpaa155

Gaquerel, E., Weinhold, A., and Baldwin, I. T. (2009). Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphigidae) and its natural host Nicotiana attenuata. VIII. An unbiased GCxGC-ToFMS analysis of the plant's elicited volatile emissions. *Plant Physiol.* 149, 1408–1423. doi: 10.1104/pp.108.130799

Gayomba, S. R., and Muday, G. K. (2020). Flavonols regulate root hair development by modulating accumulation of reactive oxygen species in the root epidermis. *Development* 147, 185819. doi: 10.1242/dev.185819

Geuns, J. M., Smets, R., Struyf, T., Prinsen, E., Valcke, R., and Van, H. (2001). Apical dominance in Pssu-ipt-transformed tobacco. *Phytochemistry* 58, 911–921. doi: 10.1016/S0031-9422(01)00338-7

Gill, S. S., and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* 48, 909–930. doi: 10.1016/j.plaphy.2010.08.016

Guo, H., Cao, P., Wang, C., Lai, J., Deng, Y., Li, C., et al. (2023). Population analysis reveals the roles of DNA methylation in tomato domestication and metabolic diversity. *Sci. China Life Sci.* 66, 1888–1902. doi: 10.1007/s11427-022-2299-5

Guo, H., Kan, Y., and Liu, W. (2011). Differential expression of miRNAs in response to topping in flue-cured tobacco (*Nicotiana tabacum*) roots. *PloS One* 6, e28565. doi: 10.1371/journal.pone.0028565

Harvaux, M., and Kloppstech, K. (2001). The protective functions of carotenoid and flavonoid pigments against excess visible radiation at chilling temperature investigated in Arabidopsis npq and tt mutants. *Planta* 213, 953–966. doi: 10.1007/s004250100572

Hernandez, I., Alegre, L., Van Breusegem, F., and Munne-Bosch, S. (2009). How relevant are flavonoids as antioxidants in plants? *Trends Plant Sci.* 14, 125–132. doi: 10.1016/j.tplants.2008.12.003

Jin, C., Fang, C., Yuan, H., Wang, S., Wu, Y., Liu, X., et al. (2015). Interaction between carbon metabolism and phosphate accumulation is revealed by a mutation of a cellulose synthase-like protein, CSLF6. *J. Exp. Bot.* 66, 2557–2567. doi: 10.1093/jxb/erv050

Katoh, A., Shoji, T., and Hashimoto, T. (2007). Molecular cloning of N-methylputrescine oxidase from tobacco. *Plant Cell Physiol.* 48, 550–554. doi: 10.1093/pcp/pcm018

Kellenberger, R. T., Byers, K., De Brito Francisco, R. M., Staedler, Y. M., Lafountain, A. M., Schonenberger, J., et al. (2019). Emergence of a floral colour polymorphism by pollinator-mediated overdominance. *Nat. Commun.* 10, 63. doi: 10.1038/s41467-018-07936-x

Lei, B., Chang, W., Zhao, H., Zhang, K., Yu, J., Yu, S., et al. (2022). Nitrogen application and differences in leaf number retained after topping affect the tobacco (*Nicotiana tabacum*) transcriptome and metabolome. *BMC Plant Biol.* 22, 38. doi: 10.1186/s12870-022-03426-x

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

Lewis, D. R., Ramirez, M. V., Miller, N. D., Vallabhaneni, P., Ray, W. K., Helm, R. F., et al. (2011). Auxin and ethylene induce flavonol accumulation through distinct transcriptional networks. *Plant Physiol.* 156, 144–164. doi: 10.1104/pp.111.172502

Li, M., Cao, Y. T., Debnath, B., Yang, H. J., Kui, X. H., and Qiu, D. L. (2021). Cloning and expression analysis of flavonoid 3', 5'-hydroxylase gene from *Brunfelsia acuminata*. *Genes (Basel)* 12, 1086. doi: 10.3390/genes12071086

Li, C., Liu, S., Yao, X., Wang, J., Wang, T., Zhang, Z., et al. (2017). PnF3H, a flavanone 3-hydroxylase from the Antarctic moss Pohlia nutans, confers tolerance to salt stress and ABA treatment in transgenic Arabidopsis. *Plant Growth Regul.* 83, 489–500. doi: 10.1007/s10725-017-0314-z

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

Liu, F., Wu, Z., Zhang, X., Xi, G., Zhao, Z., Lai, M., et al. (2021). Microbial community and metabolic function analysis of cigar tobacco leaves during fermentation. *Microbiologyopen* 10, e1171. doi: 10.1002/mbo3.v10.2

Loreti, E., Povero, G., Novi, G., Solfanelli, C., Alpi, A., and Perata, P. (2008). Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in Arabidopsis. *New Phytol.* 179, 1004–1016. doi: 10.1111/j.1469-8137.2008.02511.x

Lukaszewicz, M., Matysiak-Kata, I., Skala, J., Fecka, I., Cisowski, W., and Szopa, J. (2004). Antioxidant capacity manipulation in transgenic potato tuber by changes in phenolic compounds content. *J. Agric. Food Chem.* 52, 1526–1533. doi: 10.1021/jf034482k

Luo, P., Shen, Y., Jin, S., Huang, S., Cheng, X., Wang, Z., et al. (2016). Overexpression of rosa rugosa anthocyanidin reductase enhances tobacco tolerance to abiotic stress through increased ROS scavenging and modulation of ABA signaling. *Plant Sci.* 245, 35–49. doi: 10.1016/j.plantsci.2016.01.007

Ma, Y., Feng, Y., Diao, T., Zeng, W., and Zuo, Y. (2020). Experimental and theoretical study on antioxidant activity of the four anthocyanins. *J. Mol. Structure* 1204, 127509–127516. doi: 10.1016/j.molstruc.2019.127509

Ma, D. M., Gandra, S. V. S., Manoharlal, R., La Hovary, C., and Xie, D. Y. (2019). Untargeted metabolomics of nicotiana tabacum grown in United States and India characterizes the association of plant metabolomes with natural climate and geography. *Front. Plant Sci.* 10, 1370. doi: 10.3389/fpls.2019.01370

Miller, G., Suzuki, N., Ciftci-Yilmaz, S., and Mittler, R. (2010). Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant Cell Environ.* 33, 453–467. doi: 10.1111/j.1365-3040.2009.02041.x

Muhlemann, J. K., Younts, T. L. B., and Muday, G. K. (2018). Flavonols control pollen tube growth and integrity by regulating ROS homeostasis during high-temperature stress. *Proc. Natl. Acad. Sci. U.S.A.* 115, 11188–11197. doi: 10.1073/pnas.1811492115

Mullineaux, P. M., and Rausch, T. (2005). Glutathione, photosynthesis and the redox regulation of stress-responsive gene expression. *Photosynth Res.* 86, 459–474. doi: 10.1007/s11120-005-8811-8

Ogilvie, I. S., and Kozumplik, V. (1980). Genetic analysis of quantitative characters in cigar and pipe tobacconicotiana tabacum. i. morphological characters. *Can. J. Genet. Cytology* 22, 173–182. doi: 10.1139/g80-022

Pelletier, M. K., and Shirley, B. W. (1996). Analysis of flavanone 3-hydroxylase in Arabidopsis seedlings. coordinate regulation with chalcone synthase and chalcone isomerase. *Plant Physiol.* 111, 339–345. doi: 10.1104/pp.111.1.339

Petrussa, E., Braidot, E., Zancani, M., Peresson, C., Bertolini, A., Patui, S., et al. (2013). Plant flavonoids–biosynthesis, transport and involvement in stress responses. *Int. J. Mol. Sci.* 14, 14950–14973. doi: 10.3390/ijms140714950

Prochazkova, D., Bousova, I., and Wilhelmova, N. (2011). Antioxidant and prooxidant properties of flavonoids. *Fitoterapia* 82, 513–523. doi: 10.1016/j.fitote.2011.01.018

Qin, Y., Bai, S., Li, W., Sun, T., Galbraith, D. W., Yang, Z., et al. (2020). Transcriptome analysis reveals key genes involved in the regulation of nicotine biosynthesis at early time points after topping in tobacco (Nicotiana tabacum L.). *BMC Plant Biol.* 20, 30. doi: 10.1186/s12870-020-2241-9

Rasul, A., Millimouno, F. M., Ali Eltayb, W., Ali, M., Li, J., and Li, X. (2013). Pinocembrin: a novel natural compound with versatile pharmacological and biological activities. *BioMed. Res. Int.* 2013, 379850. doi: 10.1155/2013/379850

Ryan, K. G., Swinny, E. E., Markham, K. R., and Winefield, C. (2002). Flavonoid gene expression and UV photoprotection in transgenic and mutant Petunia leaves. *Phytochemistry* 59, 23–32. doi: 10.1016/S0031-9422(01)00404-6

Shi, Q., Li, C., and Zhang, F. (2006). Nicotine synthesis in *Nicotiana tabacum L.* induced by mechanical wounding is regulated by auxin. *J. Exp. Bot.* 57, 2899–2907. doi: 10.1093/jxb/erl051

Silva-Navas, J., Moreno-Risueno, M. A., Manzano, C., Tellez-Robledo, B., Navarro-Neila, S., Carrasco, V., et al. (2016). Flavonols mediate root phototropism and growth through regulation of proliferation-to-differentiation transition. *Plant Cell* 28, 1372–1387. doi: 10.1105/tpc.15.00857

Sparvoli, F., Martin, C., Scienza, A., Gavazzi, G., and Tonelli, C. (1994). Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera L.*). *Plant Mol. Biol.* 24, 743–755. doi: 10.1007/BF00029856

Tu, Y., Liu, F., Guo, D., Fan, L., Zhu, Z., Xue, Y., et al. (2016). Molecular characterization of flavanone 3-hydroxylase gene and flavonoid accumulation in two chemotyped safflower lines in response to methyl jasmonate stimulation. *BMC Plant Biol.* 16, 132. doi: 10.1186/s12870-016-0813-5

Turnbull, J. J., Nakajima, J., Welford, R. W., Yamazaki, M., Saito, K., and Schofield, C. J. (2004). Mechanistic studies on three 2-oxoglutarate-dependent oxygenases of flavonoid biosynthesis: anthocyanidin synthase, flavonol synthase, and flavanone 3beta-hydroxylase. *J. Biol. Chem.* 279, 1206–1216. doi: 10.1074/jbc.M309228200

Verhoeyen, M. E., Bovy, A., Collins, G., Muir, S., Robinson, S., De Vos, C. H., et al. (2002). Increasing antioxidant levels in tomatoes through modification of the flavonoid biosynthetic pathway. *J. Exp. Bot.* 53, 2099–2106. doi: 10.1093/jxb/erf044

Wang, S. C., Alseekh, S., Fernie, A. R., and Luo, J. (2019). The structure and function of major plant metabolite modifications. *Mol. Plant* 12, 899–919. doi: 10.1016/j.molp.2019.06.001

Wang, S. C., Xiao, Y., Zhou, Z., Yuan, J., Guo, H., Yang, Z., et al. (2021). High-quality reference genome sequences of two coconut cultivars provide insights into evolution of monocot chromosomes and differentiation of fiber content and plant height. *Genome Biol.* 22, 304. doi: 10.1186/s13059-021-02522-9

Wang, S. C., Yang, C., Tu, H., Zhou, J., Liu, X., Cheng, Y., et al. (2017). Characterization and metabolic diversity of flavonoids in citrus species. *Sci. Rep-UK* 7, 10549. doi: 10.1038/s41598-017-10970-2

Watkins, J. M., Hechler, P. J., and Muday, G. K. (2014). Ethylene-induced flavonol accumulation in guard cells suppresses reactive oxygen species and moderates stomatal aperture. *Plant Physiol.* 164, 1707–1717. doi: 10.1104/pp.113.233528

Wu, Y., Guo, J., Zhou, Q., Xin, Y., Wang, G., and Xu, L. A. (2018). *De novo* transcriptome analysis revealed genes involved in flavonoid biosynthesis, transport and regulation in Ginkgo biloba. *Ind. Crops Products* 124, 226–235. doi: 10.1016/j.indcrop.2018.07.060

Xu, Z., Zhou, J., Ren, T., Du, H., Liu, H., Li, Y., et al. (2020). Salt stress decreases seedling growth and development but increases quercetin and kaempferol content in Apocynum venetum. *Plant Biol. (Stuttg)* 22, 813–821. doi: 10.1111/plb.13128

Yan, S., Niu, Z., Yan, H., Zhang, A., and Liu, G. (2019). Transcriptome sequencing reveals the effect of biochar improvement on the development of tobacco plants before and after topping. *PloS One* 14, e0224556. doi: 10.1371/journal.pone.0224556

Yang, J., Chen, R., Wang, C., Li, C., Ye, W., Zhang, Z., et al. (2024). A widely targeted metabolite modificomics strategy for modified metabolites identification in tomato. *J. Integr. Plant Biol.* 66, 810–823. doi: 10.1111/jipb.13629

Yang, Z., Li, N., Kitano, T., Li, P., Spindel, J. E., Wang, L., et al. (2021). Genetic mapping identifies a rice naringenin O-glucosyltransferase that influences insect resistance. *Plant J.* 106, 1401–1413. doi: 10.1111/tpj.v106.5

Yuan, H., Tang, G., Jian, X., and Pan, H. (2015). Biosynthesis of natural products and synthetic biology of antitumor drugs. *Scientia Sin. (Vitae)* 45, 1027–1039. doi: 10.1360/N052015-00049

Zabala, G., and Vodkin, L. O. (2005). The wp mutation of glycine max carries a gene-fragment-rich transposon of the CACTA superfamily. *Plant Cell* 17, 2619–2632. doi: 10.1105/tpc.105.033506

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, S., Liu, Y., Hu, W., Liu, Y., Shen, X., and Wang, Y. (2013). Integrated transcriptional and phytochemical analyses of the flavonoid biosynthesis pathway in Epimedium. *Plant Cell Tissue Organ Culture* 115, 355–365. doi: 10.1007/s11240-013-0367-2

Zhang, Q., Kong, G., Zhao, G., Liu, J., Jin, H., Li, Z., et al. (2023a). Microbial and enzymatic changes in cigar tobacco leaves during air-curing and fermentation. *Appl. Microbiol. Biotechnol.* 107, 5789–5801. doi: 10.1007/s00253-023-12663-5

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, P., Wang, X., Lu, Q., Zhang, H., Chen, J., Zhang, H., et al. (2023b). Allantoin, a purine metabolite, confers saline–alkaline tolerance to sugar beet by triggering a self-amplifying feedback loop comprising jasmonic acid and nitric oxide. *Environ. Exp. Bot.* 206, 1–16. doi: 10.1016/j.envexpbot.2022.105172

Zhao, B., Wang, L., Pang, S., Jia, Z., Wang, L., Li, W., et al. (2020). UV-B promotes flavonoid synthesis in Ginkgo biloba leaves. *Ind. Crops Products* 151, 112483–112495. doi: 10.1016/j.indcrop.2020.112483

Zhuang, K., Shi, D., Hu, Z., Xu, F., Chen, Y., and Shen, Z. (2019). Subcellular accumulation and source of O_2^- and H_2O_2 in submerged plant Hydrilla verticillata (L.f.) Royle under NH_4^+-N stress condition. *Aquat Toxicol.* 207, 1–12. doi: 10.1016/j.aquatox.2018.11.011

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 07 March 2025
DOI 10.3389/fpls.2025.1538607

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Qinghua Shi,
Shandong Agricultural University, China
Jelli Venkatesh,
International Atomic Energy Agency, Austria

*CORRESPONDENCE
Biao Zhu
billzhu@zafu.edu.cn

†These authors have contributed equally to this work

RECEIVED 03 December 2024
ACCEPTED 04 February 2025
PUBLISHED 07 March 2025

CITATION
Ma X, Liang G, Xu Z, Lin C and Zhu B (2025) CaMYBA–CaMYC–CaTTG1 complex activates the transcription of anthocyanin synthesis structural genes and regulates anthocyanin accumulation in pepper (*Capsicum annuum* L.) leaves. *Front. Plant Sci.* 16:1538607. doi: 10.3389/fpls.2025.1538607

CaMYBA–CaMYC–CaTTG1 complex activates the transcription of anthocyanin synthesis structural genes and regulates anthocyanin accumulation in pepper (*Capsicum annuum* L.) leaves

Xiaowei Ma†, Guangbo Liang†, Ziqian Xu, Chenwei Lin and Biao Zhu*

Key Laboratory of Quality and Safety Control for Subtropical Fruit and Vegetable, Ministry of Agriculture and Rural Affairs, Collaborative Innovation Center for Efficient and Green Production of Agriculture in Mountainous Areas of Zhejiang Province, College of Horticulture Science, Zhejiang A&F University, Hangzhou, Zhejiang, China

Anthocyanins are flavonoid-derived metabolites that contribute to plant and human health. At present, few studies have studied the biosynthesis and accumulation mechanism of anthocyanins in pepper leaves. The role of CaMYBA–CaMYC–CaTTG1 complex in anthocyanin biosynthesis in pepper leaves was studied. Yeast two-hybrid and dual-luciferase experiments showed that CaMYBA, CaMYC, and CaTTG1 could form an MYB–bHLH–WD40 (MBW) complex. They also have transcriptional activation on the anthocyanin synthesis structural genes *CaCHS*, *CaCHI*, *CaF3H*, *CaF3'5'H*, *CaANS*, *CaDFR*, and *CaUFGT*. Silencing *CaMYBA* or *CaMYC* could decrease the content of anthocyanin in pepper leaves. Transient overexpression of *CaMYBA* in tobacco indicated that CaMYBA determines the function of an MBW complex. Further analysis showed that CaMYBA could activate the expression of *CaMYC* by binding to its promoter. Overall, our study expands the understanding of the regulatory mechanism of anthocyanin synthesis in pepper leaves and has important significance for creating more pepper plants with different color patterns by gene editing engineering.

KEYWORDS

pepper, anthocyanin, MBW complex, regulatory complex, transcriptional regulation

1 Introduction

Anthocyanins are one of the important factors that produce a series of colors in different tissue parts of plants such as leaves, flowers, and fruits. Anthocyanins are a major branch of flavonoid metabolism whose biological functions are diverse and play important roles in plant metabolism and breeding, including attracting insects and birds to pollinate and spread seeds and preventing photooxidative damage (Feild et al., 2001; Liu et al., 2020; Winkel-Shirley, 2001). In recent years, researchers have paid more and more attention to anthocyanin synthesis and its regulatory mechanism, which has been well studied in petunia (Albert et al., 2014), snapdragons (Naing et al., 2017), maize (Chachar et al., 2024), and *Arabidopsis* (Gonzalez et al., 2008).

The anthocyanin biosynthetic pathway branches off from the general phenylpropanoid pathway (Borovsky et al., 2004), and its synthesis requires two categories of genes: those encoding the enzymes that catalyze different reactions (i.e., structural genes) and those that regulate the expression of the structural genes (i.e., regulatory genes) (Gonzali et al., 2009). All structural genes in the pathway, including chalcone synthase (CHS), chalcone isomerase (CHl), flavanone 3-hydroxylase (F3H), flavonoid 3′5′-hydroxylase (F3′5′H), dihydroflavonol-4-reductase (DFR), anthocyanin synthase (ANS), and UDP-glucose:flavonoid 3-glucosyltransferase (UFGT), have been identified in some crops (Lalusin et al., 2006; Sunil and Shetty, 2022). *CHS*, *CHI*, and *F3H* are early biosynthetic genes (EBGs), while *F3′5′H*, *DFR*, *ANS*, and *UFGT* are late biosynthetic genes (LBGs) (Jung et al., 2019).

These structural genes usually are regulated by multiple transcription factors, especially MYB, bHLH, and WD40 proteins (Petroni and Tonelli, 2011). The three transcription factors can form a regulatory complex MYB–bHLH–WD40 (MBW) and interact with the promoters of the structural genes of the anthocyanin biosynthesis pathway to regulate their expression (Petroni and Tonelli, 2011; Spelt et al., 2000). For example, the AtMYB113/AtMYB114/AtPAP1/AtPAP2-AtEGL3/AtGL3-AtTTG1 complex can control the accumulation of anthocyanin in *Arabidopsis*, overexpression of AtMYB113 or AtMYB114 results in substantial increases in pigment production similar to the result of overexpression of AtPAP1 or AtPAP2, and pigment production in these overexpressors remains TTG1- and bHLH-dependent (Gonzalez et al., 2008). In *Petunia*, the MYB–bHLH–WD40 complex regulates anthocyanin biosynthesis at different tissue sites, and different MYBs determine the tissue-specific accumulation of anthocyanins (Albert et al., 2011). The patterning and spatial localization of anthocyanins are primarily determined by the activity of the R2R3-MYB factors in the complex, with individual gene-family members regulating separate patterns (Davies et al., 2012), which act with common bHLH and WD40 factors (Schwinn et al., 2006; Gonzalez et al., 2008; Albert et al., 2011; Lowry et al., 2012).

Pepper (*Capsicum annuum* L.) is one of the most important horticultural crops due to its culinary and ornamental applications (Liu et al., 2020). Anthocyanin-rich pepper cultivars can become special peppers with high antioxidant activity (Sharma et al., 2016). Anthocyanin pigmentation in *C. annuum* is influenced by the locus *A*, which encodes a MYB transcription factor (CaMYBA). The expression of EBGs (*CaCHS* and *CaCHI*) and LBGs (*CaDFR* and *CaANS*) of the anthocyanin pathway in pepper has been proposed to be *A*-independent (Borovsky et al., 2004). The insertion of a non-long terminal repeat (non-LTR) retrotransposon in the *CaMYBA* promoter region causes it to recruit transcription factors to activate *CaMYBA* expression, resulting in purple pigmentation in various tissues including fruits (only at the immature stage), flowers, and leaves (Jung et al., 2019). However, tissue-specific anthocyanin pigmentation is still present in some capsicum materials that carry the non-functional *CaMYBA* allele, especially in flowers and fruits. *CaAN3* can induce capsicum fruit-specific anthocyaninosis (Byun et al., 2022). The mutation of WD40 transcription factor CaTTG1 can lead to the disappearance of the hypocotyl anthocyanin phenotype of pepper (Wang et al., 2025). A recessive gene *ayw*, which encodes F3′5′H, was identified as the major candidate gene influencing the yellow color of the anthers and the green color of the stems after preliminary and fine mapping (Wang et al., 2023). Previously, virus-induced gene silencing (VIGS) of the *CaMYBA* gene induced the downregulation of most anthocyanins that synthesize structural genes (Aguilar-Barragán and Ochoa-Alejo, 2014; Zhang et al., 2015) and bHLH gene *CaMYC* (Zhang et al., 2015). However, VIGS of *CaMYC* also reduced the expression levels of most anthocyanins that synthesize structural genes, but not *CaMYBA* (Lu et al., 2019). VIGS of *CaTTG1* reduced the expression levels of some anthocyanins that synthesize structural genes (Aguilar-Barragán and Ochoa-Alejo, 2014). Although some studies have reported that *CaMYBA*, *CaMYC*, and *CaTTG1* are highly correlated with anthocyanin synthesis in pepper, they may form an MBW complex to regulate the transcriptional expression of anthocyanin synthesis structural genes. However, whether they can form an MBW complex or their specific molecular mechanism of regulating anthocyanin synthesis has not been deeply studied. This study verified that *CaMYBA*, *CaMYC*, and *CaTTG1* can form an MBW complex and then regulate the transcriptional expression of *CaCHS*, *CaCHI*, *CaF3H*, *CaF3′5′H*, *CaDFR*, *CaANS*, and *CaUFGT*, thus affecting anthocyanin synthesis in pepper leaves.

2 Materials and methods

2.1 Plant material

Green pepper (Zunla) and purple pepper (PP) were used as experimental materials. The whole plant of PP material was purple. Its stems, leaves, and flowers were purple, and its young fruit was purple, which gradually changed to orange after the fruit ripened. The materials treated in the same phase were randomly selected and divided into three groups as three biological replicates. Mature leaves of different colors were taken, immediately frozen in liquid nitrogen, and stored at −80°C for further analysis.

2.2 Virus-induced gene silencing of *CaMYBA* and *CaMYC*

According to the method of Wang et al. (2013a), gene-specific primers with restriction enzyme cleavage sites were used to amplify *CaMYBA* and *CaMYC* coding region fragments (Supplementary Table S1). The obtained product was inserted into the pTRV2 carrier to generate the pTRV2:CaMYBA and pTRV2:CaMYC. pTRV1, pTRV2, pTRV2:CaMYBA, and pTRV2:CaMYC vectors were transformed into *Agrobacterium tumefaciens* strain (GV3101). GV3101 carrying pTRV1 was mixed with the empty vector pTRV2:00, pTRV2:CaMYBA, and pTRV2:CaMYC in a 1:1 ratio. When the first set of true leaves fully unfolded, the *Agrobacterium* suspension containing pTRV1, pTRV2:00, pTRV2:CaMYBA, and pTRV2:CaMYC (OD600 = 1.0) was soaked into the fully developed cotyledons of PP using a 1.0-mL sterile needle-free syringe. The *Agrobacterium*-inoculated pepper plants were grown for 54 hours under conditions of 16°C, 75% relative humidity, and darkness and then transferred to a growth chamber at 22°C, 60% relative humidity, and 16 hours of light/8 hours of dark light cycle.

2.3 Yeast two-hybrid assay

Yeast two-hybrid (Y2H) assay was conducted according to the study of Wei et al. (2020). For yeast two-hybrid assay, the Coding sequences (CDSs) of CaMYBA (NM_001324618.1), CaMYC (XM_016686645.2), and CaTTG1 (XM_016708729.2) were cloned into the pGADT7 and pGBKT7 vectors to generate the prey and bait plasmids (Supplementary Table S1). The recombinant plasmids were co-transformed into the Y2H GOLD strain and grown on SD/-Leu/-Trp medium at 30°C for 3–4 days. Then, several dilutions of transformants were transferred to SD/-Leu/-Trp/-His/-Ade with X-α-Gal.

2.4 Luciferase complementation assay

According to the method of Luo et al. (2020), the full lengths of *CaMYBA*, *CaMYC*, and *CaTTG1* were inserted into the pCAMBIA1300-nLUC and pCAMBIA1300-cLUC vectors (Supplementary Table S1). The constructs and empty plasmids were transformed into GV3101 and transiently expressed in *Nicotiana benthamiana* leaves. The LUC fluorescence signal was observed through a Tanon 5200 imaging system (Tanon, Shanghai, China).

2.5 Bimolecular fluorescence complementation assay

Full-length coding sequences of CaMYBA, CaMYC, and CaTTG1 were cloned into the binary N-terminal fragment of yellow fluorescent protein (nYFP) and C-terminal fragment of yellow fluorescent protein (cYFP) vectors (Supplementary Table S1). According to the method of Wang et al. (2021), *Agrobacterium* strains transformed with indicated nYFP or cYFP vectors were incubated, harvested, and resuspended in infiltration buffer (0.2 mM acetosyringone, 10 mM $MgCl_2$, and 10 mM MES, pH 5.6) to identical concentrations (OD600 = 0.8). Equal volumes of an *Agrobacterium* culture containing nYFP (OD600 = 0.8) and cYFP (OD600 = 0.8) were mixed before infiltration into *N. benthamiana* leaves. After infiltration, plants were incubated at 24°C for 48 hours before observation. A confocal laser scanning microscope was used to identify yellow fluorescent protein (YFP) and 4',6-diamidino-2-phenylindole (DAPI) fluorescent signals.

2.6 Dual-luciferase assay

The CDSs of CaMYBA, CaMYC, and CaTTG1 were amplified and ligated into the pGreenII-62-SK vector for the generation of effector constructs (Supplementary Table S1). The *CaCHS*, *CaCHI*, *CaF3H*, *CaF3'5'H*, *CaDFR*, *CaANS*, and *CaUFGT* promoter fragments were ligated into the pGreenII-0800-LUC vector as reporters (Supplementary Tables S1 and S2). According to the method of Luo et al. (2020), the effector and reporter constructs were introduced into *A. tumefaciens* strain GV3101 (pSoup-p19) and transiently expressed in 4-week-old tobacco leaves as previously described. The luminescent living image was obtained using the plant living imaging system after 48 hours of incubation at 24°C following infiltration. The real-time fluorescence technique was employed for the quantitative assessment of LUC and REN expression, as well as for the calculation of their ratio to determine transcriptional activity, and the control ratio was set to 1.

2.7 Determination of anthocyanins

The content of anthocyanins was determined following the methods described in a previous study (Zhang et al., 2020). The samples were ground into powder with liquid nitrogen, followed by 24 hours of extraction in 1% HCl–methyl alcohol (3 mL, v/v) under 4°C in the dark. The absorbance at 530 nm and 657 nm was determined, and estimation of anthocyanin concentration was conducted based on the following formula: Q Anthocyanins = $(A_{530} - 0.25 * A_{657})$/fresh weight.

2.8 RNA extraction and real-time quantitative RT-PCR

Total RNA (1 μg), extracted using TRIzol reagent (Sangon Biotech, Shanghai, China), was used for oligo (dT) 18-primed cDNA synthesis according to the reverse transcription protocol (Vazyme, Nanjing, China). The resulting cDNA was subjected to real-time quantitative RT-PCR using a SYBR Premix Ex Taq kit (Vazyme) on a qTOWER3 real-time PCR machine. All quantitative primer information is shown in Supplementary Table S1. For each reported result, at least three independent biological samples were subjected to a minimum of three technical replicates. The results were normalized using the internal control *CaUBI-3*.

2.9 Statistical analysis

SPSS program version 19 (United States) was used to analyze the data. The one-way analysis of variance and *t*-test were used to determine the significant difference between groups.

3 Results

3.1 CaMYBA, CaMYC, and CaTTG1 form an MBW complex

Although the transcription factors CaMYBA, CaMYC, and CaTTG1 are expressed in the leaf of *C. annuum* L (Zhang et al., 2015; Lu et al., 2019), whether or not they regulate the anthocyanin pathway in pepper leaf as a complex remains to be confirmed. We compared CaMYBA and CaMYC with MYB and bHLH proteins that have been reported to regulate anthocyanin synthesis in other species. Similar to MYB proteins that regulate anthocyanin synthesis in other species, CaMYBA contains a highly conserved R2R3 domain in the N-terminal region, a bHLH-interaction motif and an ANDV motif in the R3 domain region, and a KPRPR[S/T]F motif in the C-terminal region (Supplementary Figure S3; Stracke et al., 2001; Yan et al., 2020). Similarly, CaMYC, like other bHLH proteins that regulate anthocyanin synthesis, has a MYB-interaction motif in the N-terminal and a conserved bHLH domain in the C-terminal (Supplementary Figure S4; Wang et al., 2019). CaTTG1 is highly similar to TTG1 protein in tomato with a typical WD40 protein domain, and it is also similar to TTG1 protein regulating anthocyanin synthesis in *Arabidopsis* and *Petunia* (Gonzalez et al., 2008; Albert et al., 2011; Wang et al., 2025).

To verify whether CaMYBA, CaMYC, and CaTTG1 can form a complex, we validated their pairwise interactions by Y2H assay. The results showed that only the yeast cells transformed with pGADT7-CaMYBA and pGBKT7-CaMYC, and pGADT7-CaTTG1 and pGBKT7-CaMYC could grow well and show GAL4 activity on the X-α-Gal-contained selective media (Figure 1A). In order to further verify the reliability of the interaction results of the Y2H assay, we conducted luciferase complementation assay (LCA) and bimolecular fluorescence complementation (BiFC) assay. The results of the LCA and BiFC assay confirmed the interaction results of the Y2H assay. Accordingly, CaMYC can interact with CaMYBA or CaTTG1 proteins, while CaMYBA cannot interact with CaTTG1 proteins, and they can form a CaMYBA–CaMYC–CaTTG1 MBW complex.

3.2 CaMYBA, CaMYC, and CaTTG1 synergistically regulate the anthocyanin pathway genes

To verify the regulatory effects of CaMYBA, CaMYC, and CaTTG1 on the transcriptional activation of structural genes in the anthocyanin synthesis pathway of pepper leaves, we collected their 5′-non-coding sequences upstream of the translation initiation site (Supplementary Table S2) to evaluate CaMYBA, CaMYC, and CaTTG1 regulatory capacity on the expression of each. The possible MYB-recognizing element (MRE) or bHLH-recognizing element (BRE) was predicted (Zhu et al., 2015), and it was found that there was at least one MRE and one BRE on the promoters of the seven anthocyanin synthesis pathway genes: *CaCHS*, *CaCHI*, *CaF3H*, *CaF3′5′H*, *CaDFR*, *CaANS*, and *CaUFGT* (Supplementary Figure S5). By single and combinatory tests of CaMYBA, CaMYC, and CaTTG1 as effectors in dual-luciferase assays (Figure 2A), we detected how they regulated the reporter gene via the 5′-non-coding regions of the pathway genes. The tests showed that CaMYBA could, alone, initiate transcription on the seven pathway gene promoters. By contrast, CaMYC alone failed to activate observable promoter activities in the same setting. When CaMYBA and CaMYC are together or when all three of them are together, greater promoter activity can be generated, and the presence of CaMYBA, CaMYC, and CaTTG1 at the same time produces the largest promoter activity (Figure 2B, Supplementary Figure S6). This suggests that CaMYBA, CaMYC, and CaTTG1 acted as an MBW complex in the transcriptional activation of the main anthocyanin pathway genes in pepper leaves, and *CaMYBA* should play a role as the main active factor.

3.3 Silencing of *CaMYBA* or *CaMYC* reduced the accumulation of anthocyanins and the expression of anthocyanin pathway genes

To verify the role of *CaMYBA* or *CaMYC* in anthocyanin synthesis in pepper leaves, *CaMYBA* or *CaMYC* expression was silenced through the VIGS technique, in which the *Agrobacterium* strain harboring pTRV1 and pTRV2:00, pTRV2:CaMYBA, and pTRV2:CaMYC were injected into cotyledons of purple pepper line PP. After 5 weeks of injection, obvious green leaves were observed in the silenced peppers with the pTRV2:CaMYBA or pTRV2:CaMYC vector. After silencing *CaMYBA* or *CaMYC*, the purple pigment in pepper leaves obviously faded. However, no obvious symptoms were observed in pepper seedlings with pTRV2:00 (Figures 3A–D). The high silencing efficiency of *CaMYBA* or *CaMYC* gene expression after 5 weeks of injection suggested the reliability of the VIGS technique in silencing pepper gene expression (Figure 3E).

Real-time quantitative PCR analyses demonstrated that *CaMYBA* or *CaMYC* silencing did not impact *CaTTG1* and *CaCHI* expression. After the silencing of *CaMYBA*, the expression of *CaMYC* decreased significantly by as much as 99%. Similarly, after the silencing of *CaMYC*, the expression of *CaMYBA* also decreased to a certain extent by approximately 48%. The expression of *CaCHS*, *CaF3H*, *CaF3′5′H*, *CaDFR*, *CaANS*, and *CaUFGT* decreased significantly after *CaMYBA* silencing or *CaMYC* silencing was conducted due to low expression observed in silenced pepper leaves compared with the negative control. In contrast, the expression of these six genes decreased even more after the silencing of *CaMYBA* (Figure 3E).

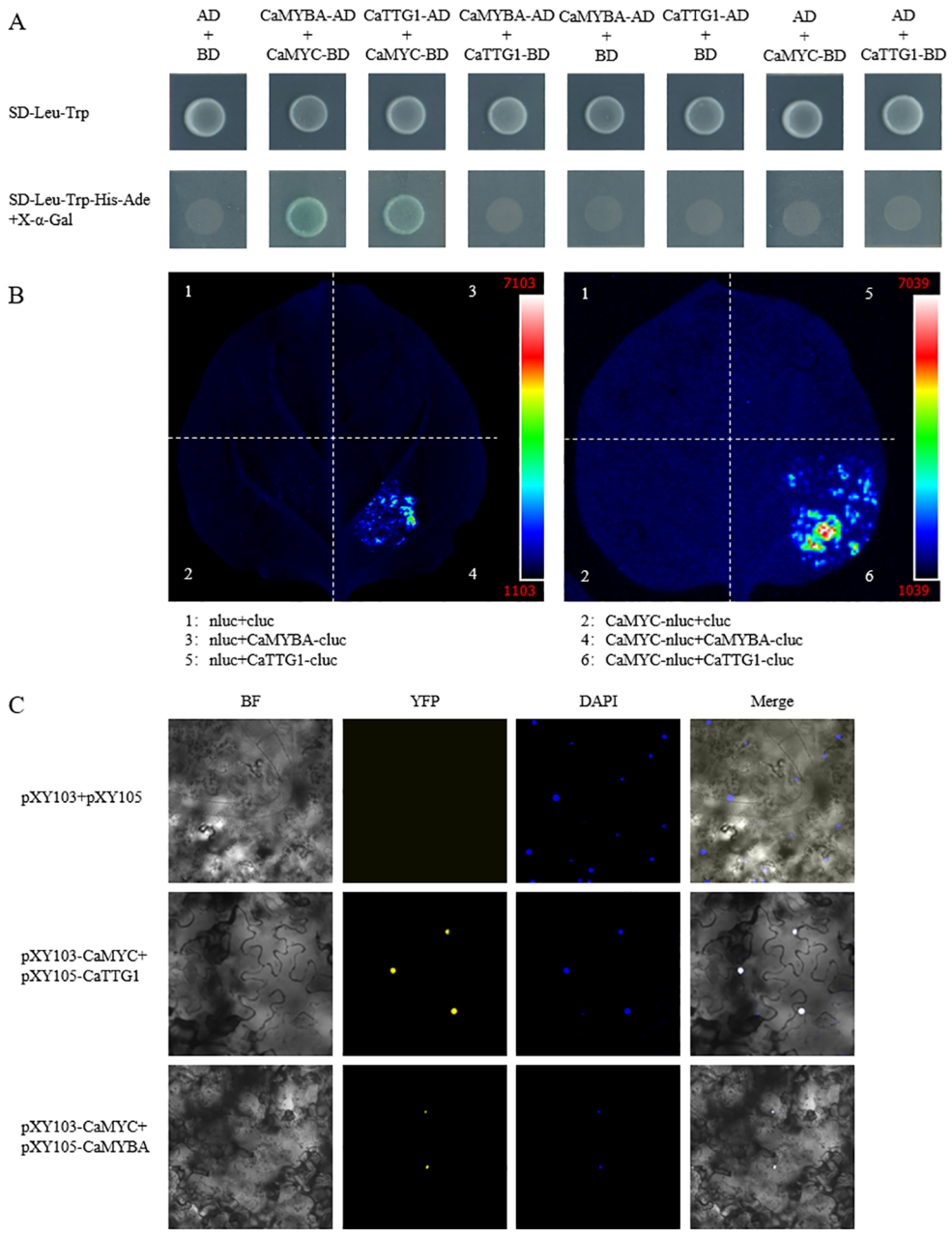


FIGURE 1
Interaction between *CaMYBA* and *CaMYC*, *CaMYC*, and CaTTG1. **(A)** Interaction of *CaMYBA* and *CaMYC*, *CaMYC*, and CaTTG1 in yeast two-hybrid (Y2H) assay. Transformed yeast cells were grown on SD-Leu-Trp and SD-Leu-Trp-His-Ade added with X-α-Gal. Transformation of pGADT7 (AD) and pGBKT7 (BD) vectors was used as a negative control. **(B)** *CaMYC* interacts with *CaMYBA* and CaTTG1 in the luciferase complementation assay (LCA). Positive luminescence observed by Charge Coupled Device (CCD) camera indicates mutual interaction. **(C)** Physical interactions of *CaMYBA* and *CaMYC*, *CaMYC*, and CaTTG1 by bimolecular fluorescence complementation (BiFC) assay. Transformation of pXY103 and pXY105 vectors was used as a negative control. Microscopic images were taken under bright field and fluorescence. Images under bright field (left one), YFP (left two), and DAPI (right two); the merged images (right one) are shown on the right. BF, bright field; YFP, yellow fluorescent protein.

3.4 *CaMYBA* in the MBW complex leads to activation of anthocyanin synthesis

To confirm the role played by *CaMYBA*, *CaMYC*, and *CaTTG1* in the synthesis of anthocyanin, *Agrobacterium* carrying 35S:00, 35S:CaMYBA, 35S:CaMYC, or 35S:CaTTG1 was injected into tobacco leaves alone or in combination. Compared with the negative control, there were obvious color changes in the leaf regions with combined infection containing *CaMYBA* but no obvious color changes in the leaf regions without combined infection containing *CaMYBA* (Figure 4A). Consistent with observed phenotypic changes, total anthocyanin levels in tobacco leaf regions without *CaMYBA* overexpression were similar to those in negative controls, with no significant changes. In contrast, anthocyanin accumulation was much higher in leaf regions with *CaMYBA* overexpression (Figure 4B). In addition, no *NtMYBA* expression was detected in all overexpression

10.3389/fpls.2025.1538607

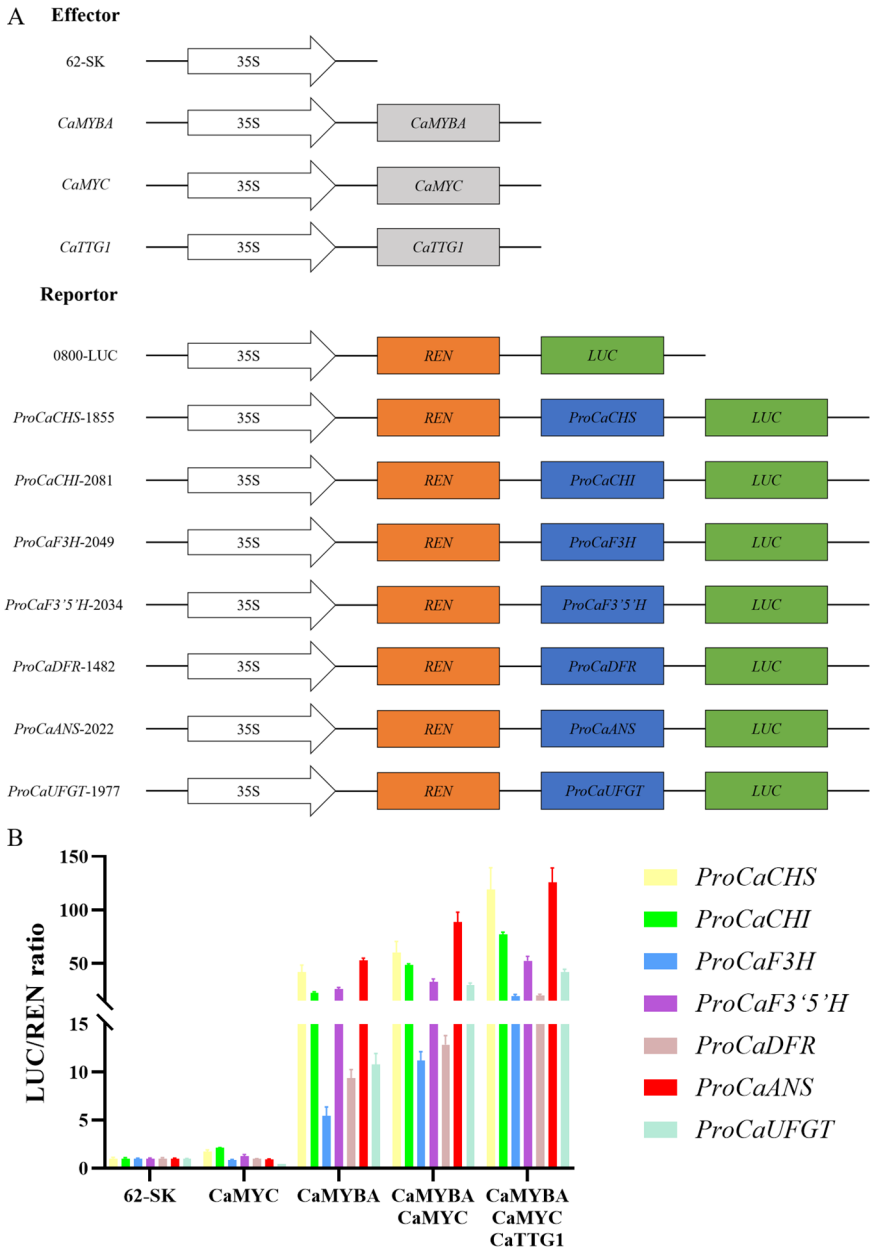


FIGURE 2
Collaborative regulation of CaMYBA, CaMYC, and CaTTG1 on the anthocyanin pathway genes of *Capsicum annuum* L. **(A)** Construct details for dual-luciferase assays. The effector constructs contain CaMYBA, CaMYC, and CaTTG1 driven by the CaMV 35 S promoter. The reporter constructs contain the firefly luciferase (LUC) driven by the promoter of *CaCHS*, *CaCHI*, *CaF3H*, *CaF3'5'H*, *CaDFR*, *CaANS*, or *CaUFGT*, and the Renilla luciferase (REN) driven by the CaMV 35 S promoter. **(B)** The effects of CaMYBA, CaMYC, and CaTTG1 individually and in combination on the promoter activity of *CaCHS*, *CaCHI*, *CaF3H*, *CaF3'5'H*, *CaDFR*, *CaANS*, or *CaUFGT* with the luciferase reporter assay. The 5′-non-coding regions are shown using colored bars for the seven genes tested. Empty effector vector (62-SK) was used as the control. LUC/REN ratio of the control (tobacco leaves co-transformed with the reporters and the empty effector vector) was taken as 1 for normalization. Error bars represent the mean ± SD of three biological replicates. Statistical significance was determined using Duncan's t-test ($p < 0.05$).

combinations, so the effect of *NtMYBA* in tobacco leaves on anthocyanin accumulation could be excluded. The expression of *CaMYBA* was higher in tobacco leaf regions with color changes (Figure 4C). Therefore, ectopic transient overexpression of *CaMYBA* in tobacco leaves induces anthocyanin accumulation.

The effects of single or combined overexpression of *CaMYBA*, *CaMYC*, and *CaTTG1* on anthocyanin biosynthesis gene expression in tobacco leaves were analyzed by qRT-PCR. The CaMYBA-overexpressing, CaMYBA-CaMYC-overexpressing, and CaMYBA–CaMYC–CaTTG1-overexpressing regions presented remarkably higher expressions of *NtMYC*, *NtCHS*, *NtCHI*, *NtF3H*, *NtF3'5'H*, *NtDFR*, *NtANS*, and *NtUFGT* relative to negative control region and CaTTG1-overexpressing region, and this consistent with the greater content of anthocyanin in CaMYBA-overexpressing, CaMYBA-CaMYC-overexpressing, and CaMYBA–CaMYC–CaTTG1-overexpressing regions (Figure 4C). The expression levels of related genes did not change significantly in CaMYC-overexpressing, CaTTG1-overexpressing, and CaMYC-CaTTG1-overexpressing

10.3389/fpls.2025.1538607

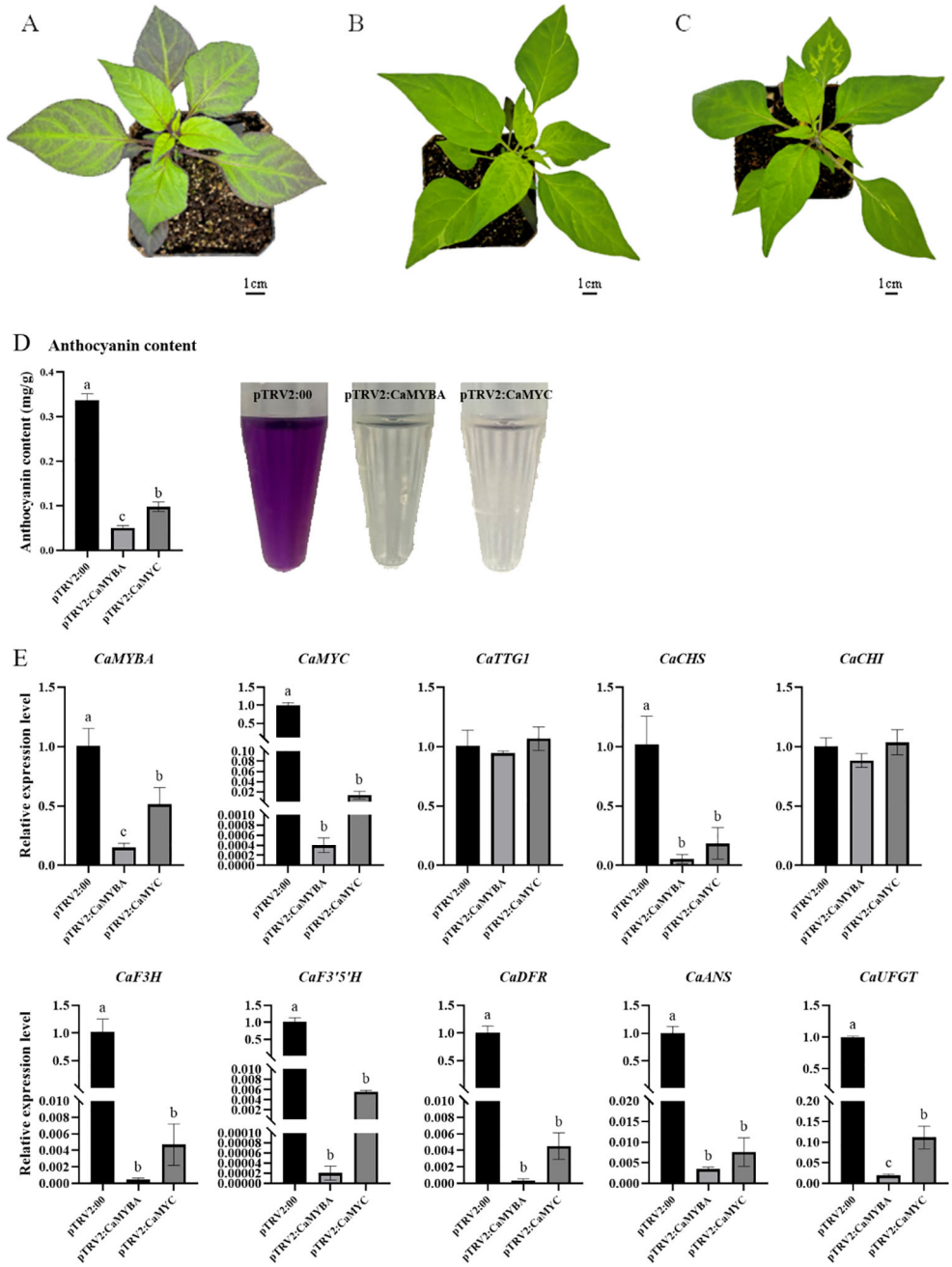


FIGURE 3
TRV-mediated silencing of CaMYBA or CaMYC in purple pepper plants (line PP). **(A)** pTRV2 empty vector infected pepper plant. **(B)** CaMYBA-silencing pepper plant. **(C)** CaMYC-silencing pepper plant. **(D)** Anthocyanin content in silencing plants. **(E)** Expression of anthocyanin synthesis pathway genes in silencing pepper plants. pTRV2:00, negative control plants; pTRV2:CaMYBA, CaMYBA-silencing plants; pTRV2:CaMYC, CaMYC-silencing plants. The expression of all genes was normalized by that of ubiquitin-conjugating protein gene *CaUBI-3*. The experiment was conducted with three biological replicates. Error bars represent the mean ± SD of three biological replicates. Statistical significance was determined using Duncan's t-test ($p < 0.05$). Lower case letters indicate significant differences at the $p<0.05$ level.

regions. Although the expression levels of *NtF3'5'H* in CaMYC-overexpressing and CaMYC-CaTTG1-overexpressing regions were higher than those in negative control, statistical analysis showed no significant difference (Figure 4C). Overall, *CaMYBA* dominated anthocyanin accumulation, and it may independently activate the expression of *NtMYC* in tobacco leaves and form a complex with it to further regulate anthocyanin synthesis.

3.5 CaMYBA binds to the *CaMYC* promoter and activates its transcription

To verify the regulatory effects of *CaMYBA* on the transcriptional activation of CaMYC, we collected 5′-non-coding sequences of *CaMYC* upstream of the translation initiation site (Supplementary Table S2) to evaluate CaMYBA regulatory capacity on their

10.3389/fpls.2025.1538607

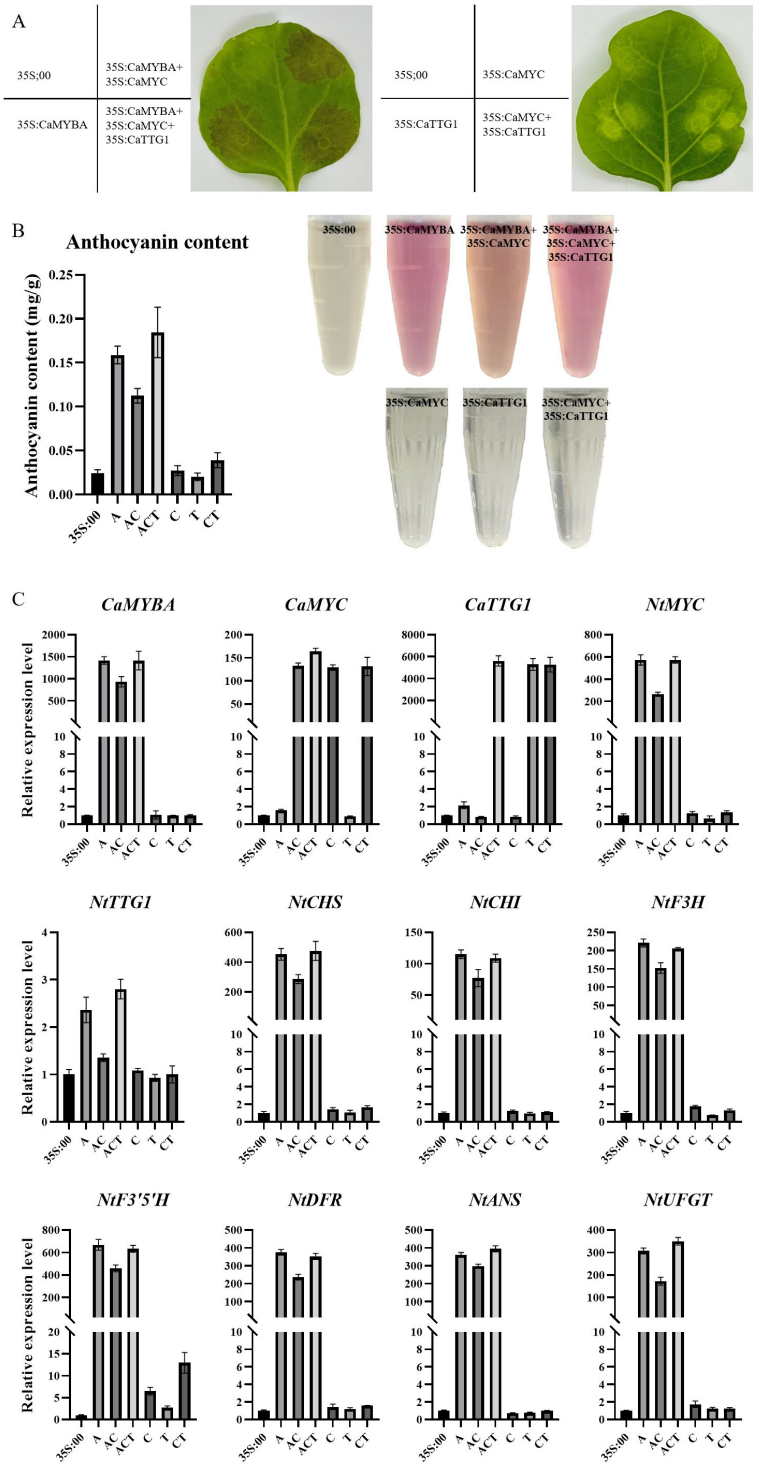


FIGURE 4
Transient overexpression phenotype in tobacco leaves. **(A)** Pigmentation of transiently overexpressed tobacco leaves in different combinations. **(B)** Determination of anthocyanin content in transiently overexpressed tobacco leaves in different combinations. 35S:00, negative control; A, 35S:CaMYBA; AC, 35S:CaMYBA+35S:CaMYC; ACT, 35S:CaMYBA+35S:CaMYC+35S:CaTTG1; C, 35S:CaMYC; T, 35S:CaTTG1; CT, 35S:CaMYC+CaTTG1. **(C)** qRT-PCR analysis of anthocyanin-related gene expression between tobacco leaves with different combinations of transient overexpression. *NtEF1α* was used as an internal control gene. Error bars represent the mean ± SD of three biological replicates. Statistical significance was determined using Duncan's t-test ($p < 0.05$).

expression (Figure 5A). Dual-luciferase assay showed that *CaMYBA* could initiate transcription on *CaMYC* promoters (Figure 5B). At the same time, we also performed transient overexpression of CaMYBA in pepper leaves. Although the phenotype change was not so obvious as in tobacco leaves, qRT-PCR data showed that the expression level of *CaMYBA* was significantly increased, and the expression levels of *CaMYC*, *CaCHS*, *CaCHI*, *CaANS*, and *CaUFGT* were also significantly increased (Supplementary Figure S7). These results all

indicate that *CaMYBA* can activate the transcriptional expression of *CaMYC*.

4 Discussion

Chili pepper (*C. annuum* L.) originated in the tropics of South America, belongs to the Solanaceae family, and has high economic values (Wang et al., 2022). It has been reported that purple pepper leaves can enhance photosynthesis and alleviate oxidative stress (Dewez and Perreault, 2013; Zhang et al., 2015). In addition, pepper can also be used as an ornamental crop, and the purple pigment increases its ornamental value. Therefore, it is of great significance to expand the understanding of the regulatory mechanism of anthocyanin biosynthesis in pepper leaves.

In *Arabidopsis* and *Petunia*, MYB, bHLH, and WD40 transcription factors form an MBW complex to regulate anthocyanin biosynthesis (Gonzalez et al., 2008; Albert et al., 2011). According to previous reports, *CaMYBA*, *CaMYC*, and *CaTTG1* have been shown to be highly correlated with anthocyanin accumulation in pepper, and they may regulate anthocyanin biosynthesis in pepper (Borovsky et al., 2004; Aguilar-Barragán and Ochoa-Alejo, 2014; Zhang et al., 2015; Lu et al., 2019; Jung et al., 2019; Wang et al., 2025; Liu et al., 2024). However, whether they can form an MBW complex has not been reported. In this study, amino acid sequence analysis showed that CaMYBA and CaMYC were highly similar in structure to MYB and bHLH transcription factors regulating anthocyanin synthesis in other species (Supplementary Figures S1 and S2); CaMYBA and CaMYC should be the corresponding members of the MBW complex regulating anthocyanin synthesis in capsicum. Meanwhile, protein interaction experiments showed that CaMYC could interact with CaMYBA or CaTTG1 (Figures 1A–C), and it was indicated that CaMYBA, CaMYC, and CaTTG1 can also form an MBW complex in pepper.

The binding of MYB and bHLH to promoters of anthocyanin structural genes was first reported with maize C1 and B (Roth et al., 1991). For the MYB part, C1 could bind to variable sites in the maize a1 gene promoter (Sainz et al., 1997). For the bHLH part, after the reported binding of CG-1 protein (Staiger et al., 1991) and human c-MYC (Blackwell et al., 1990) to CACGTG, the G-box was shown to bind to maize R (Kong et al., 2012), petunia AN1, and *Ipomoea* bHLH2 (Wang et al., 2013b). Most pathway genes identified so far do contain the *cis*-regulatory region with the necessary footings for the MBW complex; also, a 7-bp MRE (ANCNNCC) and a 6-bp bHLH-recognizing element [BRE and CACN(A/C/T)(G/T)] are required for an MBW complex to activate the promoter of target genes in the anthocyanin synthesis pathway (Zhu et al., 2015). After comparison, there was at least one MRE and

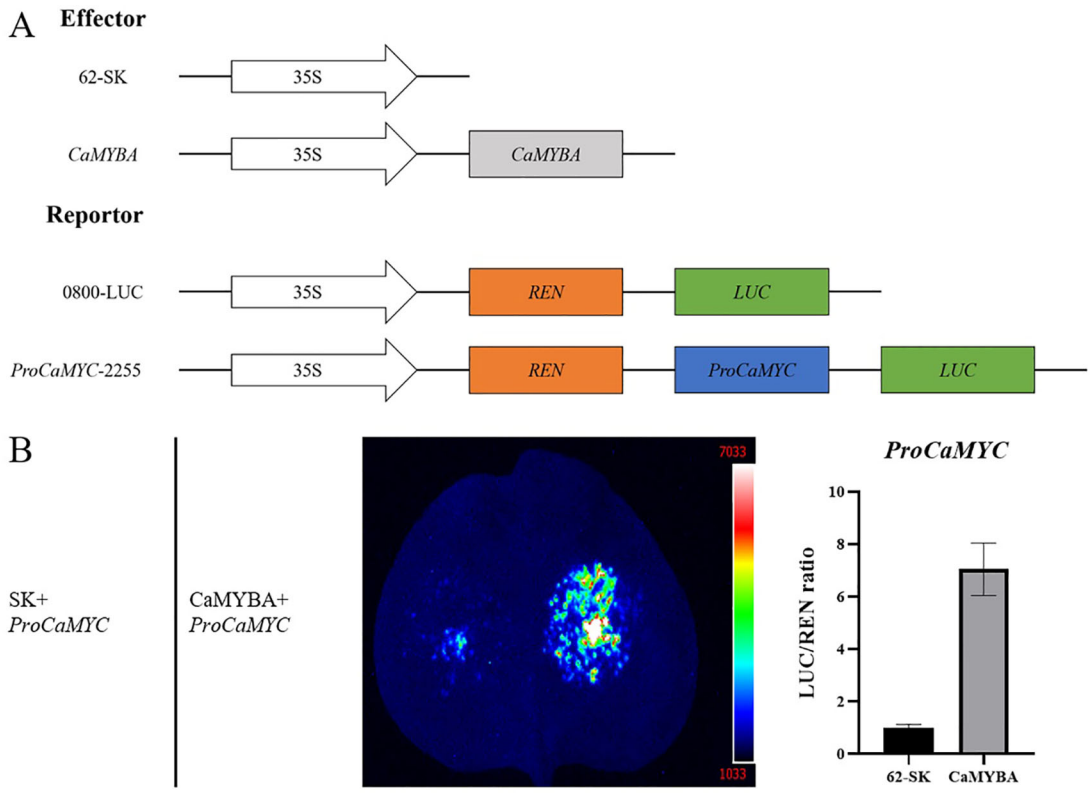


FIGURE 5
Regulation of CaMYBA on the *CaMYC*. **(A)** Construct details for dual-luciferase assays. The effector constructs contain CaMYBA driven by the CaMV35S promoter. The reporter constructs contain the firefly luciferase (LUC) driven by the promoter of *CaMYC* and the Renilla luciferase (REN) driven by the CaMV35S promoter. **(B)** Dual-luciferase detection experiments showed that CaMYBA promoted the expression of the *CaMYC*. Empty effector vector (62-SK) was used as the control. LUC/REN ratio of the control (tobacco leaves co-transformed with the reporters and the empty effector vector) was taken as 1 for normalization. Error bars represent the mean ± SD of three biological replicates. Statistical significance was determined using a t-test, $p < 0.01$ (**).

one BRE on the promoter of *CaCHS*, *CaCHI*, *CaF3H*, *CaF3′5′H*, *CaDFR*, *CaANS*, and *CaUFGT* (Supplementary Figure S5). By single and combinatory tests of CaMYBA, CaMYC, and CaTTG1 as effectors in dual-luciferase assays (Figure 2A), CaMYBA, CaMYC, and CaTTG1 acted as an MBW complex in the transcriptional activation of the main anthocyanin pathway genes in pepper leaves (Figure 2B). Studies have shown that the activity of MYB transcription factors determines the function and regulatory mode of the MBW complex, with individual gene-family members regulating separate patterns (Davies et al., 2012), which act with common bHLH and WD40 transcription factors (Schwinn et al., 2006; Gonzalez et al., 2008; Albert et al., 2011; Lowry et al., 2012). Therefore, we believe that CaMYBA also determines the specific functional mode of an MBW complex regulating anthocyanin synthesis in pepper leaves. After VIGS silencing *CaMYBA* and *CaMYC* in our purple pepper variety PP, the leaves changed significantly from purple to green (Figures 3A–C), and the expression levels of anthocyanin synthesis structural genes (such as *CaCHS*) were significantly decreased except for *CaCHI*. In the study of Zhang et al. and Lu et al., the expression levels of all anthocyanin synthesis structure genes such as *CaCHS* were significantly decreased after silencing *CaMYBA* or *CaMYC* in pepper (Zhang et al., 2015; Lu et al., 2019). The difference is that the silencing of *CaMYC* resulted in a certain degree of decline in the expression of *CaMYBA* (Figure 3B). The VIGS experiment showed that *CaMYBA* and *CaMYC* have previously been shown to play a crucial role in the process of anthocyanin accumulation in pepper leaves. The transient overexpression of *CaMYBA*, *CaMYC*, and *CaTTG1* alone or in combination in tobacco leaves produced obvious anthocyanin accumulation only when *CaMYBA* was present (Figures 4A, B). Overall, CaMYBA determined the function and regulatory mode of CaMYBA–CaMYC–CaTTG1 complex to activate anthocyanin synthesis in pepper leaves.

In *Arabidopsis*, the bHLH transcription factor TRANSPARENT TESTA8 (AtTT8) regulates its own expression through an MBW complex, which ultimately contributes to the regulation of anthocyanin and proanthocyanidin (PA) synthesis (Baudry et al., 2006; Xu et al., 2013). In *Petunia*, ectopic expression of *PhAN2* (R2R3-MYB activator) in leaves resulted in ectopic expression of the bHLH transcription factor PhAN1 (Spelt et al., 2000). Similarly, *PhAN1* transcript levels are severely reduced in the anthers of petunias that lack a functional *PhAN4* allele (R2R3-MYB activator) (Spelt et al., 2000). After the expression of *CaMYBA* was silenced in pepper leaves, the expression of *CaMYC* was also decreased, and the expression of *NtMYC* was also significantly increased after transient overexpression of *CaMYBA* in tobacco leaves (Figures 3D, 4C). This suggests that *CaMYC* may also be regulated by *CaMYBA* in pepper. In the promoter analysis of *CaMYC*, we also found several possible MRE and BRE sites (Supplementary Figure S5). Dual-luciferase assay also proved that CaMYBA binds to the *CaMYC* promoter and activates its transcription (Figure 5B). Transient overexpression of *CaMYBA* in pepper leaves showed no significant changes in leaf color phenotype, but the expression levels of *CaMYC* and some anthocyanin synthesis structural genes increased significantly (Supplementary Figure S7). Therefore, when combined with studies of bHLH transcription factors such as AtTT8 and PhAN1 in other plants, we speculated that the transcription of CaMYBA activated the expression of *CaMYC* in pepper leaves, and they formed an MBW complex together with CaTTG1 to further transcriptively activate the expression of anthocyanin synthesis structural genes such as *CaANS*, thus promoting the accumulation of anthocyanins in pepper leaves. Our results verified the formation of CaMYBA–CaMYC–CaTTG1 complex in pepper and verified its function and regulation mode in the anthocyanin synthesis of pepper leaves. However, it is not clear whether CaMYBA can stably activate anthocyanin synthesis in the absence of CaMYC. Further studies should focus on elucidating the effect of the presence of CaMYC and CaTTG1 on the functional stability of CaMYBA at the protein level and the validation of MRE and BRE on the promoter of the structural gene of anthocyanin synthesis in pepper. The verification of the CaMYBA–CaMYC–CaTTG1 complex and its function and regulatory mode will help to expand the understanding of the regulatory mechanism of anthocyanin synthesis in pepper leaves, improve the regulatory model of anthocyanin synthesis in pepper, and also help to create germplasm resources of pepper with different color patterns.

5 Conclusion

In sum, CaMYBA, CaMYC, and CaTTG1 can form an MBW complex, and the complex can directly bind promoters of anthocyanin synthesis structural genes such as *CaANS* to promote their transcription and expression, thus promoting anthocyanin accumulation in pepper leaves. We believe that CaMYBA activates *CaMYC* expression and determines the function and regulation mode of the MBW complex.

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

XM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. GL: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. ZX: Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft. CL: Validation, Visualization, Writing – original draft. BZ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Software, 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 financially supported by the Major Science and Technology Project of Plant Breeding in Zhejiang Province (2021C02065).

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.2025.1538607/full#supplementary-material

References

Aguilar-Barragán, A., and Ochoa-Alejo, N. (2014). Virus-induced silencing of MYB and WD40 transcription factor genes affects the accumulation of anthocyanins in chilli pepper fruit. *Biol. Plantarum* 58, 567–574. doi: 10.1007/s10535-014-0427-4

Albert, N. W., Davies, K. M., Lewis, D. H., Zhang, H., Montefiori, M., Brendolise, C., et al. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell* 26, 962–980. doi: 10.1105/tpc.113.122069

Albert, N. W., Lewis, D. H., Zhang, H., Schwinn, K. E., Jameson, P. E., and Davies, K. M. (2011). Members of an R2R3-MYB transcription factor family in Petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant Journal: Cell Mol. Biol.* 65, 771–784. doi: 10.1111/j.1365-313X.2010.04465.x

Baudry, A., Caboche, M., and Lepiniec, L. (2006). TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. *Plant Journal: Cell Mol. Biol.* 46, 768–779. doi: 10.1111/j.1365-313X.2006.02733.x

Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N., and Weintraub, H. (1990). Sequence-specific DNA binding by the c-Myc protein. *Sci. (New York N.Y.)* 250, 1149–1151. doi: 10.1126/science.2251503

Borovsky, Y., Oren-Shamir, M., Ovadia, R., De Jong, W., and Paran, I. (2004). The *A* locus that controls anthocyanin accumulation in pepper encodes a MYB transcription factor homologous to *Anthocyanin2* of Petunia. *TAG. Theor. Appl. Genet. Theoretische und angewandte Genetik* 109, 23–29. doi: 10.1007/s00122-004-1625-9

Byun, J., Kim, T. G., Lee, J. H., Li, N., Jung, S., and Kang, B. C. (2022). Identification of *CaAN3* as a fruit-specific regulator of anthocyanin biosynthesis in pepper (*Capsicum annuum*). *TAG. Theor. Appl. Genet. Theoretische und angewandte Genetik* 135, 2197–2211. doi: 10.1007/s00122-022-04106-y

Chachar, Z., Lai, R., Ahmed, N., Lingling, M., Chachar, S., Paker, N. P., et al. (2024). Cloned genes and genetic regulation of anthocyanin biosynthesis in maize, a comparative review. *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1310634

Davies, K. M., Albert, N. W., and Schwinn, K. E. (2012). From landing lights to mimicry: the molecular regulation of flower colouration and mechanisms for pigmentation patterning. *Funct. Plant Biology: FPB* 39, 619–638. doi: 10.1071/FP12195

Dewez, D., and Perreault, F. (2013). Effect of the anthocyanic epidermal layer on Photosystem II and I energy dissipation processes in Tradescantia pallida (Rose) Hunt. *Acta Physiol. Plant* 35, 463–472. doi: 10.1007/s11738-012-1089-5

Feild, T. S., Lee, D. W., and Holbrook, N. M. (2001). Why leaves turn red in autumn. The role of anthocyanins in senescing leaves of red-osier dogwood. *Plant Physiol.* 127, 566–574. doi: 10.1104/pp.010063

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant journal: Cell Mol. Biol.* 53, 814–827. doi: 10.1111/j.1365-313X.2007.03373.x

Gonzali, S., Mazzucato, A., and Perata, P. (2009). Purple as a tomato: towards high anthocyanin tomatoes. *Trends Plant Sci.* 14, 237–241. doi: 10.1016/j.tplants.2009.02.001

Jung, S., Venkatesh, J., Kang, M. Y., Kwon, J. K., and Kang, B. C. (2019). A non-LTR retrotransposon activates anthocyanin biosynthesis by regulating a MYB transcription factor in *Capsicum annuum*. *Plant Science: Int. J. Exp. Plant Biol.* 287, 110181. doi: 10.1016/j.plantsci.2019.110181

Kong, Q., Pattanaik, S., Feller, A., Werkman, J. R., Chai, C., Wang, Y., et al. (2012). Regulatory switch enforced by basic helix-loop-helix and ACT-domain mediated dimerizations of the maize transcription factor R. *Proc. Natl. Acad. Sci. United States America* 109, E2091–E2097. doi: 10.1073/pnas.1205513109

Lalusin, A. G., Nishita, K., Kim, S. H., Ohta, M., and Fujimura, T. (2006). A new MADS-box gene (*IbMADS10*) from sweet potato (*Ipomoea batatas* (L.) Lam) is involved in the accumulation of anthocyanin. *Mol. Genet. Genomics: MGG* 275, 44–54. doi: 10.1007/s00438-005-0080-x

Liu, J., Ai, X., Wang, Y., Lu, Q., Li, T., Wu, L., et al. (2020). Fine mapping of the *Ca3GT* gene controlling anthocyanin biosynthesis in mature unripe fruit of *Capsicum annuum* L. *TAG. Theor. Appl. Genet. Theoretische und angewandte Genetik* 133, 2729–2742. doi: 10.1007/s00122-020-03628-7

Liu, S., Yang, H., Zhang, H., Liu, J., Ma, S., Hui, H., et al. (2024). Phenotypic, genetic, variation, and molecular function of CaMYB113 in pepper (*Capsicum annuum* L.). *Int. J. Biol. Macromolecules* 281, 136300. doi: 10.1016/j.ijbiomac.2024.136300

Lowry, D. B., Sheng, C. C., Lasky, J. R., and Willis, J. H. (2012). Five anthocyanin polymorphisms are associated with an R2R3-MYB cluster in *Mimulus guttatus* (Phrymaceae). *Am. J. Bot.* 99, 82–91. doi: 10.3732/ajb.1100285

Lu, B. Y., Cheng, G. X., Zhang, Z., Sun, J. T., Ali, M., Jia, Q. L., et al. (2019). *CaMYC*, A novel transcription factor, regulates anthocyanin biosynthesis in color-leaved pepper (*Capsicum annuum* L.). *J. Plant Growth Regul.* 38, 574–585. doi: 10.1007/s00344-018-9871-2

Luo, P., Li, Z., Chen, W., Xing, W., Yang, J., and Cui, Y. (2020). Overexpression of RmICE1, a bHLH transcription factor from Rosa multiflora, enhances cold tolerance via modulating ROS levels and activating the expression of stress-responsive genes. *Environ. Exp. Bot.* 178, 104160. doi: 10.1016/j.envexpbot.2020.104160

Naing, A. H., Park, K. I., Ai, T. N., Chung, M. Y., Han, J. S., Kang, Y. W., et al. (2017). Overexpression of snapdragon *Delila* (*Del*) gene in tobacco enhances anthocyanin accumulation and abiotic stress tolerance. *BMC Plant Biol.* 17, 65. doi: 10.1186/s12870-017-1015-5

Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant science: an Int. J. Exp. Plant Biol.* 181, 219–229. doi: 10.1016/j.plantsci.2011.05.009

Roth, B. A., Goff, S. A., Klein, T. M., and Fromm, M. E. (1991). C1- and R-dependent expression of the maize Bz1 gene requires sequences with homology to mammalian myb and myc binding sites. *Plant Cell* 3, 317–325. doi: 10.1105/tpc.3.3.317

Sainz, M. B., Grotewold, E., and Chandler, V. L. (1997). Evidence for direct activation of an anthocyanin promoter by the maize C1 protein and comparison of DNA binding by related Myb domain proteins. *Plant Cell* 9, 611–625. doi: 10.1105/tpc.9.4.611

Schwinn, K., Venail, J., Shang, Y., Mackay, S., Alm, V., Butelli, E., et al. (2006). A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus Antirrhinum. *Plant Cell* 18, 831–851. doi: 10.1105/tpc.105.039255

Sharma, V. K., Chandresh, C., Kumar, R., Meena, R. D., and Mawliya, M. (2016). Influence of planting time and fruit maturity stage on antioxidant activity, phenols and anthocyanin contents in sweet pepper. *Res. Crops* 17, 360–368. doi: 10.5958/2348-7542.2016.00061.9

Spelt, C., Quattrocchio, F., Mol, J. N., and Koes, R. (2000). *anthocyanin1* of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* 12, 1619–1632. doi: 10.1105/tpc.12.9.1619

Staiger, D., Becker, F., Schell, J., Koncz, C., and Palme, K. (1991). Purification of tobacco nuclear proteins binding to a CACGTG motif of the chalcone synthase promoter by DNA affinity chromatography. *Eur. J. Biochem.* 199, 519–527. doi: 10.1111/j.1432-1033.1991.tb16150.x

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. *Curr. Opin. Plant Biol.* 4, 447–456. doi: 10.1016/s1369-5266(00)00199-0

Sunil, L., and Shetty, N. P. (2022). Biosynthesis and regulation of anthocyanin pathway genes. *Appl. Microbiol. Biotechnol.* 106, 1783–1798. doi: 10.1007/s00253-022-11835-z

Wang, J., Dai, Y., Pan, L., Chen, Y., Dai, L., Ma, Y., et al. (2025). Fine mapping and identification of *CaTTG1*, a candidate gene that regulates the hypocotyl anthocyanin accumulation in *Capsicum annuum* L. *Hortic. Plant J* 11 (1), 264–274. doi: 10.1016/j.hpj.2023.05.016

Wang, H., Guan, S., Zhu, Z., Wang, Y., and Lu, Y. (2013b). A valid strategy for precise identifications of transcription factor binding sites in combinatorial regulation using bioinformatic and experimental approaches. *Plant Methods* 9, 34. doi: 10.1186/1746-4811-9-34

Wang, J. E., Li, D. W., Gong, Z. H., and Zhang, Y. L. (2013a). Optimization of virus-induced gene silencing in pepper (Capsicum annuum L.). *Genet. Mol. Research: GMR* 12, 2492–2506. doi: 10.4238/2013.July.24.4

Wang, L., Tang, W., Hu, Y., Zhang, Y., Sun, J., Guo, X., et al. (2019). A MYB/bHLH complex regulates tissue-specific anthocyanin biosynthesis in the inner pericarp of red-centered kiwifruit *Actinidia chinensis* cv. Hongyang. *Plant Journal: Cell Mol. Biol.* 99, 359–378. doi: 10.1111/tpj.14330

Wang, Y., Wang, Z., Du, H., Chen, B., Wang, G., Wang, Q., et al. (2023). Fine mapping of the flavonoid 3',5'-hydroxylase gene controlling anthocyanin biosynthesis in pepper anthers and stems. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1232755

Wang, Z., Wong, D. C. J., Wang, Y., Xu, G., Ren, C., Liu, Y., et al. (2021). GRAS-domain transcription factor PAT1 regulates jasmonic acid biosynthesis in grape cold stress response. *Plant Physiol.* 186, 1660–1678. doi: 10.1093/plphys/kiab142

Wang, J., Yang, G., Chen, Y., Dai, Y., Yuan, Q., Shan, Q., et al. (2022). Genome-wide characterization and anthocyanin-related expression analysis of the *B-BOX* gene family in *capsicum annuum* L. *Front. Genet.* 13. doi: 10.3389/fgene.2022.847328

Wei, Z. Z., Hu, K. D., Zhao, D. L., Tang, J., Huang, Z. Q., Jin, P., et al. (2020). *MYB44* competitively inhibits the formation of the *MYB340-bHLH2-NAC56* complex to regulate anthocyanin biosynthesis in purple-fleshed sweet potato. *BMC Plant Biol.* 20, 258. doi: 10.1186/s12870-020-02451-y

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

Xu, W., Grain, D., Le Gourrierec, J., Harscoët, E., Berger, A., Jauvion, V., et al. (2013). Regulation of flavonoid biosynthesis involves an unexpected complex transcriptional regulation of *TT8* expression, in Arabidopsis. *New Phytol.* 198, 59–70. doi: 10.1111/nph.12142

Yan, S., Chen, N., Huang, Z., Li, D., Zhi, J., Yu, B., et al. (2020). *Anthocyanin Fruit* encodes an R2R3-MYB transcription factor, SlAN2-like, activating the transcription of *SlMYBATV* to fine-tune anthocyanin content in tomato fruit. *New Phytol.* 225, 2048–2063. doi: 10.1111/nph.16272

Zhang, Z., Li, D. W., Jin, J. H., Yin, Y. X., Zhang, H. X., Chai, W. G., et al. (2015). VIGS approach reveals the modulation of anthocyanin biosynthetic genes by CaMYB in chili pepper leaves. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00500

Zhang, Y., Ye, J., Liu, C., Xu, Q., Long, L., and Deng, X. (2020). Citrus PH4-Noemi regulatory complex is involved in proanthocyanidin biosynthesis via a positive feedback loop. *J. Exp. Bot.* 71, 1306–1321. doi: 10.1093/jxb/erz506

Zhu, Z., Wang, H., Wang, Y., Guan, S., Wang, F., Tang, J., et al. (2015). Characterization of the cis elements in the proximal promoter regions of the anthocyanin pathway genes reveals a common regulatory logic that governs pathway regulation. *J. Exp. Bot.* 66, 3775–3789. doi: 10.1093/jxb/erv173

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 12 March 2025
DOI 10.3389/fpls.2025.1540674

OPEN ACCESS

EDITED BY
Moonhyuk Kwon,
Gyeongsang National University, Republic of Korea

REVIEWED BY
Saet Buyl Lee,
National Institute of Agricultural Science, Republic of Korea
Soyon Park,
University of Missouri, United States
Pradeep Yerramsetty,
University at Buffalo, United States

*CORRESPONDENCE
Jungmin Ha
jungmin.ha@snu.ac.kr

RECEIVED 06 December 2024
ACCEPTED 20 February 2025
PUBLISHED 12 March 2025

CITATION
Cho Y, Kwon H, Kim BC, Shim D and Ha J (2025) Identification of genetic factors influencing flavonoid biosynthesis through pooled transcriptome analysis in mungbean sprouts.
Front. Plant Sci. 16:1540674.
doi: 10.3389/fpls.2025.1540674

Identification of genetic factors influencing flavonoid biosynthesis through pooled transcriptome analysis in mungbean sprouts

Yeonghun Cho[1], Hakyung Kwon[1,2], Byeong Cheol Kim[3], Donghwan Shim[4] and Jungmin Ha[1,2]*

[1]Department of Agriculture, Forestry and Bioresources and Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea, [2]Crop Genomics Lab, Plant Genomics and Breeding Institute, Seoul National University, Seoul, Republic of Korea, [3]Department of Plant Science, Gangneung-Wonju National University, Gangneung, Republic of Korea, [4]Department of Biological Sciences, Chungnam National University, Daejeon, Republic of Korea

Introduction: Mungbean (*Vigna radiata L.*) is gaining increasing interest among legume crops because of its nutritional value. Various secondary metabolites that act as antioxidants and bioactive compounds are beneficial for human health. The secondary metabolite content in plants is easily influenced by environmental conditions, and this influence varies depending on the genotype.

Materials and Methods: Here, we screened six genotypes with consistently high and low content of major secondary metabolites (gallic acid, chlorogenic acid, neo-chlorogenic acid, genistin, formononetin, catechin, syringic acid, and resveratrol) across environmental replicates. Transcriptome data obtained from the individual genotypes were pooled into two groups: high and low levels of secondary metabolites.

Results and Discussion: Of the 200 differentially expressed genes identified using stringent criteria, 23 were annotated in the secondary metabolite pathway. By combining the results of the secondary metabolite and transcriptome data, we identified six key genes encoding four enzymes (*CCoAOMT1*; Caffeoyl-CoA O-methyltransferase, *CYP81E1*; 4'-methoxyisoflavone 2'-hydroxylase, *DFR*; dihydroflavonol-4-reductase, and *HCT*; shikimate O-hydroxycinnamoyltransferase) that commonly influence the content of secondary metabolites (catechin, chlorogenic acid, formononetin, and genistin) in mungbeans. Through regulatory network analysis, NAC042 and MYB74 transcription factors were identified. These transcription factors regulate the expression of four key genes in mungbean, *CCoAOMT1 (Vradi02g00000724.1), CYP81E1(Vradi09g00002897.1), DFR(Vradi07g00001336.1)*, and *HCT(Vradi07g00000614.1)* leading to high flavonoid content.

Conclusion: These results provide information on the common genetic factors involved in the production of secondary metabolites, which can improve the nutritional value of mungbeans and contribute to the development of elite mungbean cultivars.

KEYWORDS

mungbean sprout, ultra-high-performance liquid chromatography, secondary metabolite content, sample pooling, RNA-Seq, gene expression

1 Introduction

Mungbean (*Vigna radiata L.*) is a major legume crop with high economic and agricultural value. It ranks third among legumes in South Asia, after chickpeas and pigeon pea, with an average annual per capita consumption of up to 2 kg (Vijayalakshmi et al., 2003). Mungbeans are an important source of vegetable proteins and an important resource for the alternative meat industry (Dahiya et al., 2015; Pataczek et al., 2018). Mungbeans are grown extensively throughout Asia, particularly in countries such as Thailand, where they account for up to 83% of the total area under legumes (Vijayalakshmi et al., 2003). In general, mungbeans are well suited to intercropping systems, complement other crops, and provide key benefits, such as fixing nitrogen from the atmosphere, improving soil fertility, and reducing the need for chemical fertilizers. The crop also has a high market value, resulting in a high price per kilogram, making it an attractive option for farmers seeking both profitability and sustainability (Nair and Schreinemachers, 2020).

Recently, mungbeans have attracted increasing interest from customers owing to their high levels of secondary metabolites (Tang et al., 2014; Pataczek et al., 2018). One of the ways to consume mungbeans is to sprout them; the key benefit of this method is that germination increases the amount of antioxidants and other bioactive compounds (Ebert et al., 2017). Specifically, carotenoids, vitamin E, and various phenolic compounds in seeds increase after germination, which helps prevent chronic diseases in humans, such as inflammation and cancer (Kim et al., 2012; Nderitu et al., 2013; Basli et al., 2017; Teodor et al., 2020; Li et al., 2021).

To increase the levels of secondary metabolites and enhance the nutritional value of mungbeans, understanding the regulatory mechanisms of the biosynthetic pathways of target compounds and the underlying genetic factors remains crucial. Although mungbeans are rich in secondary metabolites compared to other legumes, only a few studies have been conducted on the biosynthetic pathways of secondary metabolites in mungbeans, with most studies on soybeans. Furthermore, previous studies attempting to increase the antioxidant components in mungbeans have focused on controlling stress factors (Lim et al., 2022a, 2022b), as most secondary metabolites in plants are produced as defense mechanisms against biotic and abiotic stresses (Hartmann, 2007; Zandalinas et al., 2017). However, the effectiveness of this approach varies among genotypes, because each genotype may have different sensitivities or tolerances to environmental factors (Kim et al., 2023a). Therefore, investigating the genetic factors that regulate the biosynthesis of secondary metabolites in mungbeans is necessary. To achieve this, eliminating environmental factors and varietal differences to identify the underlying genetic determinants remains imperative.

Gene pooling has been used to identify useful genetic factors for plant breeding and development, thereby minimizing genotype specificity (Cronn et al., 2012; Botnari et al., 2018). Genetic pooling methods have been used in various strategies to identify common genetic factors associated with target traits regardless of individual variation. In a previous study conducted on pigeon peas (*Cajanus cajan* L.), the QTL-seq approach was used to identify candidate genes for flowering and leaf shape (Singh et al., 2022). In sesame (*Sesamum indicum* L.), the same approach was used to identify loci that control lignan content (Kim et al., 2023b). Because of the unique characteristics of gene pooling, it can also be applied to non-model plants, including cultivars or wild varieties, and their pooled transcriptomes (Ward et al., 2012; Goyal et al., 2016).

In this study, we conducted transcriptome pooling of six mungbean species to identify the key genes that regulate the biosynthetic pathways of secondary metabolites. These findings provide information on the key genetic factors involved in the biosynthesis of secondary metabolites and contribute to improving mungbean varieties with enhanced nutritional value.

2 Materials and methods

2.1 Plant material and sample preparation

A total of 12 mungbean cultivars were used in the present study (Supplementary Table S1). Seeds were germinated by soaking in distilled water for 16 h at 37°C using an incubator (ISS-4075R, Jeiotech). The germinated seeds were moved to a sprout cultivator (ST001A, Sundotcom) and cultivated for three days at 28–30°C with a water spraying interval of 4 h and water spraying time of 2 min (Kim et al., 2021). After three days, the sprouts were harvested. Each sample of the 12 mungbean varieties consisted of 30 sprouts, which were divided into two groups of 15 sprouts each for secondary metabolite phenotyping and RNA extraction, respectively. The first group was

dried at 70°C for 24 h in an oven (HQ-DO84, CORETECH) and then finely grounded for extraction with 70% ethanol (Supelco, cat. no. 1009831011) in the dark at room temperature. The other group was stored at -80°C for further study and RNA extraction.

2.2 Measurement of flavonoid content using UPLC

Ultra-high-performance liquid chromatography (UPLC) analysis was performed to measure the metabolite content (Nexera series equipped with MPM-40, SCL-40, SPD-M40, LC-40, SIL-40, and CTO-40 units from Shimadzu, Kyoto, Japan) with a photodiode array detector. Separation was conducted using a ZORBAX SB-C18 column (3.5 µm, 4.6 x 150 mm; Agilent, PN 863953-902, Santa Clara, USA). The sample injection volume was 2 µL and the column oven temperature was set at 40°C. The solvent for the mobile phase gradients was ultrapure water (Thermo Fisher, W5-4, Korea with 0.1% acetic acid solution [v/v; solvent A] and acetonitrile [solvent B]). The ratio of solvent A proceeded with 95% and solvent B flowed at 1 mL/min as follows: 0–10 min 95–90% A, 10–11 min 90–85% A, 11–15 min 85–80% A, 15–16 min 80–70% A, 16–25 min 70–65% A, 25–28 min 65–50% A, 28–32 min 50% A, 32.1 min solvent A was increased from 50 to 95%, and 32.1–40 min 95% A. The chromatograms of reference compounds, biochanin A (Chemfaces, CFN99734, Wuhan, China), catechin (Chemfaces, CFN99646), genistin (Chemfaces, CFN90250), caffeic acid (Chemfaces, CFN99646), daidzein (Chemfaces, CFN98774), daidzin (Chemfaces, CFN99101), formononetin (Chemfaces, CFN99962), gallic acid (Chemfaces, CFN99624), genistein (Chemfaces, CFN98681), glycitein (Chemfaces, CFN99106), glycitin (Chemfaces, CFN99105), isovitexin (Chemfaces, CFN98620), neochlorogenic acid (Chemfaces, CFN97472), p-coumaric acid (Chemfaces, CFN98794), resveratrol (Chemfaces, CFN98791), syringic acid (Chemfaces, CFN98884), t-ferulic acid (Chemfaces, CFN92394), vitexin (Chemfaces, CFN98601), chlorogenic acid (Chemfaces, CFN99116), coumestrol (Chemfaces, CFN96040), kaempferol (Chemfaces, CFN98838), myricetin (Chemfaces, CFN98877), and quercetin (Chemfaces, CFN99272) were extracted at wavelengths ranging from 270 to 330 nm (Supplementary Table S2). The analysis results were determined with three replicate.

2.3 RNA extraction and cDNA library construction

Total RNA was extracted from six cultivars of mungbean sprouts using a Ribospin Plant RNA extraction kit (GeneAll, Songpa-gu, South Korea) following the manufacturer's protocol. cDNA libraries for RNA-seq were constructed using a TruSeq Staranede mRNA LT Sample Prep Kit (Illumina Inc., San Diego, CA, USA). Twelve libraries were constructed with two libraries from each of the six cultivars used for replication. The size and quality of libraries used for sequencing were determined using a 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA). Sequencing runs were conducted in the paired-end mode using a Truseq SBS Kit on the Illumina NovaSeq 6000 platform. The RNA sequencing data were deposited in the NCBI SRA database (PRJNA1086206).

2.4 Identification and analysis of differentially expressed genes

The reference genome and gene annotation data for mapping were downloaded from the Seoul National University Crop Genomics Lab (http://plantgenomics.snu.ac.kr) (Ha et al., 2021). The reads were mapped to the mungbean reference genome using HiSat2 (Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype | Nature Biotechnology, n.d). Counts per million of mapped read values were calculated using FeatureCounts (Liao et al., 2014). The raw read counts were normalized using the trimmed mean of M-values (TMM) method, and differentially expressed genes (DEGs) were identified using edgeR (Robinson et al., 2010). DEGs were defined as genes with an absolute log2foldchange (|log2FC|) ≥ 1 between the two groups.

2.5 Functional annotation of DEGs and iRegNet regulatory co-expression network analysis

Gene ontology (GO) (Gene Ontology Consortium, 2004) enrichment analysis was performed using Clusterprofiler (Yu et al., 2012). Kyoto Encyclopedia of Genes and Genomes ontology pathway enrichment analysis was conducted using ClusterProfiler. iRegNet regulatory co-expression network analysis was conducted with homologous genes from *Arabidopsis thaliana* detected using BLAST.

2.6 Validation of DEGs using real-time quantitative reverse transcription-PCR

Real-time quantitative reverse transcription PCR (qRT-PCR) was performed using a PrimeScript RT Reagent Kit with gDNA Eraser (TaKaRa Bio Inc., San Jose, CA, USA) according to the manufacturer's protocols. Forty cycles of PCR amplification were performed using gene-specific primers. The relative gene expression levels were calculated using the $2^{-\Delta\Delta CT}$ method. The housekeeping gene, eukaryotic initiation factor 5A (*EIF5A*, Vradi07g21320), was used as a normalization control.

3 Results

3.1 Contents of secondary metabolites in mungbean sprouts

In a previous study on the metabolic profiling of 50 mungbean genotypes, 12 genotypes were selected based on the criteria of significant variations in the contents of 23 major secondary metabolites, and an expression analysis was conducted to identify key genes involved in the biosynthesis of secondary metabolites (Kim et al., 2023a). Building on these findings, we analyzed these 23

secondary metabolites again in the 12 genotypes to further investigate differences in metabolite accumulation between genotypes with consistently high and low secondary metabolite contents. Among the 23 commonly detected secondary metabolites in legumes, 19 were detected due to species-specific differences, comprising biochanin A, caffeic acid, catechin, coumestrol, daidzein, daidzin, formononetin, gallic acid, genistein, genistin, isovitexin, myricetin, p-coumaric acid, quercetin, resveratrol, syringic acid, sum of chlorogenic acid and neo-chlorogenic acid and vitexin; however, four metabolites, glycitein, clycitin, kaempferol, and t-ferulic acid, were not detected (Supplementary Table S3). To ensure robustness in selecting target metabolites for genetic analysis, we prioritized metabolites and genotypes at consistent levels across trials. As determined by one-way analysis of variance (ANOVA) at a significance level of 0.05, the contents of gallic acid, sum of chlorogenic acid and neo-chlorogenic acid, genistin, formononetin, catechin, syringic acid, and resveratrol were significantly higher in genotypes 205, 304, and 667 than in SH, 313, and DH, which was consistent with the results of our previous study (Figure 1). Other metabolites showed no significant differences between genotypes. Based on the consistently high and low levels of these secondary metabolites, genotypes 205, 304, and 667 were designated as the high group, whereas SH, 313, and DH were classified as the low group. These groups were then selected for further transcriptome analysis to compare expression patterns and investigate the key genetic factors involved in the biosynthetic pathways of these secondary metabolites.

3.2 Identification of DEGs between the high group and low group

To investigate the genetic factors underlying the differences in the secondary metabolite content, the total mRNA of the six mungbean genotypes, 205, 304, 667, SH, 313, and DH, were sequenced. In total, 80 Gb of reads (6.6 Gb per library on average) were obtained from the six mungbean genotypes (Supplementary Table S4). In general, 96.97% of the reads were properly aligned with the mungbean reference genome (Ha et al., 2021).

To minimize genotype-specific variations, sequencing data from each genotype were pooled into high group and low group. Between the groups (high group vs. low group), 4,217 DEGs were identified, which comprised 1,838 upregulated and 2,379 downregulated genes (Figure 2A, Supplementary Table S5).

A group-wise comparison alone may not sufficiently capture consistent expression differences because DEG detection can be disproportionately influenced by a small number of genotype pairs exhibiting extreme variations. To minimize genotype-specific variations, a more stringent selection criterion was applied by analyzing DEGs through pairwise comparisons across all possible genotype combinations within the High and Low groups (Figure 2B). In total, nine pairwise comparisons were performed: Genotype 205 vs. Genotype SH, Genotype 205 vs. Genotype 313, Genotype 205 vs. Genotype DH, Genotype 304 vs. Genotype SH, Genotype 304 vs. Genotype 313,… Genotype 667 vs. Genotype DH. Each DEG was classified according to the number of pairwise comparisons in which it was detected. D1 denotes a gene that was identified in only one of the nine comparisons, and D9 denotes a gene that was consistently detected in all nine comparisons. After cross-validation, 200 DEGs were selected based on appearing in seven (D7), eight (D8), or nine (D9) pairs of comparisons. These genes with highly stable differential expression patterns were designated as true DEGs and subjected to further functional analysis.

3.3 Functional annotation of DEGs

To determine the biological functions of DEGs identified through cross-validation, functional annotation of the 200 DEGs, which were identified at least seven pairs of combinations (D7, D8, and D9) was conducted using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and GO enrichment analyses. In the KEGG pathway analysis, 43 DEGs were annotated to five pathways and the top three pathways based on p-value harbor 14 genes; 4 genes in Isoflavonoid biosynthesis, 7 genes in Phenylpropanoid biosynthesis, and 3 genes in Tropane, piperidine, and pyridine alkaloid biosynthesis. In the GO analysis, 154 DEGs were annotated to 15 clusters. The top three clusters based on p-value include “phenylpropanoid biosynthetic process” (7 genes), “phenylpropanoid metabolic process” (7 genes), and “secondary metabolite biosynthetic process” (8 genes). Based on the two enrichment analyses, 23 DEGs were annotated as genes related to the biosynthetic pathways of secondary metabolites (Figure 3A, Supplementary Table S6).

3.4 Validation of gene expression levels using qRT-PCR

Among the 23 DEGs, six were selected for qRT-PCR as key genes in the biosynthetic pathways that regulate the content of secondary metabolites (Table 1, Supplementary Table S7). These genes include caffeoyl-CoA O-methyltransferase (CCoAOMT: Vradi02g00000724.1, Vradi03g00001121.1), 4’-methoxyisoflavone 2’-hydroxylase (CYP81E1: Vradi09g00002897.1, Vradi02g00004162.1), dihydroflavonol-4-reductase (DFR: Vradi07g00001336.1), and shikimate O-hydroxycinnamoyltransferase (HCT: Vradi07g00000614.1), all of which are enzymes involved in the biosynthesis of chlorogenic acid, genistin, formononetin, catechin, syringic acid, and resveratrol (Figure 4). Consistent with the RNA-seq results, the expression levels of CCoAOMT and HCT were upregulated (Figure 5A), while CYP81E1 and DFR were downregulated in the high group than in the low group (Figure 5B, Table 1). These findings reinforce the reliability of our RNA-seq data, highlighting clear differential expression patterns that align with the distinct metabolic profiles observed in the high group and low group.

4 Discussion

Understanding the mechanisms by which secondary metabolites accumulate is necessary to enhance the nutritional

10.3389/fpls.2025.1540674

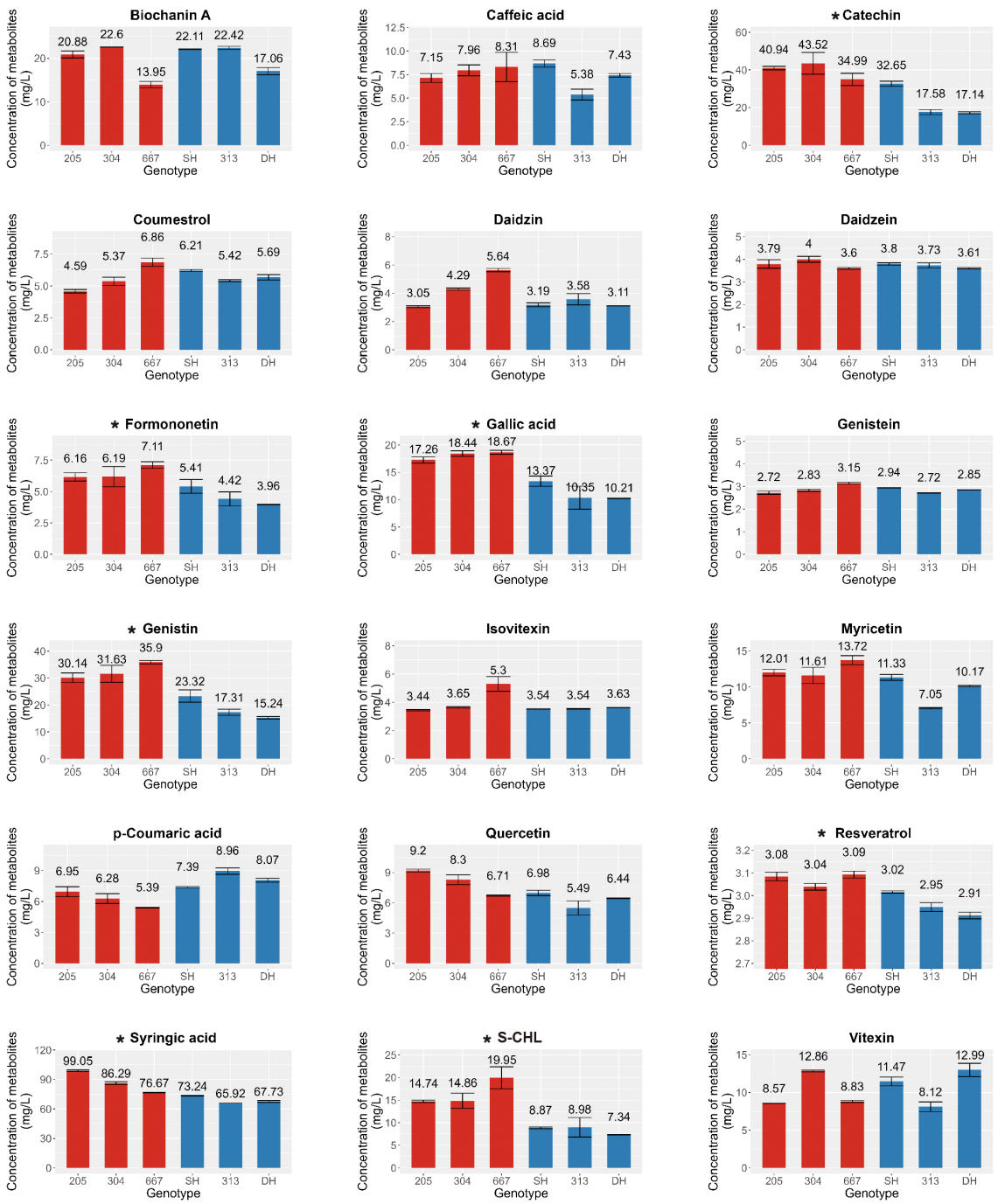


FIGURE 1
Contents of Secondary metabolites in mungbean sprouts of six genotypes. X-axis indicate mungbean sprouts genotype (205, 304, 667, SH, 313, and DH). Y-axes indicate the concentration of secondary metabolites (Catechin, Syringic acid, Genistin, Formononetin, Total Chlorogenic acid, Gallic acid, Resveratrol, Daidzein, Genistein, Vitexin, p-Coumaric acid, Caffeic acid, Daidzin, Isovitexin, Myricetin, Quercetin, Coumestrol, Biochanin A, and S-CHL, sum of chlorogenic acid and neo-chlorogenic acid.), determined with three replicates. The error bar indicates the standard deviation. The asterisk indicates metabolites that were statistically significant based on the ANOVA test ($p < 0.05$). Red and blue bars indicates high group and low group, respectively.

value of mungbeans and develop elite cultivars. Previous studies have artificially induced stress conditions, including biotic and abiotic stresses, to increase secondary metabolite content (Singh et al., 2002). However, the response of different genotypes to these stress conditions varied with different response mechanisms against external factors (Kim et al., 2022) and certain genotypes showed consistently higher metabolite content with statistical significance compared to others under different conditions, suggesting the existence of genetic regulators that control genes and enzymes that influence the production of secondary metabolites.

10.3389/fpls.2025.1540674

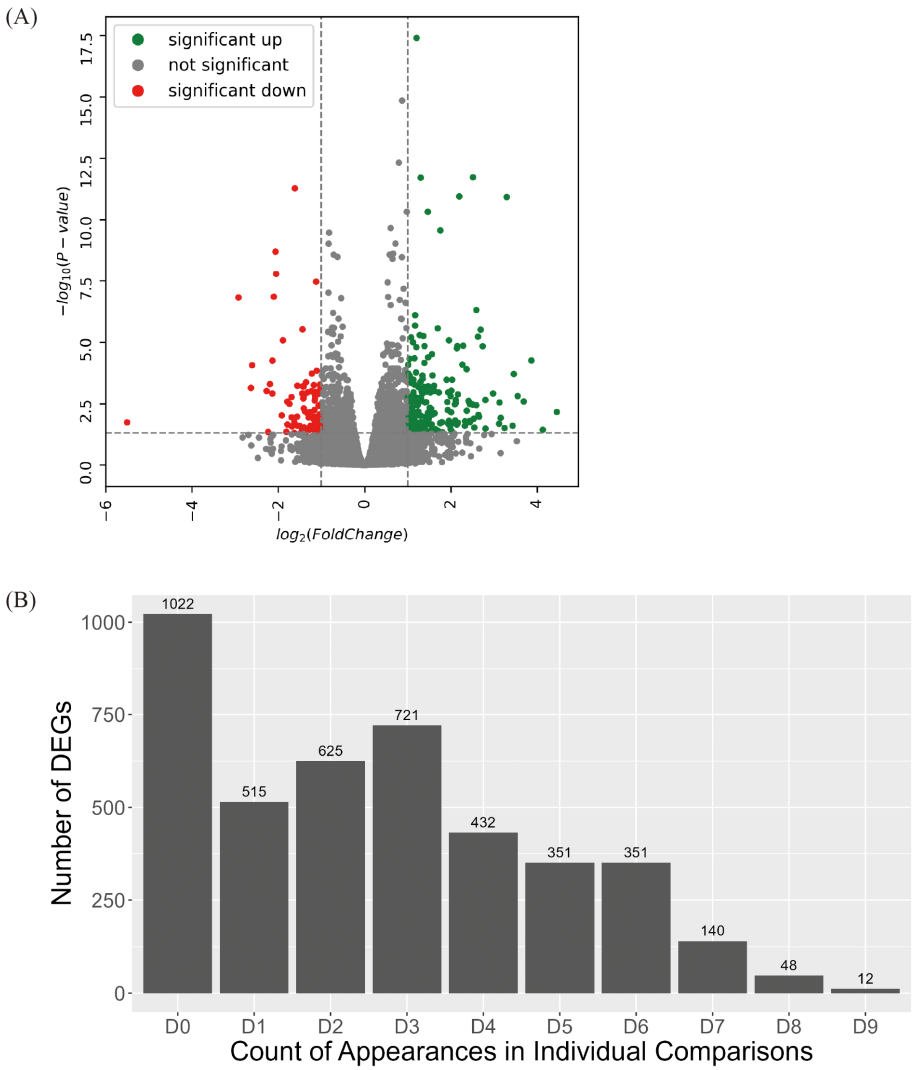


FIGURE 2
Identification of DEGs between high group and low group. **(A)** Volcano plot of DEGs identified between high group and low group. Colored dots represent DEG with value of 2foldchange. **(B)** Statistics of DEGs identified across each comparison between high group and low group. Among 4,217 DEGs between high group and low group, the number of DEGs cross-validated through individual comparison were identified. Every possible comparison between individuals in high group and low group were conducted. "D1" indicates the number of DEG identified between one genotype of high group and one genotype of low group. "D9" indicates the number of DEG appeared in every nine possible combinations between each genotype in high group and low group. "D0" indicates the number of DEGs unique to the high group and low group comparison.

Sample pooling has also been used in several genetic studies. For example, QTL-seq and genome-wide association studies use pooled DNA samples to identify candidate genes associated with target traits (Gyawali et al., 2019; Singh et al., 2022; Kim et al., 2023b). Although sample pooling offers optimized cost and statistical efficiency, previous research has revealed that it has the disadvantage of false detection caused by pooling bias (Wang and Paterson, 1994; Cutler and Jensen, 2010; Rajkumar et al., 2015). Despite this drawback, it remains a useful tool to identify common genetic variations, and researchers have tried various approaches to overcome false detections by overlapping identified DEGs from different methods (Ko and Van Raamsdonk, 2023). Therefore, in this study, we calculated DEGs using two combined approaches to minimize false detections. The first approach to calculating DEGs was RNA transcript pooling, grouping each member of the high group and low group as a replicate, resulting in 4,217 DEGs (Figure 2A). The second approach for calculating the DEGs was to compare each genotype of the high group with each genotype of the low group. With a total of nine possible combinations, DEGs detected from seven or more combinations were identified, and, in total, 200 DEGs were confirmed to be detected in HvsL DEGs (Figure 2B). Most of the 200 DEGs identified using our stringent criteria were annotated as being involved in secondary metabolite pathways according to KEGG (Figure 3A) and GO (Figure 3B) analyses, indicating that false detections were successfully removed from our DEGs.

Although the functions of major secondary metabolites in mungbeans are well known owing to their nutritional

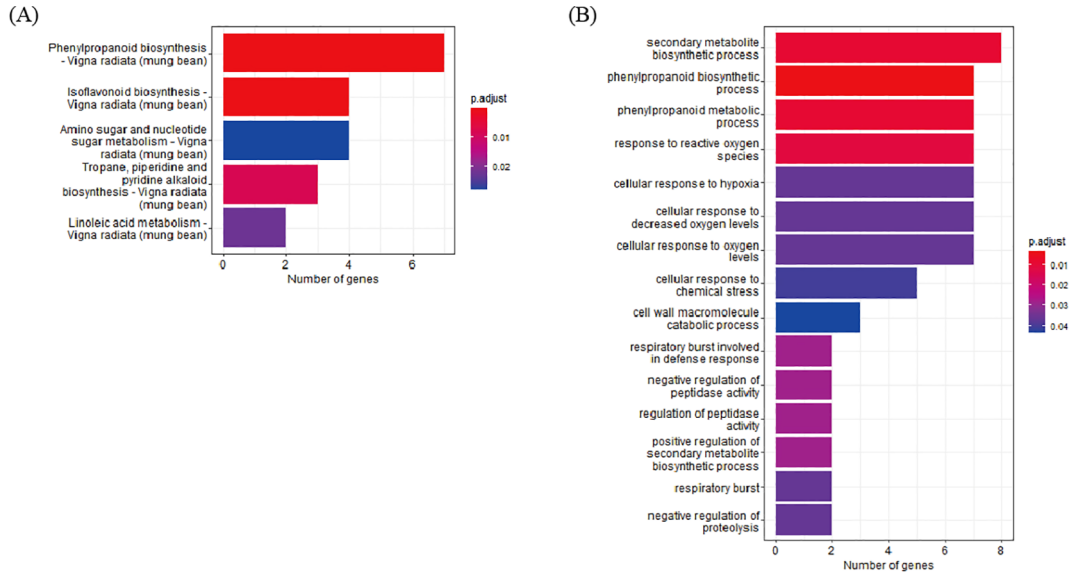


FIGURE 3
Enrichment analysis of DEG. **(A)** KEGG enrichment. **(B)** GO enrichment analysis of DEGs.

significance, the mechanisms of secondary metabolite biosynthesis in mungbeans remain unclear. Through functional annotation and the integration of metabolic and transcriptomic data, six genes encoding four key enzymes (CCoAOMT: Catechin, CYP81E1: Catechin and Chlorogenic acid, DFR: Genistin and Formononetin, and HCT: Genistin and Formononetin) were identified. *HCT* and *CCoAOMT*, which are involved in the biosynthesis of catechin and chlorogenic acid, showed higher expression levels in the high group genotypes, resulting in increased catechin and chlorogenic acid content in high group (Figure 4). These two metabolites, which are abundant in coffee and tea, offer many health benefits, including stable blood pressure and anti-cancer effects (Khalesi et al., 2014; Hayakawa et al., 2020). A previous study conducted on coffee (*Coffea canephora*) suggested that increased expression of *HCT* directly results in high levels of chlorogenic acid, which is consistent with the results of this study (Lepelley et al., 2007; Lallemand et al., 2012). In mungbean sprouts, chlorogenic acid content has been reported to be associated with the level of *HCT* expression under salinity stress (Kang, et al., 2022), which indicates that *HCT* is a key enzyme for the biosynthesis of chlorogenic acid in mungbeans. In *Arabidopsis*, eriodictyol accumulation has been reported to be caused by *CCoAOMT* (Wils et al., 2013), leading to increased catechin levels.

The expression levels of DFR and CYP81E1, which are involved in the isoflavonoid biosynthesis pathway, were significantly lower in high group with higher genistin and formononetin content than in low group (Figure 4) (Petit et al., 2007; Uchida et al., 2017). Naringenin and liquiritigenin, as intermediate precursors, are catalyzed by various flavonoids using these enzymes (Liu et al., 2021). Decreased expression of *DFR* and *CYP81E1* can activate an alternate biosynthetic pathway, resulting in the production of genistin and formononetin, which originate from naringenin and

TABLE 1 RNA-seq result of six candidate genes used for qRT-PCR validation.

Gene name	Regulation	Gene ID	Log_2 Fold Change	P-value	FDR	EC number	Description
CCoAOMT_1	Down	Vradi02g00000724.1	5.97	1.95.E-10	7.76.E-09	2.1.1.104	caffeoyl-CoA O-methyltransferase
CCoAOMT_2	Down	Vradi03g00001121.1	5.46	1.64.E-03	8.02.E-03	2.1.1.104	caffeoyl-CoA O-methyltransferase
CYP81E1_1	Up	Vradi09g00002897.1	-10.85	2.94.E-16	8.18.E-14	1.14.14.89	4'-methoxyisoflavone 2'-hydroxylase
CYP81E1_2	Up	Vradi02g00004162.1	-5.94	7.71.E-10	2.57.E-08	1.14.14.89	4'-methoxyisoflavone 2'-hydroxylase
DFR	Down	Vradi07g00001336.1	-1.24	2.11.E-05	1.91.E-04	1.1.1.219	dihydroflavonol-4-reductase
HCT	Down	Vradi07g00000614.1	3.41	9.60.E-08	1.72.E-06	2.3.1.133	shikimate O-hydroxycinnamoyltransferase

10.3389/fpls.2025.1540674

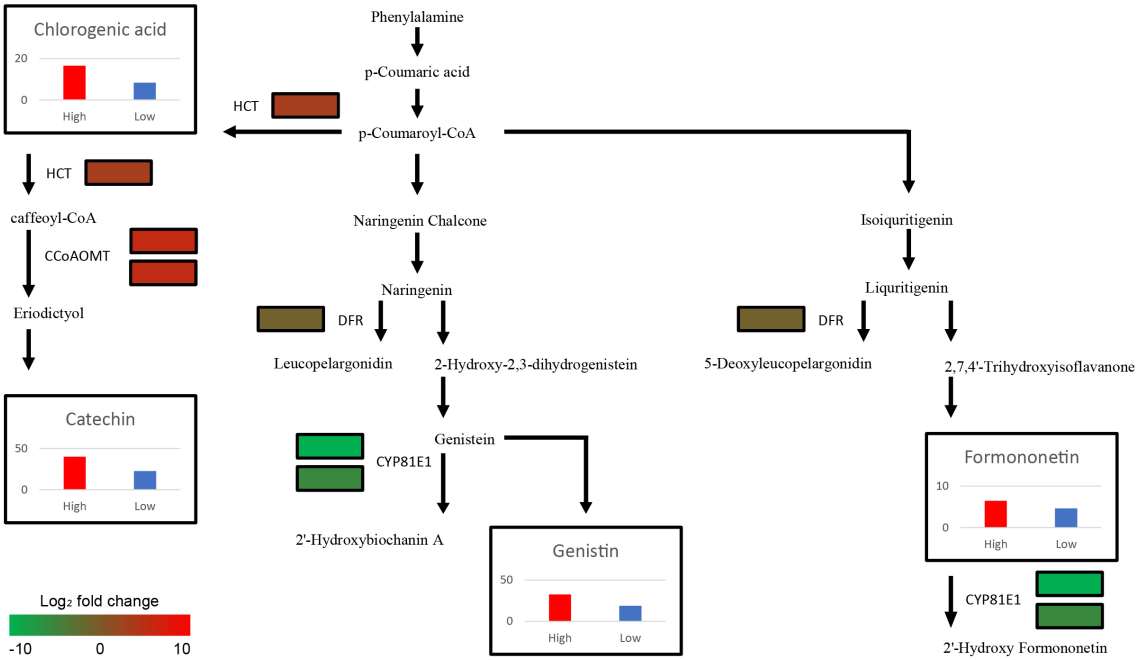


FIGURE 4
Schematic illustration of Secondary metabolite biosynthesis pathway with expression levels of 6 DEGs (CCoAOMT, CYP81E1, DFR, and HCT paralogs) and the average contents of 4 metabolites (Catechin, Chlorogenic acid, Formononetin, and Genistin) from each group. The color scale indicates fold change level of each DEGs between high group and low group. The bar plots show secondary metabolite content of High (Red bar) and Low group (Blue bar). The y-axis of four graphs represents concentration of metabolites (mg/L). CCoAOMT, caffeoyl-CoA O-methyltransferase; CYP81E1, 4′-methoxyisoflavone 2′-hydroxylase; DFR, dihydroflavonol-4-reductase; HCT, shikimate O-hydroxycinnamoyltransferase.

liquritigenin, respectively. The previous research findings, reporting that *CYP81E1* catalyzes the hydroxylation of the isoflavone formononetin, to yield 2′-hydroxyformononetin (Akashi et al., 1998), are consistent with our results, which demonstrate that the expression level of *CYP81E1* has a negative correlation with the formononetin content. The expression levels of *DFR* and *CYP81E1* in this study were lower in high group, which had higher genistin and formononetin content than low group (Figure 4). These results indicate that the downregulation of CYP81E1 and DFR may enhance the accumulation of genistin and formononetin in high group by redirecting biosynthetic pathways.

Our analysis revealed that the flavonoid content was generally elevated across different types of flavonoids in high group, with significant differences observed for chlorogenic acid, catechin, genistin, and formononetin (Figure 1). Additionally, despite the differences in the genetic backgrounds of the six mungbean genotypes, the expression levels of the six genes encoding the four key enzymes involved in the biosynthesis of these flavonoids were highly correlated. This suggests that these key enzymes are co-regulated by common genetic or transcriptional mechanisms, ultimately influencing flavonoid biosynthesis. To identify potential transcription factors upstream of these key enzymes,

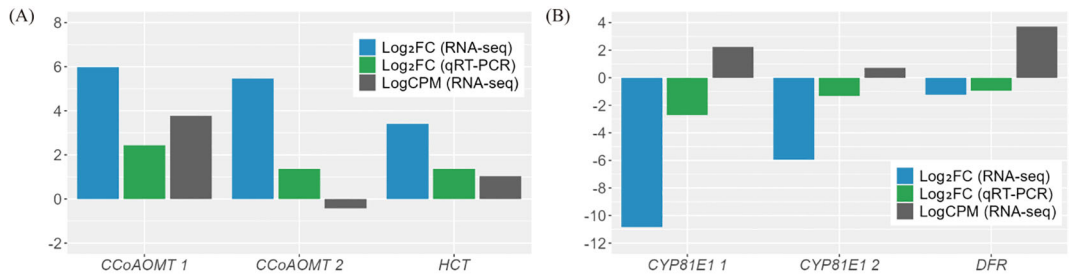


FIGURE 5
Validation of the DEGs through qRT-PCR analysis. **(A)** Up-regulated genes. **(B)** Down-regulated genes. The results of qRT-PCR were determined with three technical replicates. Log counts per million reads (LogCPM) values were used to identify key genes between paralogues for subsequent analysis. CCoAOMT, caffeoyl-CoA O-methyltransferase; CYP81E1, 4′-methoxyisoflavone 2′-hydroxylase; DFR, dihydroflavonol-4-reductase; HCT, shikimate O-hydroxycinnamoyltransferase.

10.3389/fpls.2025.1540674

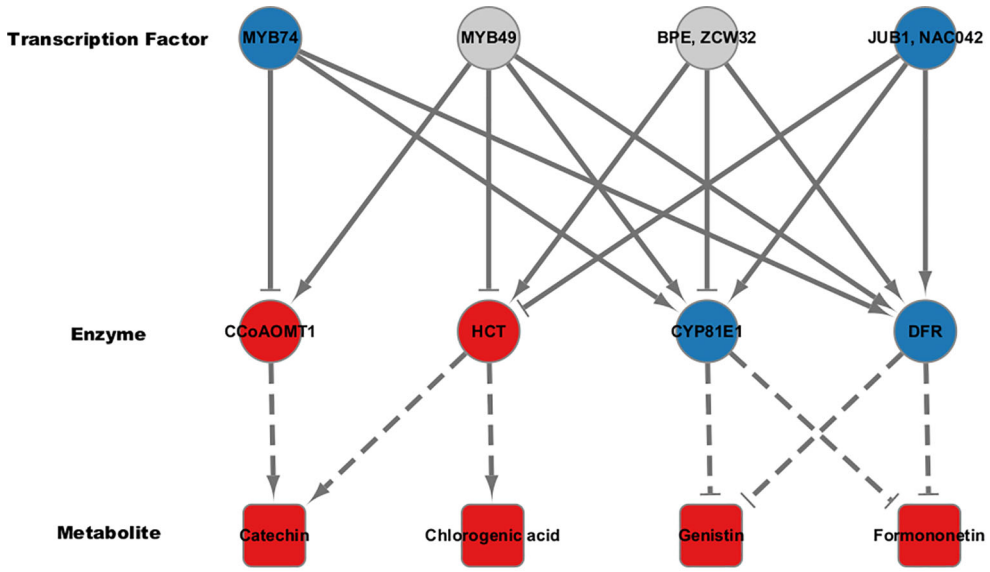


FIGURE 6
Regulatory co-expression network analysis for flavonoid biosynthetic pathway. Transcription factors affecting key enzymes in the flavonoid biosynthetic pathway were predicted using iRegNet. Arrows represent positive correlations, while T-shaped ends indicate negative correlations. Node colors reflect direction of regulation; Red for up-regulation and blue for down-regulation. Dotted lines suggest the association between the key enzymes and their corresponding secondary metabolites.

homologous *Arabidopsis* genes of these key enzymes were subjected to iRegNet analysis (Vradi02g00000724.1; AT4G34050, Vradi09g00002897.1; AT5G36220, Vradi07g00001336.1; AT5G42800, and Vradi07g00000614.1; AT2G19070) (Figure 6) (Shim et al., 2021). Four transcription factors regulate these enzymes: MYB49 (AT5G54230), MYB74 (AT4G05100), NAC042 (AT2G43000), and ZCW32 (AT1G59640). Among the mungbean homologs of these transcription factors, MYB74 (Vradi10g00000584.1) and NAC042 (Vradi06g00001115.1) showed significantly lower expression in high group compared to low group, with log2FC values of -3.68 and -4.3, respectively (Table 2). MYB74 and NAC042 are known to negatively regulate CCoAOMT1 and HCT, respectively, and positively regulate CYP81E1 and DFR. These findings suggest that the downregulation of MYB74 and NAC042 may have led to the upregulation of CCoAOMT1 and HCT and the downregulation of CYP81E1 and DFR, resulting in higher levels of catechin, chlorogenic acid, genistin, and formononetin in the high group.

While these findings provide insights into the transcriptional regulation of secondary metabolite biosynthesis, further studies are needed to validate the functional roles of key genetic factors. Such studies using CRISPR-based knockout, overexpression, and complementation test experiments will directly confirm their roles in flavonoid biosynthesis. Additionally, environmental factors influencing the expression of key genes and metabolite accumulation should be investigated to optimize growth conditions, develop elite cultivars, and enhance metabolite production in field. These approaches will provide insights into the regulatory mechanisms involved in secondary metabolism in mungbean.

In this study, we utilized metabolic data consistent with environmental conditions, along with DEGs identified using stringent criteria obtained from a pooled transcriptome. This approach aims to minimize environmental influences and eliminate variations due to genetic diversity among genotypes. The secondary metabolite content in mungbean sprouts was

TABLE 2 Fold changes of mungbean homologs of *Arabidopsis thaliana* transcription factors. "*A. thaliana*" refers to gene name from *Arabidopsis thaliana*, and "*V. radiata*" refers to gene names from *Vigna radiata*.

TF name	*A.thaliana*	*V.radiata*	Log_2FC	p-value	Avg. RPKM	
					High group	Low group
MYB49	AT5G54230	Vradi04g00003889.1	N.D.	N.D.	16.512	24.66
MYB74	AT4G05100	Vradi10g00000584.1	-3.678	1.45E-05	749.90	1260.50
NAC042	AT2G43000	Vradi06g00001115.1	-4.295	2.427E-13	3488.04	7445.27
ZCW32	AT1G59640	Vradi07g00000781.1	0.500	0.004	30475.54	27157.72

TF, Transcription Factor; Avg, average; N.D., not detected in high group vs. low group DEGs.

associated with the expression levels of the genes involved in the key biosynthetic pathways for each target metabolite. These results indicate that the accumulation or biosynthesis of secondary metabolites is regulated at the transcriptional level in mungbean sprouts. We identified the key genetic factors affecting the contents of major secondary metabolites. The findings of this study will contribute to the sustained and efficient production of antioxidant compounds in mungbean sprouts, thereby facilitating the development of high-value products with added benefits.

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

YC: Formal analysis, Visualization, Writing – original draft. HK: Data curation, Investigation, Validation, Writing – review & editing. BK: Investigation, Methodology, Writing – original draft. DS: Supervision, Validation, Writing – review & editing. JH: Conceptualization, Formal analysis, Funding acquisition, Supervision, Validation, 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 National Research Foundation of Korea (NRF) grants funded by the Korean government (MSIT; Grant No. 2022R1A4A1030348). This work was supported by the New Faculty Startup Fund from Seoul National University.

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.2025.1540674/full#supplementary-material

References

Akashi, T., Aoki, T., and Ayabe, S. (1998). CYP81E1, a Cytochrome P450 cDNA of Licorice (*Glycyrrhiza eChinata*L.), Encodes Isoflavone 2′-Hydroxylase. *Biochem. Biophys. Res. Commun.* 251, 67–70. doi: 10.1006/bbrc.1998.9414

(n.d). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype | Nature Biotechnology. Available online at: https://www.nature.com/articles/s41587-019-0201-4 (Accessed February 1, 2024).

Basli, A., Belkacem, N., Amrani, I., Basli, A., Belkacem, N., and Amrani, I. (2017). Health benefits of phenolic compounds against cancers," in. *Phenolic Compounds - Biol. Activity*. (IntechOpen) 193–210. doi: 10.5772/67232

Botnari, V., Andronic, L., and Cotenco, E. (2018). Bases for genetic improvement of grain legumes and cereals. In: *International congress on oil and protein crops, 20-24 mai 2018, Chişinău*. Chişinău, Republica Moldova: Tipografia "Artpoligraf", p. 64.

Cronn, R., Knaus, B. J., Liston, A., Maughan, P. J., Parks, M., Syring, J. V., et al. (2012). Targeted enrichment strategies for next-generation plant biology. *Am. J. Bot.* 99, 291–311. doi: 10.3732/ajb.1100356

Cutler, D. J., and Jensen, J. D. (2010). To pool, or not to pool? *Genetics* 186, 41–43. doi: 10.1534/genetics.110.121012

Dahiya, P. K., Linnemann, A. R., Van Boekel, M. A. J. S., Khetarpaul, N., Grewal, R. B., and Nout, M. J. R. (2015). Mung bean: technological and nutritional potential. *Crit. Rev. Food Sci. Nutr.* 55, 670–688. doi: 10.1080/10408398.2012.671202

Ebert, A. W., Chang, C.-H., Yan, M.-R., and Yang, R.-Y. (2017). Nutritional composition of mungbean and soybean sprouts compared to their adult growth stage. *Food Chem.* 237, 15–22. doi: 10.1016/j.foodchem.2017.05.073

Gene Ontology Consortium (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* 32, D258–D261. doi: 10.1093/nar/gkh036

Goyal, E., Amit, S. K., Singh, R. S., Mahato, A. K., Chand, S., and Kanika, K. (2016). Transcriptome profiling of the salt-stress response in Triticum aestivum cv. Kharchia Local. *Sci. Rep.* 6, 27752. doi: 10.1038/srep27752

Gyawali, A., Shrestha, V., Guill, K. E., Flint-Garcia, S., and Beissinger, T. M. (2019). Single-plant GWAS coupled with bulk segregant analysis allows rapid identification and corroboration of plant-height candidate SNPs. *BMC Plant Biol.* 19, 412. doi: 10.1186/s12870-019-2000-y

Ha, J., Satyawan, D., Jeong, H., Lee, E., Cho, K.-H., Kim, M. Y., et al. (2021). A near-complete genome sequence of mungbean (Vigna radiata L.) provides key insights into the modern breeding program. *Plant Genome* 14, e20121. doi: 10.1002/tpg2.20121

Hartmann, T. (2007). From waste products to ecochemicals: Fifty years research of plant secondary metabolism. *Phytochemistry* 68, 2831–2846. doi: 10.1016/j.phytochem.2007.09.017

Hayakawa, S., Ohishi, T., Miyoshi, N., Oishi, Y., Nakamura, Y., and Isemura, M. (2020). Anti-cancer effects of green tea epigallocatchin-3-gallate and coffee chlorogenic acid. *Molecules* 25, 4553. doi: 10.3390/molecules25194553

Khalesi, S., Sun, J., Buys, N., Jamshidi, A., Nikbakht-Nasrabadi, E., and Khosravi-Boroujeni, H. (2014). Green tea catechins and blood pressure: a systematic review and meta-analysis of randomised controlled trials. *Eur. J. Nutr.* 53, 1299–1311. doi: 10.1007/s00394-014-0720-1

Kim, D.-K., Jeong, S. C., Gorinstein, S., and Chon, S.-U. (2012). Total polyphenols, antioxidant and antiproliferative activities of different extracts in mungbean seeds and sprouts. *Plant Foods Hum. Nutr.* 67, 71–75. doi: 10.1007/s11130-011-0273-x

Kim, S., Lee, E., Lee, J., An, Y. J., Oh, E., Kim, J. I., et al. (2023b). Identification of QTLs and allelic effect controlling lignan content in sesame (Sesamum indicum L.) using QTL-seq approach. *Front. Genet.* 14. doi: 10.3389/fgene.2023.1289793

Kim, B. C., Lim, I., and Ha, J. (2023a). Metabolic profiling and expression analysis of key genetic factors in the biosynthetic pathways of antioxidant metabolites in mungbean sprouts. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1207940

Kim, B. C., Lim, I., Jeon, S. Y., Kang, M., and Ha, J. (2021). Effects of irrigation conditions on development of mungbean (Vigna radiata L.) sprouts. *Plant Breed. Biotechnol.* 9, 310–317. doi: 10.9787/PBB.2021.9.4.310

Ko, B., and Van Raamsdonk, J. M. (2023). RNA sequencing of pooled samples effectively identifies differentially expressed genes. *Biology* 12, 812. doi: 10.3390/biology12060812

Lallemand, L. A., Zubieta, C., Lee, S. G., Wang, Y., Acajjaoui, S., Timmins, J., et al. (2012). A structural basis for the biosynthesis of the major chlorogenic acids found in coffee. *Plant Physiol.* 160, 249–260. doi: 10.1104/pp.112.202051

Lepelley, M., Cheminade, G., Tremillon, N., Simkin, A., Caillet, V., and McCarthy, J. (2007). Chlorogenic acid synthesis in coffee: An analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. *Plant Sci.* 172, 978–996. doi: 10.1016/j.plantsci.2007.02.004

Li, J., Lu, Y., Chen, H., Wang, L., Wang, S., Guo, X., et al. (2021). Effect of photoperiod on vitamin E and carotenoid biosynthesis in mung bean (*Vigna radiata)* sprouts. *Food Chem.* 358, 129915. doi: 10.1016/j.foodchem.2021.129915

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Lim, I., Kang, M., Kim, B. C., and Ha, J. (2022a). Metabolomic and transcriptomic changes in mungbean (Vigna radiata (L.) R. Wilczek) sprouts under salinity stress. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1030677

Lim, I., Kim, B. C., Park, Y., Park, N. I., and Ha, J. (2022b). Metabolic and Developmental Changes in Germination Process of Mung Bean (*Vigna radiata* (L.) R. Wilczek) Sprouts under Different Water Spraying Interval and Duration. *J. Food Qual.* 2022, e6256310. doi: 10.1155/2022/6256310

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

Nair, R., and Schreinemachers, P. (2020). "Global Status and Economic Importance of Mungbean," in *The Mungbean Genome*. Eds. R. M. Nair, R. Schafleitner and S.-H. Lee (Springer International Publishing, Cham), 1–8. doi: 10.1007/978-3-030-20008-4_1

Nderitu, A. M., Dykes, L., Awika, J. M., Minnaar, A., and Duodu, K. G. (2013). Phenolic composition and inhibitory effect against oxidative DNA damage of cooked cowpeas as affected by simulated *in vitro* gastrointestinal digestion. *Food Chem.* 141, 1763–1771. doi: 10.1016/j.foodchem.2013.05.001

Pataczek, L., Zahir, Z. A., Ahmad, M., Rani, S., Nair, R., Schafleitner, R., et al. (2018). Beans with benefits—The role of mungbean (Vigna radiate) in a changing environment. *Am. J. Plant Sci.* 9, 1577–1600. doi: 10.4236/ajps.2018.97115

Petit, P., Granier, T., d'Estaintot, B. L., Manigand, C., Bathany, K., Schmitter, J.-M., et al. (2007). Crystal structure of grape dihydroflavonol 4-reductase, a key enzyme in flavonoid biosynthesis. *J. Mol. Biol.* 368, 1345–1357. doi: 10.1016/j.jmb.2007.02.088

Rajkumar, A. P., Qvist, P., Lazarus, R., Lescai, F., Ju, J., Nyegaard, M., et al. (2015). Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genomics* 16, 548. doi: 10.1186/s12864-015-1767-y

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Singh, K. B., Foley, R. C., and Oñate-Sánchez, L. (2002). Transcription factors in plant defense and stress responses. *Curr. Opin. Plant Biol.* 5, 430–436. doi: 10.1016/S1369-5266(02)00289-3

Singh, V., Sinha, P., Obala, J., Khan, A. W., Chitikineni, A., Saxena, R. K., et al. (2022). QTL-seq for the identification of candidate genes for days to flowering and leaf shape in pigeonpea. *Heredity* 128, 411–419. doi: 10.1038/s41437-021-00486-x

Tang, D., Dong, Y., Ren, H., Li, L., and He, C. (2014). A review of phytochemistry, metabolite changes, and medicinal uses of the common food mung bean and its sprouts (Vigna radiata). *Chem. Cent. J.* 8, 4. doi: 10.1186/1752-153X-8-4

Teodor, E. D., Ungureanu, O., Gatea, F., and Radu, G. L. (2020). The potential of flavonoids and tannins from medicinal plants as anticancer agents. *Anti-Cancer Agents Medicinal Chem.* 20, 2216–2227. doi: 10.2174/1871520620666200516150829

Uchida, K., Akashi, T., and Aoki, T. (2017). The missing link in leguminous pterocarpan biosynthesis is a dirigent domain-containing protein with isoflavanol dehydratase activity. *Plant Cell Physiol.* 58, 398–408. doi: 10.1093/pcp/pcw213

Vijayalakshmi, P., Amirthaveni, S., Devadas, R. P., Weinberger, K., Tsou, S. C. S., and Shanmugasundaram, S. (2003). *Enhanced bioavailability of iron from mungbeans and its effects on health of schoolchildren*. (Shanhua, Taiwan: AVRDC-WorldVegetableCenter).

Wang, G. L., and Paterson, A. H. (1994). Assessment of DNA pooling strategies for mapping of QTLs. *Theoret. Appl. Genet.* 88–88, 355–361. doi: 10.1007/BF00223645

Ward, J. A., Ponnala, L., and Weber, C. A. (2012). Strategies for transcriptome analysis in nonmodel plants. *Am. J. Bot.* 99, 267–276. doi: 10.3732/ajb.1100334

Wils, C. R., Brandt, W., Manke, K., and Vogt, T. (2013). A single amino acid determines position specificity of an *Arabidopsis thaliana* CCoAOMT-like *O*-methyltransferase. *FEBS Lett.* 587, 683–689. doi: 10.1016/j.febslet.2013.01.040

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zandalinas, S. I., Sales, C., Beltrán, J., Gómez-Cadenas, A., and Arbona, V. (2017). Activation of secondary metabolism in citrus plants is associated to sensitivity to combined drought and high temperatures. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01954

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 19 May 2025
DOI 10.3389/fpls.2025.1545011

OPEN ACCESS

EDITED BY
Xiaomeng Liu,
Wuhan Polytechnic University, China

REVIEWED BY
Ke Wang,
Anhui Agricultural University, China
Yueya Zhang,
The University of Chicago, United States

*CORRESPONDENCE
Michael Handford
mhandfor@uchile.cl

RECEIVED 13 December 2024
ACCEPTED 14 April 2025
PUBLISHED 19 May 2025

CITATION
Covarrubias MP, Uribe F, Arias-G D, Cabedo P and Handford M (2025) Fruit-specific overexpression of lipoyl synthase increases both bound and unbound lipoic acid and alters the metabolome of tomato fruits. *Front. Plant Sci.* 16:1545011. doi: 10.3389/fpls.2025.1545011

Fruit-specific overexpression of lipoyl synthase increases both bound and unbound lipoic acid and alters the metabolome of tomato fruits

María Paz Covarrubias, Felipe Uribe, Daniela Arias-G, Pamela Cabedo and Michael Handford*

Centro de Biología Molecular Vegetal, Department of Biology, Faculty of Sciences, University of Chile, Santiago, Chile

Introduction: Lipoic acid (LA) is a key, yet overlooked player in primary metabolism, due to its role as a cofactor for various multi enzymatic complexes such as the E2 subunits of pyruvate dehydrogenase (PDH) and alpha-ketoglutarate dehydrogenase (kGDH). In recent years, this molecule has seen renewed interest given its strong antioxidant properties and its applications as a dietary supplement. The mechanisms behind the synthesis of LA in vivo have been elucidated, identifying lipoyl synthase (LIP1) as the key enzyme required for this process.

Methods: Therefore, in this work, we used the fruit-specific polygalacturonase (PG) promoter to guide *Solanum lycopersicum* (tomato) *LIP1* (*SlLIP1*) overexpression in stably transformed tomatoes.

Results: The resulting plants presented higher transcript levels of *SlLIP1* in a fruit-specific manner, accumulated more bound and unbound LA yet lacked major phenotypic defects at both the vegetative and reproductive growth stages. Furthermore, changes in the expression of genes related to LA synthesis were explored and a metabolomic study was carried out. Specific metabolite patterns were clearly distinguishable between untransformed and stably transformed lines. For instance, trehalose 6-phosphate, GABA and proline levels were generally higher, whilst glucose 6-phosphate and UDP-glucose levels were lower in fruits of the *SlLIP1* transformants.

Discussion: In addition, as the overexpression of *SlLIP1* results in lower transcript levels of *E2 PDH* and *E2 kGDH*, and enhanced amounts of LA-bound targets, we speculate that the proportion of unlipoylated E2 subunits of PDH and kGDH may have decreased. This work could assist in obtaining crops with a higher LA content and therefore improved health benefits.

KEYWORDS

antioxidant, lipoylation, polygalacturonase promoter, *Solanum lycopersicum*, TCA cycle

1 Introduction

Antioxidants are molecules known generally for their ability to provide reducing power and neutralize reactive oxygen species (ROS; Turrens, 2003). Although ROS are a natural byproduct of metabolism, they are often produced in electron transport chains, and even act as signaling molecules; excessive ROS production is also associated with cellular damage in organisms and aging (Turrens, 2003). Because of this, research into antioxidant molecules has surged in recent years for their applications in many industries, including in the food, pharmaceutical, agronomical, and cosmetic sectors, among others (Jimenez-Lopez et al., 2020; Lupo, 2001; Neha et al., 2019; Shahidi, 2000; Sun et al., 2024).

Lipoic acid (LA) and its reduced form (dihydrolipoic acid, DHLA) are antioxidants with several unique properties. LA is an eight-carbon molecule characterized by its two thiol groups in the carbon six and carbon eight positions that are responsible for its antioxidant nature. Like many antioxidants, LA is capable of neutralizing a wide range of ROS, regenerating other antioxidant molecules such as glutathione, and chelating heavy metals (Gorąca et al., 2011; Matsugo et al., 1996; Packer et al., 1995; Scott et al., 1994). However, unlike other antioxidants, LA is both water and lipid soluble in its free form (Packer et al., 1995). For these reasons, LA has been the focus of many studies that have demonstrated its beneficial influence over human health, such as preventing the effects of diabetic neuropathies, and stabilizing cognitive functions in Alzheimer's patients, or in other neurodegenerative diseases induced by oxidative stress (Borcea et al., 1999; Gilgun-Sherki et al., 2001; Packer et al., 2001). Other organisms, including plants, have also been shown to benefit from applications of LA. For example, exogenous LA ameliorates the effects of oxidative stress caused by high salinity in wheat seedlings, and promotes water-deficit tolerance in maize (Gorcek and Erdal, 2015; Saruhan Guler et al., 2021).

An additional feature of LA is that it can be found as an essential cofactor. In this role, LA is necessary for the functioning of various enzymatic complexes that participate in primary and secondary metabolism. These lipoylated targets include the E2 subunits of pyruvate dehydrogenase (PDH), alpha-ketoglutarate dehydrogenase (kGDH), and branched-chain 2-oxoacid dehydrogenase (BCDH), as well as the H-protein of the glycine decarboxylase complex (GDC) (Harmer et al., 2014). In plants, LA can be added onto its cognate enzyme through two distinct pathways: the direct ligation of an exogenous LA molecule onto a cognate enzyme in a 'salvage' pathway via lipoate protein ligase (LplA), or through the assembly of a lipoate cofactor *de novo*, from an octanoyl precursor followed by the addition of two sulfur atoms by lipoyl synthase (LIP1) (Ewald et al., 2014a; Yasuno and Wada, 2002). In the latter case, the octanoyl precursor is bound firstly to an acyl carrier protein (ACP) by octanoyltransferase (LIP2), enabling the 8-carbon precursor to be shuttled to either a E2 subunit (PDH, kGDH, BCDH) or H-protein (GDC) before being thiolyated by LIP1 in a rare example of a cofactor being assembled *in situ* (Ewald et al., 2014a; Sarvananda et al., 2023). Therefore, both the unbound (LA) and bound (lipoylated proteins) forms, acting as an antioxidant and cofactor, respectively, require the addition of thiol groups by LIP1, making it the key step in the metabolism of this molecule (Ewald et al., 2014).

It is known that the E2 subunit of PDH is a target for lipoylation and there are two PDH complexes in plants, with one being present in mitochondria (PDHm) and participating in the TCA cycle, while the other is found in chloroplasts (PDHc) and plays a role in *de novo* lipid biosynthesis (Tovar-Méndez et al., 2003). Studies of plastidial sunflower (LIP1p; *Helianthus annuum* L.) lipoyl synthase in *Arabidopsis* revealed that expression of this gene results in a redistribution of the glycerolipid species being produced, attributed to the depletion of S-adenosylmethionine (SAM or Ado-Met), whose reductive cleavage by LIP1 is required for thioylation to take place (Harmer et al., 2014; Martins-Noguerol et al., 2020). SAM also participates in various other cellular processes, including as a precursor for ethylene synthesis, a phytohormone associated with senescence and stress responses, as well as regulatory processes such as DNA methylation (Cronan, 2014; Harmer et al., 2014; Lu, 2000; Sauter et al., 2013). Indeed, expression of cotton (*Gossypium hirsutum* L.) LIP1 delayed senescence in *Arabidopsis* (Chen et al., 2021). Because LA is an antioxidant and participates in a central role by lipoylating several enzymatic complexes, it is expected that modifications in LA metabolism would lead to widespread metabolic changes.

In this work, we generated a system that allowed us to examine the effects of the organ-specific over-expression of LIP1. To do so, we overexpressed mitochondrial LIP1 from *Solanum lycopersicum* (tomato; *SlLIP1*, Solyc07g054540; Araya-Flores et al., 2020) in a fruit-specific manner in stably-transformed tomato. This was achieved by guiding the expression of *SlLIP1* with the promoter region of the tomato *polygalacturonase* gene (*PG*) which is triggered during fruit ripening in normal development (Montgomery et al., 1993). PG is a cell-wall degrading enzyme, and publicly-available transcriptome data (bar.utoronto.ca/eplant_tomato/) demonstrates that *PG* expression is detected exclusively in maturing tomato fruits. Changes in phenotype, gene expression, protein accumulation, and the fruit metabolome were assessed to elucidate the global impact of altered LA metabolism in plants, and especially in fruits.

2 Materials and methods

2.1 Vector construction

The full-length CDS of *SlLIP1* was previously cloned into the pGWB8 binary vector (Araya-Flores et al., 2020). Using primers containing the *Asc*I restriction site (primer pair 1-2; Supplementary Table 1), *SlLIP1* was amplified from the ATG start codon to the stop codon of the vector, including a 6xHis tag, and subcloned into the pCR™8/GW/TOPO® TA (Invitrogen), generating the pCR8-*SlLIP1* construct. The pCP binary vector is based on the backbone of the pB7FWG2-AtDXR vector, containing a 7,854 bp fragment with the T-DNA right and left borders (RB and LB) and the BAR gene, which confers resistance to BASTA® or glufosinate (herbicide) in plants (Arias et al., 2021). This pCP vector already contained an 804 bp polygalacturonase (PG) promoter fragment driving the transcription of the *DcLCYB* gene. To excise *DcLCYB*, 1

µg of pCP was digested with 10 U *Asc*I (New England Biolabs), following the manufacturer's protocol, generating linearized pCP with the PG promoter. The same digestion was performed on the pCR8-*SlLIP1* construct to obtain the *SlLIP1* insert. Finally, the *SlLIP1* insert was ligated overnight at 4°C into linearized pCP using 3 U of T4 ligase (Promega) at a 1:1 vector-to-insert ratio, resulting in the PG-*SlLIP1* vector (Figure 1A). This final vector was verified by sequencing (Macrogen Co.) and transformed into *Agrobacterium tumefaciens* (EHA105 strain). Positive clones were selected and used for the stable transformation of *Solanum lycopersicum* var. 'Micro-Tom'.

2.2 Tomato stable transformation and *in vitro* culture

S. lycopersicum var. 'Micro-Tom' was cultivated *in vitro* on solid Murashige & Skoog (MS) medium (4.4% MS salts with vitamins, 1.5% sucrose, and 1.5% agar, pH 5.8) and transformed according to Cruz-Mendívil et al. (2011). Briefly, young leaves from three-week-old plants were excised, and the tips were removed with a sterile scalpel. These explants were incubated in co-culture medium (4.4% MS salts with vitamins, 3% sucrose, 100 µM acetosyringone, 0.1 µM naphthaleneacetic acid, 2 µg/mL trans-zeatin) for 3–4 days before transformation. The explants were then immersed in an *Agrobacterium* culture harboring pCP-*SlLIP1* for 20 min, transferred to fresh co-culture medium, and incubated for 2 days at 23°C in darkness.

Next, the explants were transferred to induction medium (4.4% MS salts with vitamins, 3% sucrose, 2 µg/mL trans-zeatin, 0.1 µg/mL indole-3-butyric acid [IBA], 400 µg/mL timentin, and 0.25 µg/mL BASTA®) for callus induction and initial shoot development, which typically takes 4 to 6 weeks. The emerging shoots were then transferred to elongation medium (4.4% MS salts with vitamins, 1.5% sucrose, 400 µg/mL timentin, and 0.35 µg/mL BASTA®). Once the shoots reached a height of 8 cm, they were transferred to rooting medium (2.2% MS salts with vitamins, 1.5% sucrose, 5 µg/mL IBA, 400 µg/mL timentin, and 0.5 µg/mL BASTA®) until roots formed.

The plants were then acclimated in a greenhouse (16 h light/8 h dark photoperiod with white fluorescent light at 150 µmol/m²/s, 22–25°C) in plastic pots (20 x 10 cm) containing a mix of soil and vermiculite (2:1) or rock wool soaked with hydroponic medium (1.25 mM KNO_3, 1.5 mM $Ca(NO_3)_2$, 0.75 mM $MgSO_4$, 0.5 mM KH_2PO_4, 50 µM H_3BO_3, 10 µM $MnSO_4$, 2 µM $ZnSO_4$, 1.5 µM $CuSO_4$, 0.075 µM $(NH_4)_6Mo_7O_{24}$, 0.1 mM Na_2O_3Si, iron diethylenetriamine pentaacetate, pH 6).

To confirm the insertion of PG-*SlLIP1* into the plant genome, PCR was performed on DNA extracted from leaves using primers designed to amplify the region from the PG promoter to the 6xHis tag (primer pair 3-4, Supplementary Table 1).

2.3 RNA extraction and quantitative expression analysis

RNA was isolated from fully-expanded leaves and transgenic and wild-type (WT) tomato fruits at the green, breaker, pink and mature red stages (CFR, 1991) using a modified CTAB method (Arias et al.,

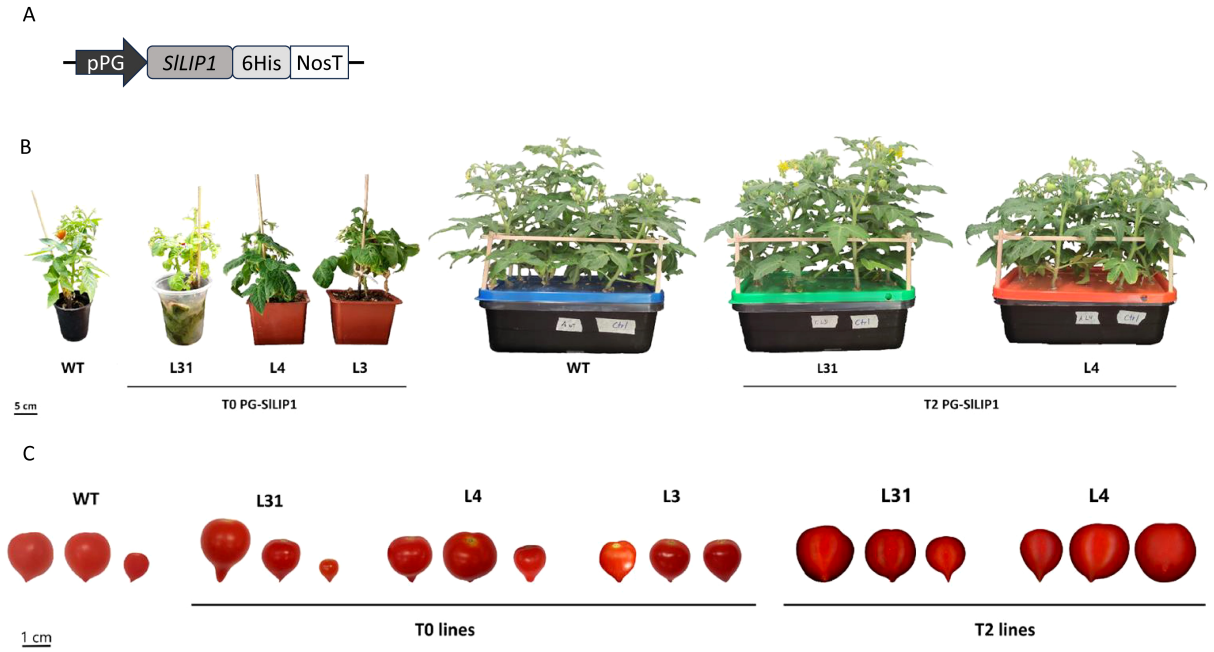


FIGURE 1
Generation of tomato plants overexpressing *SlLIP1*. **(A)** Diagram of the pCP vector used to transform tomato cv 'Micro-Tom'. **(B)** Three independent T0 lines (4-months old, L31, L4 and L3) and two in the T2 generation (3-months old, L31 and L4) exhibited normal development compared to WT. **(C)** Three representative fruits of PG-SlLIP1 lines at T0 and T2 generations.

2021, 2022), and cDNA synthesis was carried out using the ImProm-II™ Reverse Transcription System (Promega), as described in Arias et al. (2021). Quantitative expression analysis (qRT-PCR) was conducted using a Stratagene MX300P system (Agilent Technologies) with Forget-Me-Not™ qPCR Master Mix containing ROX (Biotium). From fruits, the primers used to analyze transcript levels of *SlLIP1* (primer pair 5-6), *SlLIP1p* (primer pair 7-8), *SlLIP2* (primer pair 9-10), *SlLplA* (primer pair 11-12), *E2-PDHm* (primer pair 13-14), *E2-PDHc* (primer pair 15-16), *E2-kGDH* (primer pair 17-18), and *SlSAMS1* (primer pair 19-20) are listed in Supplementary Table 1. *Actin7* (primer pair 21-22, Supplementary Table 1) was selected as the reference gene (Bokhale et al., 2023). To analyze the introduced *SlLIP1* in leaves, RT-PCR was carried out using primers specific to *SlLIP1* and the 6xHis tag (primer pair 5-4) and *GAPDH* was used as a housekeeping gene (primer pair 23-24). Each qRT-PCR was performed with three biological replicates and two technical replicates. Quantification was calculated using the method described by Pfaffl (2001).

2.4 Analysis of lipoic acid bound to proteins

Protein extraction from mature red transgenic and WT tomato fruits was performed according to Nilo-Poyanco et al. (2013). Protein concentrations were determined using the BCA method (Thermo Scientific), and samples were loaded onto an SDS-PAGE gel to confirm protein integrity via Coomassie Blue staining. Subsequently, immunoblot analysis was conducted using an anti-LA antibody (1:3000, Calbiochem 437695) and an anti-actin antibody (1:2000, Sigma-Aldrich SAB4301137). Both membranes were incubated with a secondary antibody conjugated to HRP (1:80000, Biorad 170-6515) and visualized using SuperSignal™ West Femto (Thermo Scientific). Signals were captured using a UVITEC Cambridge (Biomolecular Imaging-Alliance © Software) photodocumentation system, and images were processed using ImageJ software.

For the anti-LA immunoblot, the total signal intensity of each lane was recorded, while for the anti-actin immunoblot, the 42 kDa band was quantified (Jimenez-Lopez et al., 2014). Protein-bound LA quantification was calculated as the ratio between the sum of pixel intensity per area from the LA immunoblot and the pixel intensity per area from the 42 kDa band of the actin immunoblot for each sample, using 3 (T0 and T2) and 5 (WT) replicates. The final result for each transgenic line was normalized relative to the WT sample.

2.5 Analysis of unbound lipoic acid

To extract hydrophilic antioxidants, 10 mg of dry weight (DW) from mature red transgenic and WT tomato samples (B+15 days approximately) were ground using liquid nitrogen with a pestle and mortar. Then, 250 µL 0.5% acetic acid in methanol were added, followed by vortexing and sonication for 1 hour. The samples were centrifuged at 7,500 g for 10 min, and the supernatants were freeze-dried. The resulting pellet was resuspended in 10 µL of water to obtain the antioxidant solution.

To detect unbound LA, 1 µL of the antioxidant solution was loaded onto a nitrocellulose membrane, and a dot blot assay was performed. The membrane was blocked with 5% non-fat milk in TBS-T for 30 min. After three washes with TBS-T, the membrane was incubated with an anti-LA antibody (1:3000, Calbiochem 437695) for 30 min, followed by three additional washes with TBS-T. A secondary antibody conjugated to HRP (1:20000, Biorad 170-6515) was then applied, and signals were detected as described in section 2.4. As a positive control, 2 µL of total protein from the T0 PG-SlLIP1 L4 line was used.

2.6 Targeted metabolomic analysis

Polar metabolite extraction was performed following the protocol described by Alonso et al. (2010). Briefly, 10 mg DW from mature red transgenic and WT tomato samples were ground using liquid nitrogen with a pestle and mortar. As internal standards, 5 mM ^{13}C-glycine, 20 mM ^{13}C-mannose, and 1 mM ^{13}C-fumarate were added to each tube. Polar metabolites were extracted using 1 mL of boiling water and then incubated in a water bath at over 90°C for 10 min. The extracts were then placed on ice and centrifuged at 17,000 g for 5 min at 4°C. The supernatants were passed through a 0.22 µm filter, and the remaining pellets were subjected to a second extraction. Finally, the water-soluble metabolites were freeze-dried overnight.

The metabolites were resuspended in 500 µL of water, vortexed, and 150 µL were loaded onto a 0.2 µm nanosep MF centrifugal device and centrifuged at 17,000 g for 10 min at 4°C to detect sugars and sugar alcohols. The remaining volume was loaded onto a 3 kDa Amicon Ultra column, centrifuged at 14,000 g for 75 min at 4°C, and kept on ice until injection into the LC-MS/MS.

LC-MS/MS quantification was performed following the method described by Cocuron et al. (2014) to determine sugars, sugar alcohols, amino acids, phosphorylated compounds, and organic acids.

2.7 Statistical analysis

Transcript, bound and unbound LA, and metabolite levels were analyzed using a t-test ($p < 0.05$), with WT serving as the reference parameter for comparison.

Metabolomic data were uploaded to MetaboAnalyst 6.0 (Pang et al., 2024) for statistical analysis (one-factor), including PCA, heatmap, and pathway enrichment analysis on normalized data. Metabolites were used as predictor variables, while genotype (WT and transgenic lines) was used as the response variable. For clustering analysis, the Euclidean distance similarity measure and Ward's clustering algorithm were applied.

3 Results

3.1 Stable transformation of tomato with PG-SlLIP1 reduces seed production

Considering the essential role of lipoylation in several enzymes in primary metabolism, the overexpression of lipoyl synthase, the key enzyme required for this process, may be expected to generate pleiotropic phenotypic effects, even if limited to specific organs in plants. Therefore, to increase the LA content in fruits, we performed a stable transformation of tomato cv 'Micro-Tom' plants with *SlLIP1* (Solyc07g054540), which encodes mitochondrial lipoyl synthase (Araya-Flores et al., 2020). The CDS of *SlLIP1* was cloned into the pCP vector (Arias et al., 2021), a binary vector that contains an 804 bp fragment upstream of the transcription start site of the *Polygalacturonase* (*SlPG14*, Solyc10g080210.1.1) promoter from *S. chilense* (tomatillo). The expression of *PG* in tomato cv 'Micro-Tom' is fruit-specific starting in breaker and peaking at the pink stage (Supplementary Figure 1A). pCP also harbors a Nopaline synthase Terminator (NosT), and the *bar* gene, which confers BASTA® resistance. The *SlLIP1* CDS was cloned with a 6xHis tag at the 3' end (Figure 1A).

Using *Agrobacterium*-mediated transformation, young leaves were transformed with pCP harboring PG-SlLIP1, and the regeneration of the explants was carried out *in vitro*. Upon root elongation, plants were transferred *ex vitro* until the appearance of fruits. Three T0 transgenic lines, referred to as L31, L4 and L3, were analyzed and compared to WT plants (Figure 1B). By monitoring the growth and development of all lines, it was determined that the height of the aerial part of all plants at the end of the third month of acclimatization (4 months post-transformation) in the greenhouse was similar between L31, L4 and L3 and WT, indicating no significant impact on overall vegetative performance (Table 1). When the fruits were analyzed, the main difference observed was in the number of seeds; while WT fruits had seeds in all fruits, a reduction was found in L31 and L4, and no seeds were present in L3 (Table 1, Figure 1C). Due to this feature, transgenic lines L31 and L4 were propagated until the T2 generation, for further evaluation.

3.2 Overexpression of *SlLIP1* in tomato fruits increases lipoylated protein and unbound LA levels

The level of overexpression of *SlLIP1* was analyzed using qRT-PCR in mature red fruits from WT and transgenic PG-SlLIP1 lines across two different generations. Since the introduced *SlLIP1* is an endogenous gene in tomato, overexpression was defined as transcript levels higher than those in WT. Fruits from the transgenic lines PG-SlLIP1 L31, L4 and L3 at the T0 stage accumulated 2–3 fold higher transcript levels of *SlLIP1* than in WT. This pattern persisted into the T2 generation of L31 and L4 fruits (Figure 2A). Transcripts of introduced *SlLIP1* in fully-expanded leaves of all lines were absent, demonstrating that the PG promoter was not active in these organs (Supplementary Figure 2).

To determine whether this increase in *SlLIP1* transcript levels affected LA content, two experiments were conducted. First, lipoylated proteins were analyzed by immunoblotting, using an anti-LA antibody. The signal obtained was normalized to the signal of the anti-actin immunoblot, and the results were compared relative to WT, such that an increase in the LA/actin ratio suggests an increase in the lipoylation of target proteins (Supplementary Figure 3). In this context, fruits from T0 transgenic lines L31 and L4 showed up to ~2-fold higher lipoylated protein levels than WT. This rise was stably transmitted to the T2 generation (Figure 2B). The second approach measured free LA in extracts of hydrophilic antioxidants obtained from tomato fruits. To do so, a dot blot was performed using the same anti-LA antibody (Supplementary Figure 4). The results showed a significantly stronger signal in fruits from L31 and L4 at the T2 generation compared to WT fruits (Figure 2C). Overall, these results indicate that increasing *SlLIP1* levels in tomato fruits enhances bound and unbound LA content.

3.3 Transcript levels of LA biosynthesis and related genes are altered in transgenic tomato fruits with higher LA content

As an increase in *SlLIP1* transcript levels triggered a corresponding increase in bound and unbound LA content in

TABLE 1 Phenotypic description of WT and T0 and T2 PG-SlLIP1 transgenic tomato lines.

		Plant phenotype		Fruit phenotype	
	Genotype	Plant height (cm)	Flowering time (d)	Diameter (mm)	Presence of seeds[c]
	WT	18.8	58	13.3 ± 3.8[b]	++
T0 PG-SlLIP1	L31	16.5	40	15.7 ± 2.9[b]	+
	L4	10	67	16.3 ± 2.5[b]	+
	L3	11.5	67	13.0 ± 1.0[b]	–
T2 PG-SlLIP1	L31	22.3[a]	67	18.1 ± 2.3[b]	+
	L4	18.6[a]	60	17.8 ± 2.2[b]	+

a. Average of 18–21 plants.
b. Average of at least three fruits.
c. ++, >10 seeds/fruit; +, <5 seeds/fruit; -, absent.

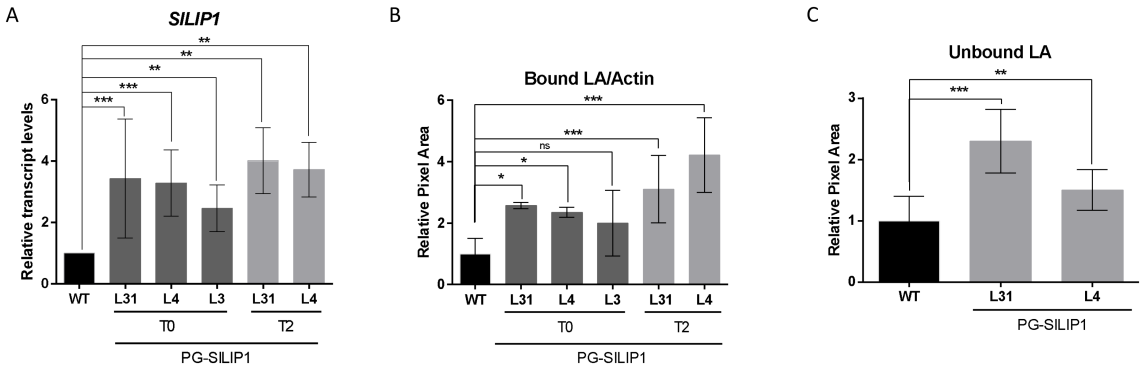


FIGURE 2
SlLIP1 transcript levels and LA content in fruits of T0 and T2 PG-SlLIP1 transgenic tomato lines. **(A)** qRT-PCR analysis of *SlLIP1* was performed on three different fruits, each with two technical replicates (n=6 ± SD), normalized against *Actin7* and calibrated to WT levels. **(B)** Quantification of immunoblot signals of protein-bound LA using an anti-LA antibody, normalized to the signal of an anti-actin antibody in WT and PG-SlLIP1 transgenic fruits. All T0 and T2 samples were subjected to this analysis at least three times, whilst 5 replicates are considered for WT. Bars show means ± SD. **(C)** Detection of unbound LA in WT and T2 PG-SlLIP1 fruits (n=9 ± SD). $p \leq 0.05$ (*), $p \leq 0.005$ (**), and $p \leq 0.0005$ (***) and not significant (n.s.).

fruits, the expression of various genes related to LA biosynthesis was analyzed to understand how transgenic fruits respond to the overexpression of *SlLIP1*. SlLIP1p (Solyc12g099700) is the plastidial lipoyl synthase isoform (Araya-Flores et al., 2020). *SlLIP1p* transcript levels in L31 and L4 at T0 were significantly higher than in WT fruits, while no differences were observed in L3 T0 or in both T2 lines (Figure 3A). On the other hand, the LA biosynthetic genes *SlLIP2* and *SlLplA* were more highly expressed in fruits from all the transgenic lines analyzed in the T0 generation (Figures 3B, C). Nevertheless, no differences were observed in the accumulation of *SlLIP2* transcripts in L31 and L4 at T2 (Figure 3B). In contrast, L4 at T2 showed higher transcript levels of *SlLplA* than WT and L31 fruits (Figure 3C).

Regarding the genes encoding lipoylated subunits, transcript levels of *E2 PDHm* (mitochondrial), *E2 PDHc* (plastidial) and *E2 kGDH* were evaluated. For *E2 PDHm*, transcript levels were higher in L31 and L3 at T0 compared to WT and L4. However, L31 and L4 at T2 showed lower transcript levels than WT (Figure 3D). In the case of *E2 PDHc*, L31 expressed this transcript more than WT, L4 and L3 fruits at T0, while L31 and L4 at T2 showed lower levels of transcript than WT (Figure 3E). For *E2 kGDH*, L31 and L4 at T0 showed higher transcript accumulation compared to WT, but at T2, L31 accumulated less of this transcript (Figure 3F).

Finally, during the enzymatic reaction of LIP1, SAM is consumed. Therefore, transcript levels of SAM synthetase 1 (*SAMS1*), the enzyme responsible for SAM synthesis (Zhang et al., 2020) were analyzed. L4 T0 fruits accumulated higher levels of this transcript, but this trend was not observed at T2. In contrast, L31 T2 fruits accumulated lower levels of *SAMS1* (Figure 3G). Taken together, these findings suggest that the fruit-specific overexpression of *SlLIP1* results in knock-on effects in transcript levels of lipoylation-related genes.

3.4 Primary metabolism is affected by the overexpression of *SlLIP1* in transgenic tomato fruits

To understand how the overexpression of *SlLIP1* and the increase in LA content modified the metabolite profiles in the PG-SlLIP1 tomato lines, a targeted metabolomic analysis of polar metabolites was performed using LC-MS/MS. A total of 99 metabolites were identified in mature red fruits, consisting of 10 sugars, 27 amino acids and their derivatives, and 62 organic acids and phosphorylated compounds. The data were evaluated through an unsupervised multivariate statistical analysis using the genotype as the response variable and metabolites as the predictor variable. Principal components analysis (PCA) of metabolites from transgenic PG-SlLIP1 (L31, L4 and L3) and WT tomato fruits explained 41% of the variance for PC1 and 21.3% for PC2. The projection clearly separated the transgenic and WT groups, with the three transgenic lines clustering together on the negative side of the plot (Figure 4A).

Next, the top 50 metabolites identified by ANOVA as being significantly-altered were clustered and visualized with a heatmap (Figure 4B). The three transgenic PG-SlLIP1 lines grouped together, distinct from WT. Thirty-five metabolites accumulated more in WT fruits than in transgenic lines, representing 70% of the metabolites identified. The remaining 15 metabolites accumulated more in the transgenic lines than in WT fruits, with four metabolites (asparagine, PEP, citrulline and trehalose) showing higher levels in L31 only. L4 and L3 fruits were more similar to each other, with four metabolites showing different patterns compared to L31 and WT fruits (UDP-arabinose, fructose 1,6-biphosphate, azelaic acid and ATP). Three metabolites (ADP-glucose, glutamate and GDP)

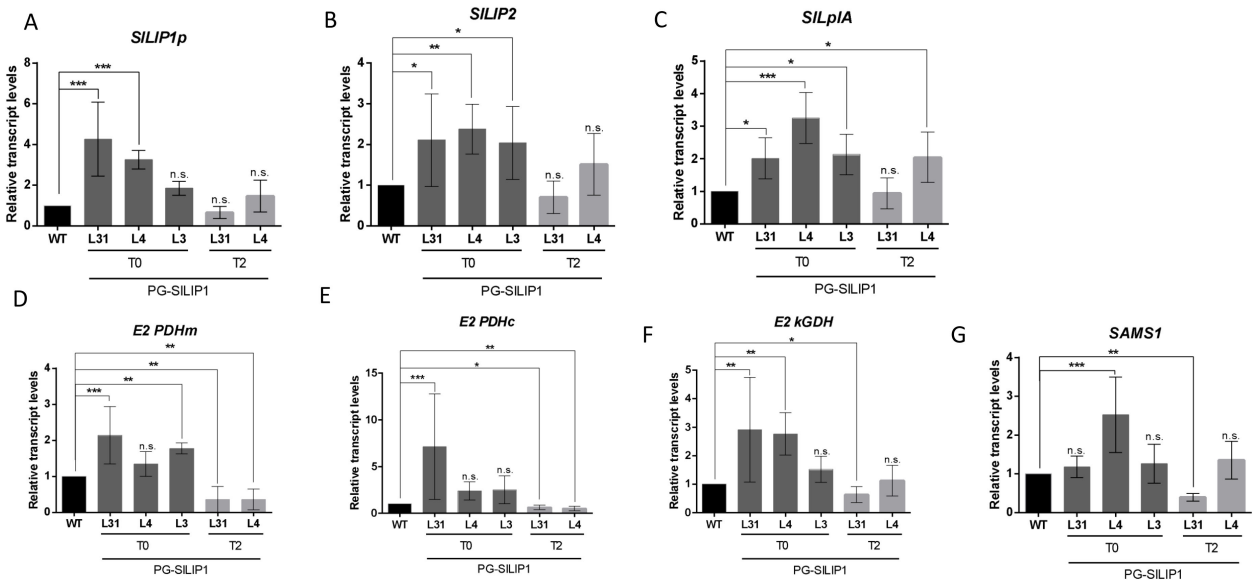


FIGURE 3
Relative transcript levels of genes involved in LA metabolism in fruits of T0 and T2 PG-SlLIP1 transgenic tomato lines. The analyzed genes are: **(A)** Plastid lipoyl synthase (*SlLIP1p*): an isoform of lipoyl synthase that contains a signal peptide targeting to chloroplasts. **(B)** Octanoyl transferase (*SlLIP2*): participates in *de novo* LA biosynthesis. **(C)** Lipoate protein ligase (*SlLplA*): involved in the salvage pathway. **(D)** Mitochondrial E2 subunit of pyruvate dehydrogenase (*E2 PDHm*). **(E)** Plastidial E2 subunit of pyruvate dehydrogenase (*E2 PDHc*). **(F)** E2 subunit of alpha-keto glutarate dehydrogenase (*E2 kGDH*). **(G)** SAM synthetase (*SAMS1*): involved in SAM biosynthesis. Transcript levels were normalized against *Actin7* and calibrated to WT levels. n=6 (T0 and T2) or 12 (WT) ± SD. $p \leq 0.05$ (*), $p \leq 0.005$ (**), $p \leq 0.0005$ (***) and not significant (n.s.).

accumulated more in all three transgenic lines than in WT. Erythrose 4-phosphate showed similar levels in L31 and L3 fruits.

The changes in metabolite accumulation across the different transgenic lines provide insights into the metabolic pathways that may be altered due to the overexpression of *SlLIP1*. To explore this, a pathway enrichment analysis was performed by comparing metabolites of each transgenic line with WT (Table 2). The pentose phosphate pathway, starch and sucrose metabolism,

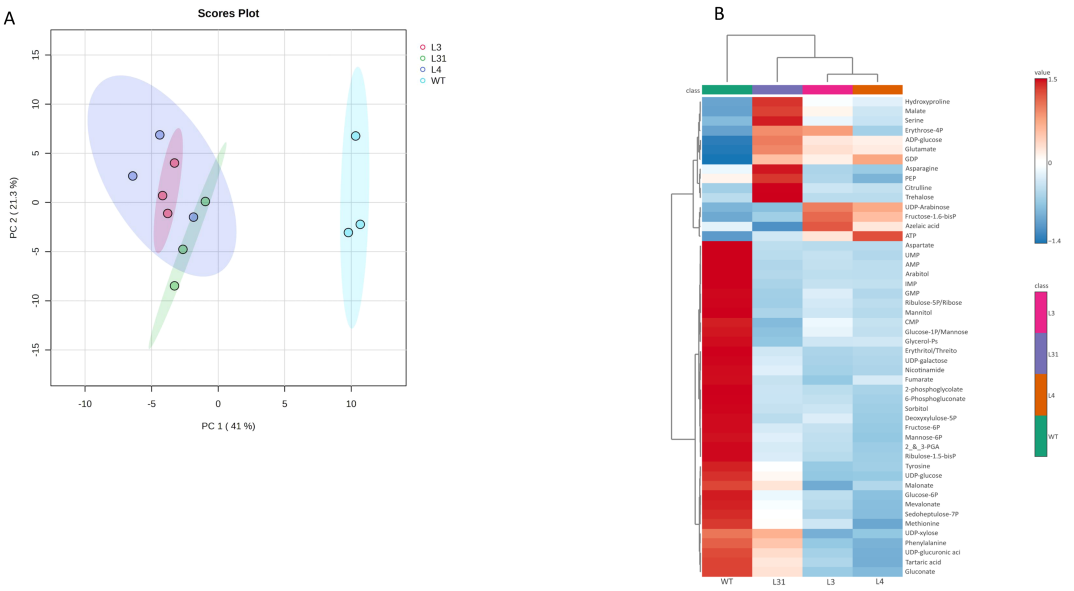


FIGURE 4
Targeted metabolomic analysis of fruits of T0 PG-SlLIP1 transgenic tomato lines. **(A)** Principal component analysis (PCA) was performed using detected metabolites as the predictor variable and genotype as the response variable. **(B)** Heatmap analysis of the top 50 significantly-altered metabolites identified by ANOVA. Columns represent the average of three biological replicates for each genotype (WT and PG-SlLIP1 lines L31, L4 and L3). The distance measure for clustering was Euclidean, and the clustering algorithm used was Ward.

TABLE 2 Metabolic pathways showing statistical differences in pathway enrichment analysis comparing WT to PG-SlLIP1 transgenic fruit metabolites.

Pathway Name	Match Status			*p*		
	L3	L4	L31	L3	L4	L31
Pentose phosphate pathway	6 of 19	6 of 19	6 of 19	4.92e-5	1.2e-03	1.2e-04
Starch and sucrose metabolism	9 of 22	9 of 22	10 of 22	6.83e-4	0.014	3.1e-04
Fructose and mannose metabolism	6 of 18	6 of 18	6 of 18	1.50e-4	1.2e-03	0.005
Alanine, aspartate and glutamate metabolism	9 of 22	9 of 22	9 of 22	0.007	0.038	0.007
Amino sugar and nucleotide sugar metabolism	14 of 52	14 of 52	14 of 52	5.29e-4	0.003	0.02

fructose and mannose metabolism, alanine, aspartate and glutamate metabolism, and amino sugar and nucleotide sugar metabolism were the five metabolic pathways that were most represented in the three transgenic lines compared to WT fruits. The *p*-values from the enrichment analysis were statistically significant in all cases. This information was used to create a map illustrating the effect of *SlLIP1* overexpression in tomato fruits compared with WT (Figure 5).

The main differences were observed in the pentose phosphate pathway, where four metabolites (ribulose 5-phosphate, ribulose 1,5-biphosphate, sedoheptulose 7-phosphate and erythritol/threitol) accumulated more in WT fruits, while erythrose 4-phosphate was more abundant in L3 fruits than in WT. Conversely, UDP-glucose, glucose 6-phosphate, and fructose 6-phosphate, related to starch and sucrose metabolism, along with metabolites related to amino sugar and nucleotide sugar metabolism (e.g., UDP-glucose, UDP-glucuronic acid, mannose 6-phosphate, UDP-xylose, fructose 6-phosphate, and UDP-galactose) accumulated more in WT fruits than in transgenic fruits (Figure 5). In contrast, trehalose 6-phosphate was more abundant in L4 fruits than WT, similar to ADP-glucose and UDP-arabinose. Regarding fructose and mannose metabolism, sorbitol, mannose 6-phosphate, and fructose 6-phosphate accumulated more in WT fruits, while fructose 1,5-biphosphate experienced the opposite. In alanine, aspartate and glutamate metabolism, aspartate and fumarate were more represented in WT fruits, whereas asparagine, glutamate (L31, L4 and L3), glutamine (L31), and GABA and proline (L31) were more abundant in transgenic lines than WT fruits (Figure 5). Therefore, these data highlight the central role of lipoylation in both primary and secondary metabolic pathways, such that overlipoylation results in a series of changes in polar metabolites.

4 Discussion

Lipoyl synthase is the enzyme that catalyzes the final step in LA biosynthesis by adding the two sulfur atoms to four octanoylated targets, PDH, kGDH, BCDH and GDC. Consistently, the effects of manipulating the expression of this enzyme throughout the whole plant are expected to be severe. Indeed, Ewald et al. (2014a) and Ewald et al. (2014) demonstrated that interruption of LIP2, LplA, LIP1p and LIP2p through T-DNA-mediated insertion in *Arabidopsis* resulted in a lethal phenotype when the T-DNA insertion was homozygous. This suggests that manipulating LA metabolism hinders normal plant growth. In this study, we used the PG promoter to overexpress tomato lipoyl synthase (*SlLIP1*) specifically during the breaker and pink fruit stages (Supplementary Figure 1B), thereby avoiding potential negative effects on vegetative organs. The increase in *SlLIP1* expression driven by the PG promoter in T0 and T2 generations was approximately 2- to 3- fold higher than in WT plants (Figure 2), similar to the ranges described by Arias et al. (2021) using the same pCP vector system. Additionally, no additional *SlLIP1* transcripts were detected in fully-expanded leaves of transformed plants (Supplementary Figure 2). The fruit-specific promoter chosen for this study corresponds to *SlPG14* (Solyc10g080210.1.1); indeed, a detailed analysis of the expression pattern of this gene shows that it is fruit-specific (Ke et al., 2018), consistent with publicly-available data (bar.utoronto.ca/efp2/Tomato/Tomato_eFPBrowser2).

The greater levels of *SlLIP1* transcripts triggered concomitant rises in both bound and unbound LA (Figure 2). These increases in LA content were associated with modifications in the transcript levels of various genes involved in LA biosynthesis (Figures 3A, B, C). To the best of our knowledge, there is no existing evidence detailing how this feedback mechanism operates between LIP1 and other genes in the LA biosynthesis pathway. However, in this study, we propose that elevated SlLIP1 activity results in depletion of octanoylated proteins which act as reaction substrates for lipoylation, leading to a positive feedback signal to upregulate SlLIP2 activity. This upregulation would enhance the transfer of octanoyl groups from acyl carrier protein (ACP) to target proteins, thus providing more potential substrates for SlLIP1. In the case of SlLplA, an enzyme associated with the salvage pathway that directly adds exogenous octanoyl or lipoyl groups onto target domains (Ewald et al., 2014a), we observed increased transcript levels in all three transgenic lines (Figure 3C). Therefore, we suggest a similar mechanism to that of SlLIP2, where substrate depletion caused by SlLIP1 activity upregulates SlLplA transcription through positive feedback, facilitating the uptake of free octanoyl and/or lipoyl from the medium and its transfer to target proteins.

Regarding the target proteins, we observed a significant increase in the transcript levels of mitochondrial and plastidial *E2-PDH*, as well as *E2 kGDH*, in some transgenic lines in the T0 generation. However, there were no differences observed in the T2 generation (Figures 3D–F). This is noteworthy, as these subunits must be lipoylated for them to be active (Douce et al., 2001; Mooney et al.,

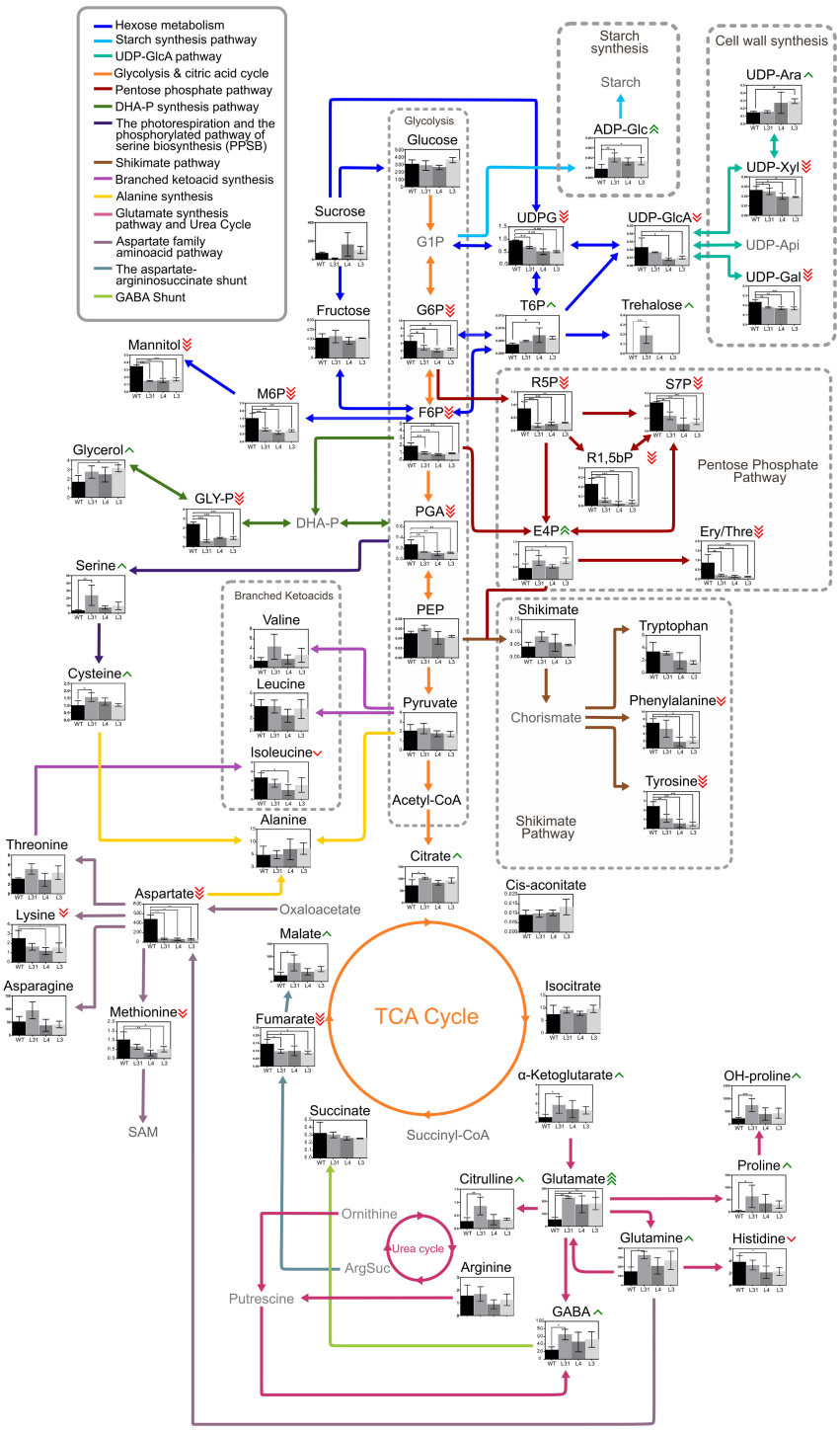


FIGURE 5
Metabolic comparison of fruits of PG-SlLIP1 transgenic and WT tomato lines. The graphical representation illustrates the variance in the metabolic profile of each transgenic fruit (n=3 ± SD) compared to WT (n=3 ± SD). Green arrows beside metabolite names indicate an increase in accumulation in transgenic fruits compared to WT, while red arrows indicate a decrease, according to the t-test with significance set at $p \leq 0.05$ (*), (**) and (***). Gray metabolites were not measured but are included in the scheme to facilitate understanding. G1P, Glucose 1-phosphate; UDPG, UDP-glucose; UDP-GlcA, UDP-glucuronic acid; UDP-Ara, UDP-arabinose; UDP-Xyl, UDP-xylose, UDP-Api, UDP-apiose; UDP-Gal, UDP-galactose; ADP-Glc, ADP-glucose; G6P, Glucose 6-phosphate; T6P, Trehalose 6-phosphate; F6P, Fructose 6-phosphate; M6P, Mannose 6-phosphate; R5P, Ribulose 5-phosphate; S7P, Sedoheptulose 7-phosphate; R1,5-bP, Ribulose 1,5-biphosphate; E4P, Erythrose 4-phosphate; Ery/Thre, Erythritol/Threitol; PGA, 2–3 phosphoglyceric acid; DHA-P, Dihydroxyacetone phosphate; GLY-P, Glycerol phosphate; PEP, Phosphoenolpyruvate; SAM, S-adenosyl methionine; GABA, gamma-aminobutyric acid; OH-proline, hydroxyproline; ArgSuc, Argininosuccinate.

2002; Perham, 2000). In this context, there may be a positive feedback loop between LA biosynthesis and the transcription of genes encoding these target proteins. A similar phenomenon was observed in hepatocytes supplemented with LA, where researchers found an increase in mitochondrial PDH activity due to inhibition of its antagonist, pyruvate dehydrogenase kinase (PDK), through an unknown mechanism (Korotchkina et al., 2004). In contrast to the T0 lines, in the T2 generation the transcript levels of *E2-PDHm*, *E2-PDHc* and *E2 kGDH* were generally lower than in WT fruits. Along these lines, it has been demonstrated that there is an unlipoylated pool of E2 subunits of PDH and kGDH proteins in normal conditions in *Arabidopsis* (Ewald et al., 2014a). Therefore, we suggest that in the T2 generation, due to the stable overexpression of *SlLIP1* (Figure 2A) and the lower levels of transcripts of the genes encoding these lipoylated targets (Figures 3D–F; *E2-PDHm*, *E2-PDHc* and *E2 kGDH*), that the pool of unlipoylated E2 subunits has been reduced in the transgenic fruits. Indeed, and in support of this hypothesis, the immunoblot analysis shows that there are overall higher levels of bound LA in PG-SlLIP1 tomatoes compared to untransformed controls (Figure 2B).

Using targeted metabolome analysis of transgenic compared to WT fruits, we observed that the three PG-SlLIP1 lines were similar to each other but distinct from WT in terms of polar metabolite accumulation (Figure 4). Based on this data, we were able to identify which metabolic pathways were most affected by the overexpression of *SlLIP1* (Table 2) and subsequently created a diagram illustrating central metabolism and the metabolite accumulation in each genotype (Figure 5). Regarding starch and sucrose metabolism, glucose 6-phosphate (G6P) and UDP-glucose (UDPG) are two essential compounds in central metabolism within fruits. G6P belongs to the hexose pool and serves as a precursor for the pentose phosphate pathway, providing reducing power within the cells (Jiang et al., 2022). As for UDPG, it plays a key role in sucrose synthesis in source organs, as well as in the formation of cellulose and callose in the apoplast (Kleczkowski et al., 2010). Additionally, it is involved in the synthesis of hemicelluloses and pectins in the cell wall (Amor et al., 1995; Feingold and Barber, 1990; Gibeaut, 2000; Johansson et al., 2002; Kleczkowski, 1994; Reiter, 2008). A mutant of the *ugpA* gene in *Arabidopsis*, which encodes UDP-glucose pyrophosphorylase responsible for synthesizing UDPG from glucose 1-phosphate, exhibited 50% less seed production compared to WT plants (Meng et al., 2009). This finding is consistent with observations in PG-SlLIP1 lines, which showed reduced or absent seed production (Table 1). Metabolomic data indicated that G6P and UDPG levels were lower in PG-SlLIP1 fruits compared to WT. Additionally, trehalose 6-phosphate (T6P) accumulated more in PG-SlLIP1 L4 and tended to be higher in L31 and L3 fruits (Figure 5). G6P and UDPG serve as precursors to T6P (Kerbler et al., 2023), which is subsequently dephosphorylated by T6P phosphatase to produce trehalose. Meanwhile, fatty acid synthesis is regulated by the WRI1 transcription factor (Zhai et al., 2018). When phosphorylated by SnRK1 kinase, WRI1 cannot translocate to the nucleus. T6P can bind to the SnRK1 kinase catalytic subunit, inhibiting its activation, thus allowing WRI1 to enter the nucleus and activate fatty acid synthesis (Zhai et al., 2018). Since LA is derived from fatty acid biosynthesis (octanoylated-ACP) when synthesized *de novo*, we propose that the observed increase in LA is, at least in part, being sustained by the elevated T6P levels.

As previously mentioned, SAM is a key molecule in LIP1 activity, as one molecule of SAM is consumed for every two sulfur atoms inserted into octanoylated proteins (Harmer et al., 2014). The biosynthesis of SAM is catalyzed by SAMS, with methionine serving as a precursor. Previous studies have mutated one of the four *SAMS* genes in *Arabidopsis*, leading to an increase in methionine levels. Consequently, mutant plants exhibit a severe growth defect due to reduced SAM levels, which are essential for methylation reactions (Goto et al., 2002; Jander and Joshi, 2010; Shen et al., 2002). Lysine, or a related metabolite, regulate SAMS at the transcriptional level and modulate its enzymatic activity (Hacham et al., 2007). Inhibition of SAMS reduces SAM biosynthesis and consequently leads to methionine accumulation, as SAM is a negative regulator of cystathionine gamma-synthase, which is involved in the initial step of methionine biosynthesis (Jander and Joshi, 2010). Our results showed a decrease in lysine levels in transgenic tomato fruits (Figure 5), suggesting that this may enable SAMS to become more active, thereby consuming more methionine to generate additional SAM for elevated LIP1 activity.

The metabolomic data showed an increase in glutamate levels in the three transgenic lines, along with a statistically significant increase in glutamine and GABA levels in L31, and a trend towards increased levels in L4 and L3 (Figure 5). These results suggest an activation of the "GABA shunt", which bypasses two steps of the TCA cycle, specifically succinyl-CoA ligase and kGDH (Bouché and Fromm, 2004), the latter of which is a lipoylated target. This pathway involves three key enzymes: glutamate decarboxylase (GAD), GABA transaminase (GABA-T) and succinic semialdehyde dehydrogenase (SSADH) (Michaeli and Fromm, 2015). The metabolomic results indicate a potential activation of the shunt, possibly due to deregulation of kGDH as a result of higher LIP1 activity. Another possible explanation is that the transgenic fruits were responding to oxidative stress, which could have triggered an increase in Ca^{2+} levels, subsequently activating the GABA shunt, as GAD is regulated by intracellular Ca^{2+}. Notably, changes in intracellular Ca^{2+} levels are known to be related to ROS response and signaling in plants (Ansari et al., 2021). Indeed, high proline content is an indicator of the stress response (Kavi Kishor and Sreenivasulu, 2014). This was corroborated by our metabolomic data, where L31 exhibited statistically-significantly higher proline levels compared to WT (Figure 5). Similar findings were reported by (Ewald et al., 2014a), where *LplA* RNAi plants showed elevated proline and GABA levels, indicating metabolic stress due to decreased E2-PDH lipoylation in leaves and roots in *Arabidopsis*. In contrast, our results showed higher levels of lipoylated proteins, suggesting that both greater and lower lipoylation may induce metabolic stress in plant cells. Nevertheless, the metabolomic analysis also demonstrates the uniformity of the mature red fruits that were chosen for these assays, as the levels of Glc and Fru [sugars known to accumulate in mature red fruits, Figure 5; Carrari and Fernie (2006)] were identical between transgenic and WT fruits, consistent with our expression data showing that transcript levels of *SlPG14* and

SlNAP2, genes whose expression correlates positively with fruit development (Ke et al., 2018; Kou et al., 2016) are similar between transgenic and WT lines (Supplementary Figures 1A, B).

In conclusion, the use of the PG promoter has made it possible to obtain stably transformed tomato plants whose plant phenotypes did not differ significantly from WT. Using this promoter, *LIP1* was overexpressed in mature fruits, but not in leaves. Overexpression of *SlLIP1* was observed to not only increase bound and unbound levels of LA but additionally it raised the expression of many other genes associated with LA metabolism. Higher *SlLIP1* expression was also observed to have effects on metabolite abundance associated with various metabolic routes, among which we highlight general reductions in metabolites participating in glycolysis and the pentose phosphate pathway, and a marked increase in glutamate and some of its downstream products such as glutamine and GABA, pointing towards activation of the GABA shunt. The applications of this work mean that SlLIP1 is a potential target for increasing LA levels in antioxidant-poor crops in an organ-specific manner, without incurring substantial physiological costs on their performance.

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

MC: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing, Data curation, Investigation, Methodology. FU: Formal analysis, Investigation, Methodology, Writing – review & editing. DA: Formal analysis, Investigation, Methodology, Writing – review & editing. PC: Formal analysis, Writing – original draft, Writing – review & editing. MH: Formal analysis, Writing – review & editing, Conceptualization, Funding acquisition, Resources, Supervision, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was funded by doctoral scholarships from the Fundación María Ghilardi Venegas (MPC, PC), ANID Doctoral Scholarship 21210768 (FU), ANID Fondecyt Postdoctoral 3220609 (DA), ANID Fondecyt Regular 1231417 (MH) and ANID-Anillo ACT210025 project (MH).

Acknowledgments

We thank all members of CBMV for helpful discussions, especially Simon Miranda and Jorge Araya. The authors acknowledge the BioAnalytical Facility at the University of North Texas for the support with mass spectrometry analyses during this work.

Conflict of interest

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

Generative AI statement

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1545011/full#supplementary-material

References

Alonso, A. P., Piasecki, R. J., Wang, Y., LaClair, R. W., and Shachar-Hill, Y. (2010). Quantifying the labeling and the levels of plant cell wall precursors using ion chromatography tandem mass spectrometry. *Plant Physiol.* 153, 915–924. doi: 10.1104/pp.110.155713

Amor, Y., Haigler, C. H., Johnson, S., Wainscott, M., and Delmer, D. P. (1995). A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. *Proc. Natl. Acad. Sci.* 92, 9353–9357. doi: 10.1073/pnas.92.20.9353

Ansari, M. I., Jalil, S. U., Ansari, S. A., and Hasanuzzaman, M. (2021). GABA shunt: a key-player in mitigation of ROS during stress. *Plant Growth Regul.* 94, 131–149. doi: 10.1007/s10725-021-00710-y

Araya-Flores, J., Miranda, S., Covarrubias, M. P., Stange, C., and Handford, M. (2020). Solanum lycopersicum (tomato) possesses mitochondrial and plastidial lipoyl synthases capable of increasing lipoylation levels when expressed in bacteria. *Plant Physiol. Biochem.* 151, 264–270. doi: 10.1016/j.plaphy.2020.03.031

Arias, D., Arenas-M, A., Flores-Ortiz, C., Peirano, C., Handford, M., and Stange, C. (2021). Daucus carota DcPSY2 and DcLCYB1 as tools for carotenoid metabolic engineering to improve the nutritional value of fruits. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.677553

Arias, D., Ortega, A., González-Calquin, C., Quiroz, L. F., Moreno-Romero, J., Martínez-García, J. F., et al. (2022). Development and carotenoid synthesis in dark-

grown carrot taproots require *PHYTOCHROME RAPIDLY REGULATED1*. *Plant Physiol.* 189, 1450–1465. doi: 10.1093/plphys/kiac097

Bokhale, M., Mwaba, I., and Allie, F. (2023). The selection and validation of reference genes for quantitative real-time PCR studies in near-isogenic susceptible and resistant tomato lines, infected with the geminivirus tomato curly stunt virus. *PloS One* 18. doi: 10.1371/journal.pone.0284456

Borcea, V., Nourooz-Zadeh, J., Wolff, S. P., Klevesath, M., Hofmann, M., Urich, H., et al. (1999). α-lipoic acid decreases oxidative stress even in diabetic patients with poor glycemic control and albuminuria. *Free Radical Biol. Med.* 26, 1495–1500. doi: 10.1016/S0891-5849(99)00011-8

Bouché, N., and Fromm, H. (2004). GABA in plants: Just a metabolite? *Trends Plant Sci.* 9, 110–115. doi: 10.1016/j.tplants.2004.01.006

Carrari, F., and Fernie, A. (2006). Metabolic regulation underlying tomato fruit development. *J. Exp. Bot.* 57, 1883–1897. doi: 10.1093/jxb/erj020

CFR (1991). *Code of Federal Regulations (CFR). United States standards for grades of fresh tomatoes* (Washington, DC: USDA Agricultural Marketing Service).

Chen, E., Hu, H., Yang, X., Li, D., Wei, Q., Zhou, F., et al. (2021). GhLIP1, a lipoic acid synthase gene, negatively regulates leaf senescence in cotton. *Plant Growth Regul.* 94, 73–85. doi: 10.1007/s10725-021-00697-6

Cocuron, J. C., Anderson, B., Boyd, A., and Alonso, A. P. (2014). Targeted metabolomics of *Physaria fendleri*, an industrial crop producing hydroxy fatty acids. *Plant Cell Physiol.* 55, 620–633. doi: 10.1093/pcp/pcu011

Cronan, J. E. (2014). "*Escherichia coli* as an experimental organism," in *Encyclopedia of Life Sciences* (Wiley). doi: 10.1002/9780470015902.a0002026.pub2

Cruz-Mendívil, A., Rivera-López, J., Germán-Báez, L. J., López-Meyer, M., Hernández-Verdugo, S., López-Valenzuela, J. A., et al. (2011). A simple and efficient protocol for plant regeneration and genetic transformation of tomato cv. micro-tom from leaf explants. *HortScience* 46, 1655–1660. doi: 10.21273/hortsci.46.12.1655

Douce, R., Bourguignon, J., Neuburger, M., and Rébeillé, F. (2001). The glycine decarboxylase system: a fascinating complex. *Trends Plant Sci.* 6, 167–176. doi: 10.1016/S1360-1385(01)01892-1

Ewald, R., Hoffmann, C., Florian, A., Neuhaus, E., Fernie, A. R., and Bauwe, H. (2014a). Lipoate-protein ligase and octanoyltransferase are essential for protein lipoylation in mitochondria of arabidopsis. *Plant Physiol.* 165, 978–990. doi: 10.1104/pp.114.238311

Ewald, R., Hoffmann, C., Neuhaus, E., and Bauwe, H. (2014). Two redundant octanoyltransferases and one obligatory lipoyl synthase provide protein-lipoylation autonomy to plastids of Arabidopsis. *Plant Biol.* 16, 35–42. doi: 10.1111/plb.12028

Feingold, D. S., and Barber, G. A. (1990). "Nucleotide sugars," in *Methods in Plant Biochemistry*, vol. 2 . Ed. P. Dey, Academic Press, 39–78. doi: 10.1016/B978-0-12-461012-5.50008-2

Gibeaut, D. M. (2000). Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides. *Plant Physiol. Biochem.* 38, 69–80. doi: 10.1016/S0981-9428(00)00167-4

Gilgun-Sherki, Y., Melamed, E., and Offen, D. (2001). Oxidative stress induced-neurodegenerative diseases: The need for antioxidants that penetrate the blood brain barrier. *Neuropharmacology* 40, 959–975. doi: 10.1016/S0028-3908(01)00019-3

Gorąca, A., Huk-Kolega, H., Piechota, A., Kleniewska, P., Ciejka, E., and Skibska, B. (2011). Lipoic acid - Biological activity and therapeutic potential. *Pharmacol. Rep.* 63, 849–858. doi: 10.1016/S1734-1140(11)70600-4

Gorcek, Z., and Erdal, S. (2015). Lipoic acid mitigates oxidative stress and recovers metabolic distortions in salt-stressed wheat seedlings by modulating ion homeostasis, the osmo-regulator level and antioxidant system. *J. Sci. Food Agric.* 95, 2811–2817. doi: 10.1002/jsfa.7020

Goto, D. B., Ogi, M., Kijima, F., Kumagai, T., Werven, F., Onouchi, H., et al. (2002). A single-nucleotide mutation in a gene encoding S-adenosylmethionine synthetase is associated with methionine over-accumulation phenotype in *Arabidopsis thaliana*. *Genes Genet. Syst.* 77, 89–95. doi: 10.1266/ggs.77.89

Hacham, Y., Song, L., Schuster, G., and Amir, R. (2007). Lysine enhances methionine content by modulating the expression of *S*-adenosylmethionine synthase. *Plant J.* 51, 850–861. doi: 10.1111/j.1365-313X.2007.03184.x

Harmer, J. E., Hiscox, M. J., Dinis, P. C., Fox, S. J., Iliopoulos, A., Hussey, J. E., et al. (2014). Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions. *Biochem. J.* 464, 123–133. doi: 10.1042/BJ20140895

Jander, G., and Joshi, V. (2010). Recent progress in deciphering the biosynthesis of aspartate-derived amino acids in plants. *Mol. Plant* 3, 54–65. doi: 10.1093/mp/ssp104

Jiang, Z., Wang, M., Nicolas, M., Ogé, L., Pérez-Garcia, M.-D., Crespel, L., et al. (2022). Glucose-6-phosphate dehydrogenases: The hidden players of plant physiology. *Int. J. Mol. Sci.* 23, 16128. doi: 10.3390/ijms232416128

Jimenez-Lopez, C., Fraga-Corral, M., Carpena, M., García-Oliveira, P., Echave, J., Pereira, A. G., et al. (2020). Agriculture waste valorisation as a source of antioxidant phenolic compounds within a circular and sustainable bioeconomy. *Food Funct.* 11, 4853–4877. doi: 10.1039/D0FO00937G

Jimenez-Lopez, J. C., Wang, X., Kotchoni, S. O., Huang, S., Szymanski, D. B., and Staiger, C. J. (2014). Heterodimeric capping protein from *Arabidopsis* is a membrane-associated, actin-binding protein. *Plant Physiol.* 3, 1312–1328. doi: 10.1104/pp.114.242487

Johansson, H., Sterky, F., Amini, B., Lundeberg, J., and Kleczkowski, L. A. (2002). Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation. *Biochim. Biophys. Acta (BBA) - Gene Structure Expression* 1576, 53–58. doi: 10.1016/S0167-4781(02)00292-0

Kavi Kishor, P. B., and Sreenivasulu, N. (2014). Is proline accumulation *per se* correlated with stress tolerance or is proline homeostasis a more critical issue? *Plant Cell Environ.* 37, 300–311. doi: 10.1111/pce.12157

Ke, X., Wang, H., Li, Y., Zhu, B., Zang, Y., He, Y., et al. (2018). Genome-aide identification and analysis of polygalacturonase genes in *Solanum lycopersicum*. *Int. J. Mol. Sci.* 19, 2290–2318. doi: 10.3390/ijms19082290

Kerbler, S. M., Armijos-Jaramillo, V., Lunn, J. E., and Vicente, R. (2023). The trehalose 6-phosphate phosphatase family in plants. *Physiologia Plantarum* 175. doi: 10.1111/ppl.14096

Kleczkowski, L. A. (1994). Glucose activation and metabolism through UDP-glucose pyrophosphorylase in plants. *Phytochemistry* 37, 1507–1515. doi: 10.1016/S0031-9422(00)89568-0

Kleczkowski, L. A., Kunz, S., and Wilczynska, M. (2010). Mechanisms of UDP-glucose synthesis in plants. *Crit. Rev. Plant Sci.* 29, 191–203. doi: 10.1080/07352689.2010.483578

Korotchkina, L. G., Sidhu, S., and Patel, M. S. (2004). R-Lipoic acid inhibits mammalian pyruvate dehydrogenase kinase. *Free Radical Res.* 38, 1083–1092. doi: 10.1080/10715760400004168

Kou, X., Liu, C., Han, L., Wang, S., and Xue, Z. (2016). NAC transcription factors play an important role in ethylene biosynthesis, reception and signaling of tomato fruit ripening. *Mol. Genet. Genomics* 291, 1205–1217. doi: 10.1007/s00438-016-1177-0

Lu, S. C. (2000). S-adenosylmethionine. *Int. J. Biochem. Cell Biol.* 32, 391–395. doi: 10.1016/S1357-2725(99)00139-9

Lupo, M. P. (2001). Antioxidants and vitamins in cosmetics. *Clinics Dermatol.* 19, 467–473. doi: 10.1016/S0738-081X(01)00188-2

Martins-Noguerol, R., Moreno-Pérez, A. J., Sebastien, A., Troncoso-Ponce, M. A., Garcés, R., Thomasset, B., et al. (2020). Impact of sunflower (*Helianthus annuus* L.) plastidial lipoyl synthases genes expression in glycerolipids composition of transgenic *Arabidopsis* plants. *Sci. Rep.* 10, 1–17. doi: 10.1038/s41598-020-60686-z

Matsugo, S., Konishi, T., Matsuo, D., Tritschler, H. J., and Packer, L. (1996). Reevaluation of superoxide scavenging activity of dihydrolipoic acid and its analogues by chemiluminescent method using 2-methyl-6-[p-methoxyphenyl]-3,7-dihydroimidazo-[1,2-a]pyrazine-3-one (MCLA) as a superoxide probe. *Biochem. Biophys. Res. Commun.* 227, 216–220. doi: 10.1006/bbrc.1996.1492

Meng, M., Geisler, M., Johansson, H., Harholt, J., Scheller, H. V., Mellerowicz, E. J., et al. (2009). UDP-Glucose pyrophosphorylase is not rate limiting, but is essential in *Arabidopsis*. *Plant Cell Physiol.* 50, 998–1011. doi: 10.1093/pcp/pcp052

Michaeli, S., and Fromm, H. (2015). Closing the loop on the GABA shunt in plants: are GABA metabolism and signaling entwined? *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00419

Montgomery, J., Pollard, V., Deikman, J., and Fischer, R. L. (1993). Positive and negative regulatory regions control the spatial distribution of polygalacturonase transcription in tomato fruit pericarp. *Plant Cell* 5, 1049–1062. doi: 10.2307/3869626

Mooney, B. P., Miernyk, J. A., and Randall, D. D. (2002). The complex fate of α-ketoacids. *Annu. Rev. Plant Biol.* 53, 357–375. doi: 10.1146/annurev.arplant.53.100301.135251

Neha, K., Haider, M. R., Pathak, A., and Yar, M. S. (2019). Medicinal prospects of antioxidants: A review. *Eur. J. Medicinal Chem.* 178, 687–704. doi: 10.1016/j.ejmech.2019.06.010

Nilo-Poyanco, R., Olivares, D., Orellana, A., Hinrichsen, P., and Pinto, M. (2013). Proteomic analysis of grapevine (*Vitis vinifera* L.) leaf changes induced by transition to autotrophy and exposure to high light irradiance. *J. Proteomics* 91, 309–330. doi: 10.1016/j.jprot.2013.07.004

Packer, L., Kraemer, K., and Rimbach, G. (2001). Molecular aspects of lipoic acid in the prevention of diabetes complications. *Nutrition* 17, 888–895. doi: 10.1016/S0899-9007(01)00658-X

Packer, L., Witt, E. H., and Tritschler, H. J. (1995). Alpha-lipoic acid as a biological antioxidant. *Free Radical Biol. Med.* 19, 227–250. doi: 10.1016/0891-5849(95)00017-R

Pang, Z., Lu, Y., Zhou, G., Hui, F., Xu, L., Viau, C., et al. (2024). MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Res.* 52, W398–W406. doi: 10.1093/nar/gkae253

Perham, R. N. (2000). Swinging arms and swinging domains in multifunctional enzymes: Catalytic machines for multistep reactions. *Annu. Rev. Biochem.* 69, 961–1004. doi: 10.1146/annurev.biochem.69.1.961

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29 (9), e45. doi: 10.1093/nar/29.9.e45

Reiter, W. (2008). Biochemical genetics of nucleotide sugar interconversion reactions. *Curr. Opin. Plant Biol.* 11, 236–243. doi: 10.1016/j.pbi.2008.03.009

Saruhan Guler, N., Ozturk, K., Sezgin, A., Altuntas, C., Kadioglu, A., and Terzi, R. (2021). Alpha lipoic acid application promotes water-deficit tolerance by modulating osmoprotectant metabolism-related genes in maize. *Russian J. Plant Physiol.* 68, 1152–1160. doi: 10.1134/S1021443721060042

Sarvananda, L., Mendis, B., Madhuranga, H., and Premarathna, A. D. (2023). Mitochondrial fatty acid synthesis in plants: enzymes and biological roles. *Modern Phytomorphology* 17, 41–49. doi: 10.5281/zenodo.7939690

Sauter, M., Moffatt, B., Saechao, M. C., Hell, R., and Wirtz, M. (2013). Methionine salvage and S -adenosylmethionine: essential links between sulfur, ethylene and polyamine biosynthesis. *Biochem. J.* 451, 145–154. doi: 10.1042/BJ20121744

Scott, B. C., Aruoma, O. I., Evans, P. J., O'Neill, C., van der Vliet, A., Cross, C. E., et al. (1994). Lipoic and dihydrolipoic acids as antioxidants. A critical evaluation. *Free Radical Res.* 20, 119–133. doi: 10.3109/10715769409147509

Shahidi, F. (2000). Antioxidants in food and food antioxidants. *Nahrung/Food* 44, 158–163. doi: 10.1002/1521-3803(20000501)44:3<158::AID-FOOD158>3.0.CO;2-L

Shen, B., Li, C., and Tarczynski, M. C. (2002). High free-methionine and decreased lignin content result from a mutation in the *Arabidopsis S-adenosyl-L-methionine synthetase* 3 gene. *Plant J.* 29, 371–380. doi: 10.1046/j.1365-313X.2002.01221.x

Sun, Y., Xie, X., and Jiang, C.-J. (2024). Antioxidant agriculture for stress-resilient crop production: Field practice. *Antioxidants* 13, 164. doi: 10.3390/antiox13020164

Tovar-Méndez, A., Miernyk, J. A., and Randall, D. D. (2003). Regulation of pyruvate dehydrogenase complex activity in plant cells. *Eur. J. Biochem.* 270, 1043–1049. doi: 10.1046/j.1432-1033.2003.03469.x

Turrens, J. F. (2003). Mitochondrial formation of reactive oxygen species. *J. Physiol.* 552, 335–344. doi: 10.1111/j.1469-7793.2003.00335.x

Yasuno, R., and Wada, H. (2002). The biosynthetic pathway for lipoic acid is present in plastids and mitochondria in *Arabidopsis thaliana*. *FEBS Lett.* 517, 110–114. doi: 10.1016/S0014-5793(02)02589-9

Zhai, Z., Keereetaweep, J., Liu, H., Feil, R., Lunn, J. E., and Shanklin, J. (2018). Trehalose 6-phosphate positively regulates fatty acid synthesis by stabilizing WRINKLED1. *Plant Cell* 30, 2616–2627. doi: 10.1105/tpc.18.00521

Zhang, X., Bao, Z., Gong, B., and Shi, Q. (2020). S-adenosylmethionine synthetase 1 confers drought and salt tolerance in transgenic tomato. *Environ. Exp. Bot.* 179. doi: 10.1016/j.envexpbot.2020.104226

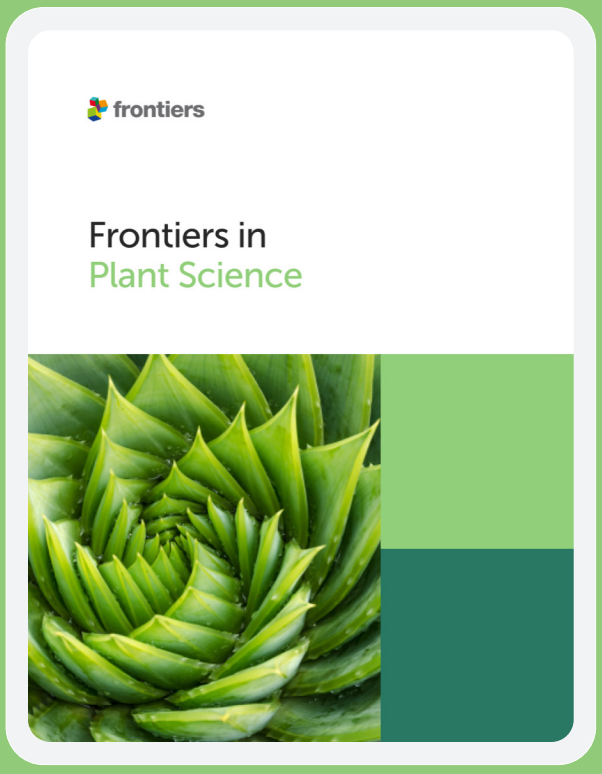
Frontiers in
Plant Science

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

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact