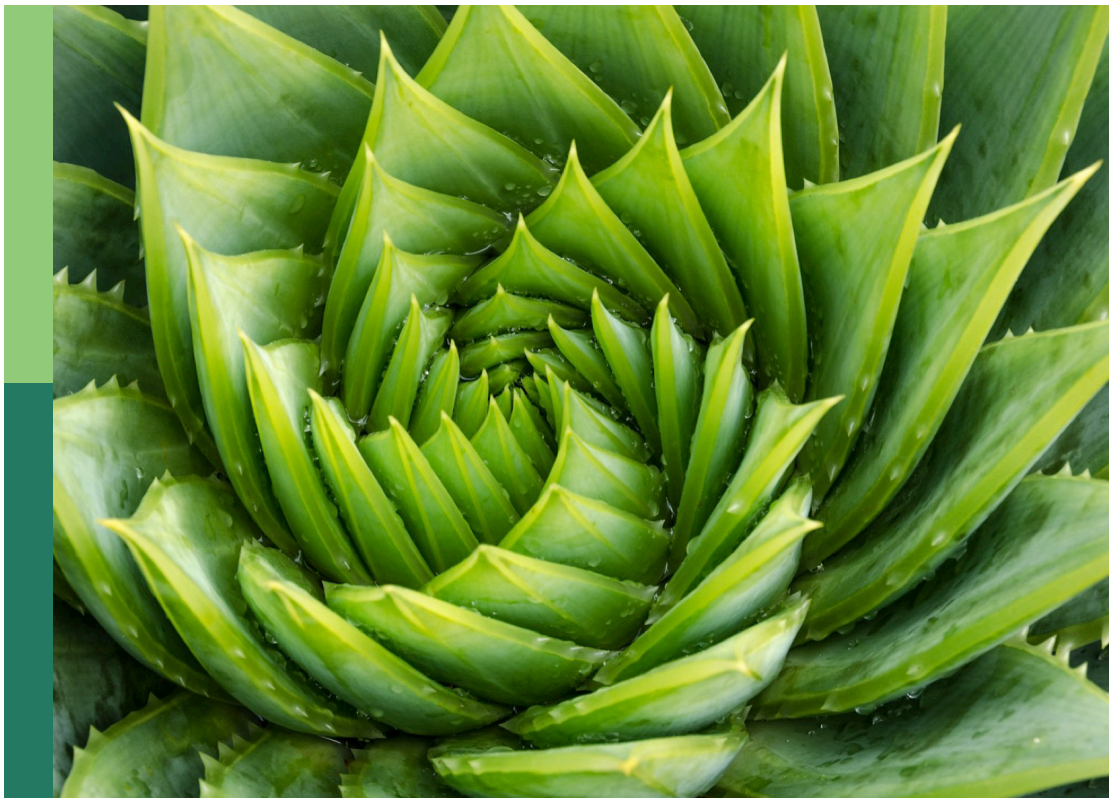
frontiers | Research Topics

Biosynthesis and regulation of plant specialized metabolisms

Edited by
Weiwei Zhang, Feng Xu and Moonhyuk Kwon

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-3992-7
DOI 10.3389/978-2-8325-3992-7

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

Biosynthesis and regulation of plant specialized metabolisms

Topic editors

Weiwei Zhang — Yangtze University, China

Feng Xu — Yangtze University, China

Moonhyuk Kwon — Gyeongsang National University, Republic of Korea

Citation

Zhang, W., Xu, F., Kwon, M., eds. (2023). *Biosynthesis and regulation of plant specialized metabolisms*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-3992-7

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 09 November 2023
DOI 10.3389/fpls.2023.1319666

OPEN ACCESS

EDITED AND REVIEWED BY
Yansheng Zhang,
Shanghai University, China

*CORRESPONDENCE
Feng Xu
xufeng@yangtzeu.edu.cn
Moonhyuk Kwon
mkwon@gnu.ac.kr

RECEIVED 11 October 2023
ACCEPTED 25 October 2023
PUBLISHED 09 November 2023

CITATION
Zhang W, Xu F and Kwon M (2023) Editorial: Biosynthesis and regulation of plant specialized metabolisms. *Front. Plant Sci.* 14:1319666. doi: 10.3389/fpls.2023.1319666

Editorial: Biosynthesis and regulation of plant specialized metabolisms

Weiwei Zhang[1], Feng Xu[1]* and Moonhyuk Kwon[2]*

[1]College of Horticulture and Gardening, Yangtze University, Jingzhou, China, [2]Division of Life Science, Anti-aging Bio Cell factory Regional Leading Research Center (ABC-RLRC), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University, Jinju, Republic of Korea

KEYWORDS

specialized metabolites, flavonoids, terpenoids, alkaloids, biosynthetic pathway, regulatory factor

Editorial on the Research Topic
Biosynthesis and regulation of plant specialized metabolisms

Plant-specialized metabolites such as flavonoids, polyketides, terpenoids, glucosinolates, and phenolic compounds, more commonly known as secondary metabolites are organic compounds produced via specialized metabolism. They are multifunctional metabolites that play an important role in the adaptation of plants to their environment, such as protection against natural enemies and pests, assistance in pollination, and protection against ultraviolet light damage. Based on the physiological activities, and pharmacological effects, plant-specialized metabolites have been widely used for the prevention and treatment of human diseases like tumors, aging, and cardiovascular disorders. The distribution of specialized metabolites in plants is unique to species, organs, tissues, and growth stages, resulting in complex and diverse biosynthetic pathways. The synthesis of plant-specialized metabolites is controlled by genetic material and regulated by a variety of biotic and abiotic factors in the environment.

During the growth and development of plants, they produce a large number of specialized metabolites that serve various cellular functions necessary for physiological processes. The synthesis and proper accumulation of these specialized metabolites are tightly controlled, and a complex regulatory network formed by enzyme genes and transcription factors (TFs) plays a key role in the metabolism of specialized metabolites. In this Research Topic, some of the articles provide some new insights into the regulation of the synthesis of different specialized metabolites. Wang et al. combined metabolomic and transcriptomic analyses to identify four structural genes involved in glucosinolate and soluble sugar biosynthesis (*BraA05gAOP1*, *BraA04gAOP4*, *BraA03gHT7*, and *BraA01gHT4*) and two transcription factors (*BraA01gCHR11* and *BraA07gSCL1*). Xu et al. discovered that *TcbZIP60* directly binds to the E-box/G-box motifs in the promoters of pyrethroid synthesizing genes *TcCHS* and *TcAOC* and activate their expression. Further studies revealed that transient overexpression of *TcbZIP60* increased the expression level of pyrethrin biosynthesis genes, resulting in a significant accumulation of pyrethrins. Conversely, silencing of *TcbZIP60* led to reduced pyrethrin

accumulation and the expression of related genes. These results suggest that *TcbZIP60* is able to regulate the biosynthesis of pyrethrins. Xu et al. revealed the relationship between jasmonic acid (JA) and aspalathoside VI biosynthesis in *Dipsacus asperoides*. They discovered that triterpenoids, JA and TFs together regulate aspalathoside VI biosynthesis in *Dipsacus asperoides*.

Plants achieve functional flexibility under the influence of environmental factors through the production of specialized metabolites. A large number of studies have shown that the production of these specialized metabolites is associated with stress and defense response signals and that changes in external conditions will cause changes in specialized metabolites at the physiological and molecular levels. With the development of biological techniques, the application of metabolomics and transcriptomics technologies has greatly facilitated these studies. Specialized metabolites usually act as defense molecules to protect plants under various adverse conditions. In this Research Topic, Park et al. found that anthocyanins and two flavonoids were significantly increased, while all types of caffeoylquinic acids (CQAs) and flavonols were significantly decreased in *Ligularia fischeri* under drought stress; RNA sequencing led to the final identification of potential drought-responsive genes including up-regulated flavonoid synthetase (*LfFNS*) and anthocyanidin-O-glucosyltransferase (*LfA5GT1*), which are important for flavonoids and anthocyanins production by *Ligularia fischeri* in response to drought stress. In another study, Lim et al. treated mungbean (*Vigna radiata*) sprouts with salt stress and found that the content of several phenylpropanoid metabolites including catechin, chlorogenic acid, p-coumaric acid, and ferulic acid was significantly increased in mungbean sprouts under mild salt stress conditions, and at the same time, several key enzyme genes and transcription factors involved in the biosynthesis of phenylpropanoid and flavonoid compounds were significantly upregulated. The composition of light can influence the metabolism of various specialized metabolites. Lim et al. evaluated the changes of metabolites in soybean seedlings under different light conditions, finding that the duration of light exposure regulates the content of kaempferol glycoside, the predominant flavonol in soybeans, with longer light exposure resulting in higher levels. Most isoflavones increased in response to red and blue lights, but daidzein increased only in response to red light. Blue light was observed to stimulate the accumulation of kaempferol-3-O-(2,6-dirhamnosyl)-galactoside more effectively than red light. Short-term red light irradiation (12 and 36 h) significantly increased the levels of malonyl daidzin and malonyl genistin, the predominant isoflavones in soybeans, along with higher expressions of flavonoid biosynthetic genes. The accumulation of specialized metabolites under stress conditions is regulated at the molecular level by various genes and transcription factors, including phytohormonal pathways. Xu et al. found that treatment of *Tanacetum cinerariifolium* with phytohormones (MeJA, abscisic acid) significantly increased the expression of *TcbZIP60*, which was further shown to regulate both terpenoid and JA pathways of pyrethroid biosynthesis in *Tanacetum cinerariifolium*. Similarly, Xu et al. found that the content of aspalathoside VI in *Dipsacus asperoides* could be increased by MeJA treatment, and the results of transcriptome analysis showed that MeJA treatment could promote the expression of genes such as *acetyl-CoA acetyltransferase* (*DaAACT*), *3-hydroxy-3-methylglutaryl-coenzyme A synthase* (*DaHMGCS*), and *3-hydroxy-3-methylglutaryl-coenzyme A reductase* (*DaHMGCR*) in triterpene biosynthesis pathway, and consequently promote the biosynthesis of aspalathoside VI.

In addition to the metabolism study, Zhang et al. verified the involvement of *CsPUPs* in the regulation of caffeine transport in tea trees by applying a heterologous expression system from yeast and Arabidopsis. On the other hand, obtaining high yields of valuable metabolites through synthetic biology is a goal we all aspire to. Nguyen et al. achieved efficient production of spirooxindoles by expressing a P450 enzyme from *Mitragyna speciosa* in *Saccharomyces cerevisiae*. This accomplishment paves the way for obtaining the well-known spectrum of spirooxindoles and their derivatives.

The articles in this Research Topic provide valuable insights into the biosynthesis and regulation of plant-specialized metabolites. There is a diversity of natural plant-specific metabolites, and the regulation mechanisms of their synthesis are unique and complicated. In the future, more studies should focus on the function of key genes of specific metabolite anabolism *in vitro* and *in vivo*, and reveal the synthesis mechanism and metabolic pathway of specific metabolites under different physiological environmental conditions.

Author contributions

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

Funding

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

Acknowledgments

We thank all the authors for contributing their scholarly work to our Research Topic, and we would like to express our deep gratitude to all reviewers for their valuable help in the evaluation of the 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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 17 October 2022
DOI 10.3389/fpls.2022.1030677

OPEN ACCESS

EDITED BY
Moonhyuk Kwon,
Gyeongsang National University,
South Korea

REVIEWED BY
Sang Un Park,
Chungnam National University,
South Korea
Yue Wang,
Zhejiang University, China
Sang Min Kim,
Korea Institute of Science and
Technology (KIST), South Korea

*CORRESPONDENCE
Jungmin Ha
j.ha@gwnu.ac.kr

SPECIALTY SECTION
This article was submitted to
Plant Metabolism and Chemodiversity,
a section of the journal
Frontiers in Plant Science

RECEIVED 29 August 2022
ACCEPTED 27 September 2022
PUBLISHED 17 October 2022

CITATION
Lim I, Kang M, Kim BC and Ha J (2022)
Metabolomic and transcriptomic
changes in mungbean (*Vigna
radiata* (L.) R. Wilczek) sprouts
under salinity stress.
Front. Plant Sci. 13:1030677.
doi: 10.3389/fpls.2022.1030677

Metabolomic and transcriptomic changes in mungbean (*Vigna radiata* (L.) R. Wilczek) sprouts under salinity stress

Insu Lim, Minseo Kang, Byeong Cheol Kim and Jungmin Ha*

Department of Plant Science, Gangneung-Wonju National University, Gangneung, South Korea

Mungbean (*Vigna radiata*) sprouts are consumed globally as a healthy food with high nutritional values, having antioxidant and anticancer capacity. Under mild salinity stress, plants accumulate more secondary metabolites to alleviate oxidative stress. In this study, metabolomic and transcriptomic changes in mungbean sprouts were identified using a reference cultivar, sunhwa, to understand the regulatory mechanisms of secondary metabolites in response to salinity stress. Under salinity conditions, the contents of phenylpropanoid-derived metabolites, including catechin, chlorogenic acid, isovitexin, *p*-coumaric acid, syringic acid, ferulic acid, and vitexin, significantly increased. Through RNA sequencing, 728 differentially expressed genes (DEGs) were identified and 20 DEGs were detected in phenylpropanoid and flavonoid biosynthetic pathways. Among them, 11 DEGs encoding key enzymes involved in the biosynthesis of the secondary metabolites that increased after NaCl treatment were significantly upregulated, including dihydroflavonol 4-reductase (log_2FC 1.46), caffeoyl-CoA O-methyltransferase (1.38), chalcone synthase (1.15), and chalcone isomerase (1.19). Transcription factor families, such as *MYB*, *WRKY*, and *bHLH*, were also identified as upregulated DEGs, which play a crucial role in stress responses in plants. Furthermore, this study showed that mild salinity stress can increase the contents of phenylpropanoids and flavonoids in mungbean sprouts through transcriptional regulation of the key enzymes involved in the biosynthetic pathways. Overall, these findings will provide valuable information for molecular breeders and scientists interested in improving the nutritional quality of sprout vegetables.

KEYWORDS

mungbean sprout, salinity stress, ultra-high-performance liquid chromatography, phenylpropanoid compound, RNA-seq, gene expression

Introduction

Mungbean (*Vigna radiata* (L.) R. Wilczek) is an important legume crop grown over 6 million hectares worldwide (about 8.5% of the global pulse cultivation area) (Gayacharan et al., 2020). Mature mungbean seeds are valuable food sources of protein and starch in many developing countries in Asia (Nair et al., 2013). Mungbean sprouts have long been consumed globally because of their excellent texture and abundant nutrients, including ascorbic acid, dietary fiber, essential amino acids, and vitamins (Gan et al., 2016; Ebert et al., 2017; Singh et al., 2017). In particular, six polyphenolic compounds, including caffeic acid, ferulic acid, gallic acid, *p*-coumaric acid, catechin, and rutin, have been identified in mungbean sprouts (Gan et al., 2016), which help maintain human health and prevent chronic diseases (Kim et al., 2012; Nderitu et al., 2013; Pudenz et al., 2014; Stagos, 2020). Owing to these nutritional values, mungbean sprout consumption has increased, and the quality of sprout vegetables has attracted consumers' attention globally.

Plants accumulate polyphenolic compounds as defense metabolites against harmful environments, and this accumulation varies depending on different stress exposures, such as high/low temperature, salinity, ultraviolet, and drought (Ksouri et al., 2008; Yang et al., 2018; Sharma et al., 2019; Chrysargyris et al., 2020). Under salinity conditions, osmotic, ionic, and oxidative stresses affect most plants, inducing morphological and biochemical changes (Yang and Guo, 2018). While severe salinity disrupts general plant development through ionic toxicity and water imbalance (Yang and Guo, 2018), mild NaCl treatment can considerably enhance the polyphenol contents and antioxidant activity in many plant species, such as rice seedlings as well as mungbean, broccoli, buckwheat, and radish sprouts (Yuan et al., 2010; Lim et al., 2012; Guo et al., 2014; Koodkaew, 2019). Similarly, many metabolic studies have shown that anthocyanin and flavonoid metabolisms are significantly influenced by mild NaCl treatment, followed by enhanced production of catechin, chlorogenic acid, ferulic acid, *p*-coumaric acid, syringic acid, and vitexin (Lim et al., 2012; Hassini et al., 2017; Sarker and Oba, 2018; Banik and Bhattacharjee, 2020; Benincasa et al., 2021). Therefore, salinity stress can play a crucial role in regulating the quality of the nutritional values of plant ingredients.

With RNA-seq-based transcriptome analysis, the gene expression patterns and regulation network in response to salinity have been widely investigated in several plants, including cotton (Wang et al., 2012), rice (Chandran et al., 2019), buckwheat (Ma et al., 2019), soybean (Fan et al., 2013), and medicine plants (Ben Abdallah et al., 2016). The expression of the genes involved in the biosynthetic pathways of flavonoids and phenylpropanoids can be increased by NaCl treatment in *solanum nigrum*, including phenylalanine ammonialyase, chalcone synthase, and flavonol synthase (Ben Abdallah et al., 2016). Furthermore, various transcription factor (TF) families play a crucial role in salinity stress responses in plants, such as NACs, MYBs, WRKYs, and bHLHs (Eulgem and Somssich, 2007). Overexpression of TaNAC29, AtMYB20, FtMYB10, and OsMYB3R-2 enhanced salt tolerance in transgenic Arabidopsis (Dai et al., 2007, 3; Cui et al., 2013, 3; Xu et al., 2015; Gao et al., 2016). TaMYB56-B, MdSIMYB1, and OsNAC6 were associated with salinity stress responses in wheat, apple, and rice, respectively (Nakashima et al., 2007; Shen et al., 2017).

During the mungbean germination stage, polyphenol accumulation is affected by growth conditions and stresses, such as temperature, salinity, and drought (Kim et al., 2009b). When mungbean sprouts are exposed to NaCl, the total phenolic and flavonoid contents increase, resulting in enhanced antioxidant activities (Koodkaew, 2019). However, only a few transcriptome studies have explored resistance to fungi and drought stress in mungbean seeds/leaves (Kumar et al., 2020), and little is known about the genetic regulation of secondary metabolites for abiotic stresses in mungbean sprouts.

This study's aim is to investigate the effects of NaCl treatment on morphological, metabolomic, and transcriptomic changes in mungbean sprouts using two cultivars, dahyeon and sunhwa. Dahyeon has been reported to accumulate relatively higher polyphenols in its sprouts compared to other cultivars and sunhwa is a mungbean reference cultivar (Kim et al., 2009a; Habibzadeh and Yagoob, 2014; Kang et al., 2014; Yoseph Ganta et al., 2021). We qualified and quantified the polyphenolic compounds in mungbean sprouts after NaCl treatment and identified the key enzymes involved in the biosynthetic pathways of these compounds. Our findings provide further insights into the molecular mechanisms associated with salinity stress and can help improve the nutritional quality and morphological characteristics of sprout vegetables through genetic engineering in legume crops.

Materials and methods

Plant materials and NaCl treatment

Two mungbean cultivars, dahyeon and sunhwa (VC1973A), were used in our experiments. Seeds were harvested at the Gangneung-Wonju National University Experimental Farm in Gangneung, South Korea (37.77°N, 128.86°E). The seeds were soaked in distilled water for 17 h at 37°C using an incubator (Jeiotech, ISS-4075R, Korea) for germination. The germinated seeds were transplanted into a plant growth chamber (Sundotcom, ST001A, Korea) and cultivated for three days at 28 ± 2°C in the dark. Water spraying automatically occurred for 4 min every 2 h (Kim et al., 2021). During cultivation, mungbean sprouts were soaked in different concentrations (0, 50, 100, and 200 mM) of NaCl solution for 30 min every 12 h (KisanBio, MB-

S4636, Korea). Mungbean sprouts were harvested on the third day after germination and stored at −70°C for RNA extraction (Lim et al., 2022). Fifty seeds were cultured for each experimental group, 30 sprouts were randomly picked for further analysis, and five sprouts were used for each biological replication of RNA extraction. The lengths of the root and hypocotyl and the thickness of the hypocotyl were measured using ImageJ (Schneider et al., 2012).

Extraction procedure

Mungbean sprouts were fully dried at 70°C for 24 h using an incubator (Gan et al., 2017). The sample was ground into a fine powder using mortar and pestle and 0.05 g powder of each sample was extracted using 70% ethanol (w/v, 1:10) for 24 h in the dark. Mixtures were vortexed and sonicated for 15 min and centrifuged at 13,000 rpm for 10 min. The supernatant was filtered through syringe filters (SMART I LAB, SPF0213-1, USA, 0.22 um) and diluted with 70% ethanol for further analysis.

Determination of total phenol content

The contents of total phenol were measured using Folin-Ciocalteu colorimetric method with slight modification (Singleton and Rossi, 1965). Gallic acid (0, 50, 100, and 200 mg/L) was used as a standard for quantification. 50 μL Folin–Ciocalteu (Sigma-Aldrich®, 47641, USA) was added to each mungbean sprout extraction (100 μL, 10,000 ppm) in an EP tube. After 5 min, 300 μL of 20% Na_2CO_3 (FUJIFILM, 199-01605, Japan) solution was added. After 15 min in dark condition, distilled water was added up to 1 mL. After 2 min of centrifuge, 200 ul of the supernatants were transferred to a 96-well plate, and absorbance was measured at 740 nm wavelength using a spectrophotometer (Thermo Scientific MIB, Multiskan FC, Korea, 738 nm), with three replications. The results were expressed as milligrams gallic acid equivalent (GAE) per 1 g of dry weight (GAE mg/g).

Determination of total flavonoid content

The contents of total flavonoids were measured using the aluminum nitrate colorimetric method with slight modifications (Zhishen et al., 1999). Quercetin (Sigma-Aldrich®, Q4951, USA) (0, 50, 100, and 200 mg/L) was used as standard. The mungbean sprouts extract (500 μL, 50,000 ppm) was mixed with a 10% aluminum nitrate (100 μL) solution and 1 M potassium acetate solution (100 μL). After 40 min in dark conditions and 2 min of centrifuge, 150 μL of the supernatants were transferred to a 96-well plate, and absorbance was measured at 429 nm wavelength using a spectrophotometer, with three replications. The results were expressed as milligrams quercetin equivalent (QAE) per 1 g of dry weight (QAE mg/g).

Determination of antioxidant activity through DPPH and ABTS assays

DPPH radical scavenging activity was measured using OxiTec™ DPPH Antioxidant Assay Kit (BIOMAX, BO-DPH-500, Korea) according to the manufacturer's protocol. Trolox (0, 40, 60, 80, and 100 mg/L) was used as the standard. Each extraction (20 μL, 50,000 ppm) was mixed with assay buffer (80 μL) and DPPH radical (100 μL) in 96-well plates, with three replications. The samples were preserved in the dark for 30 min, and absorbance was measured at 520 nm using a spectrophotometer. The result was expressed as a percentage of scavenging achieved by the DPPH.

ABTS radical scavenging activity was measured as described by Yen (Yen and Chen, 1995). Ascorbic acid (FUJIFILM, 012–04802, Japan) (0, 1, 5, 10, 25, 50, and 100 mg/L) was used as the standard. ABTS (Roche, Cat. No. 10102946001, Swiss) 7.4 mM solution and 2.6 mM potassium persulfate (YAKURI, 28718, Korea) were diluted to a 1:1 ratio (v/v). ABTS solution (180 μL) was added to each extraction (20 μL, 1,000 ppm) in 96-well plates, with three replications. The samples were kept in dark for 10 min, their absorbance was measured at 720 nm using a spectrophotometer. The result was expressed as a percentage of scavenging achieved by the ABTS.

Ultra-high-performance liquid chromatography analysis

Ultra-high-performance liquid chromatography (UHPLC) analysis was conducted using the Shimadzu UHPLC system (Nexera series equipped with MPM-40, SCL-40, SPD-M40, LC-40, SIL-40, and CTO-40 unit from Shimadzu, Kyoto, Japan) with a UV-vis detector. Separation was conducted using a C18 column (Shimadzu, Kyoto, Japan; 250 × 4.6 mm, 5 μm). Ultrapure water (Thermo Fisher, W5-4, Korea) with 0.1% ortho-phthalaldehyde buffer and acetonitrile (Thermo Fisher, 022927 M6, Korea) were used as solvents A and B, respectively. The flow rate was 0.2 mL/min and the column temperature was set at 40°C. The ratio of solvent B proceeded with 5%, increased to 30% for 0–8 min, increased to 36% for 8–20 min, and decreased to 5% for 20–25 min. 14 compounds were used as standard compounds: caffeic acid (CFN99190), catechin (CFN99646), chlorogenic acid (CFN99116), gallic acid (CFN99624), isovitexin (CFN98620), kaempferol (CFN98838), myricetin (CFN98877), neochlorogenic acid (CFN97472), *p*-coumaric acid (CFN97218), quercetin (CFN99272), resveratrol (CFN98791), syringic acid (CFN98884), t-ferulic acid (CFN99158), and vitexin (CFN98601). The standard

compounds were purchased from ChemFaces (Wuhan, China). The results were obtained based on standard calibration curves (10-100 ppm) with a tri-repeat and expressed as milligrams per 100 g of dry weight (mg/100g).

cDNA library construction and transcriptome sequencing

Total RNA was isolated from whole mungbean sprouts using a Ribospin^{TM} Plant RNA extraction Kit (GeneAll, Songpa-gu, South Korea) according to the manufacturer's protocol. The cDNA libraries for RNA-seq were constructed using a TruSeq Stranded mRNA LT Sample Prep Kit (Illumina, Inc., San Diego, CA, USA). Six libraries were constructed from the treatment and control groups, with three biological replications in each group. The size and quality of the libraries used for sequencing were checked using a 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA, USA). Sequencing runs were conducted in paired-end mode using a Truseq SBS Kit on the Illumina NovaSeq 6000 platform. RNA sequencing data were deposited in the NCBI SRA database (PRJNA853140).

Identification and analysis of differentially expressed genes

Raw reads were filtered out (q-value > 20) using trimmomatic version 0.36 (http://www.usadellab.org/) with default parameters (Bolger et al., 2014). The reference genome and gene annotation data were downloaded from the Seoul National University Crop Genomics Lab (http://plantgenomics.snu.ac.kr) (Ha et al., 2021). The filtered reads were mapped to the mungbean reference genome using Tophat2 (Kim et al., 2013, 2). Counts per million mapped read values were calculated for 30,999 genes using FeatureCounts (Liao et al., 2014). DEGs were counted using Edge-R (Robinson et al., 2010), which were defined as genes with an absolute $\log_2$ fold change (FC) value greater or equal to 1 between the two groups (control over treatment).

Functional annotation of DEGs

Gene ontology (GO) enrichment analysis was performed using BINGO software (http://www.psb.ugent.be/cbd/papers/BiNGO/Home.html) (Maere et al., 2005). The Kyoto Encyclopedia of Genes and Genomes (KEGG) ontology (KO) pathway enrichment analysis was conducted using DAVID (https://david.ncifcrf.gov/) (Sherman et al., 2022). The TFs were identified using PlantTFDB (http://planttfdb.gao-lab.org/) (Tian et al., 2020).

Validation by 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 protocol. Thirty cycles of PCR amplification were conducted using gene-specific primers. Cytochrome P450 family 20 (CYP20) was used as an internal control gene. The primer sequences are listed in the supplementary information (Table S3).

Statistical analysis

Statistical analysis was performed in R (R Core Team, 2017). The significant differences were calculated by a one-way analysis of variance followed by Duncan's multiple test; p-value <0.05 was considered statistically significant.

Results

Development of mungbean sprouts under NaCl treatments

We first assessed the effect of NaCl treatment on the development of mungbean sprouts using a gradient of different NaCl concentrations (0, 50, 100, and 200 mM) (Figure 1A). The length and thickness of hypocotyl and the length of the roots were measured in dahyeon and sunhwa (Figure 1B). The length of the roots and hypocotyl significantly decreased in both cultivars when the NaCl concentration increased (Figures 1B, C). While hypocotyl thickness showed no significance in dahyeon, it increased significantly in sunhwa under NaCl treatments (Figure 1D).

Dynamic changes in total phenol and flavonoid contents, and antioxidant activities under NaCl treatment

We also evaluated the total phenol and flavonoid contents and antioxidant activity under different NaCl concentrations (Figure 2). In general, for both cultivars, the phenol contents decreased as the NaCl concentration increased (Figure 2A). However, for total flavonoid content, the two cultivars had opposite responses under low concentrations of NaCl treatment (50 mM, Na50). The flavonoid contents of dahyeon showed a significant decrease when NaCl concentrations increased, while those of sunhwa peaked under Na50 and then

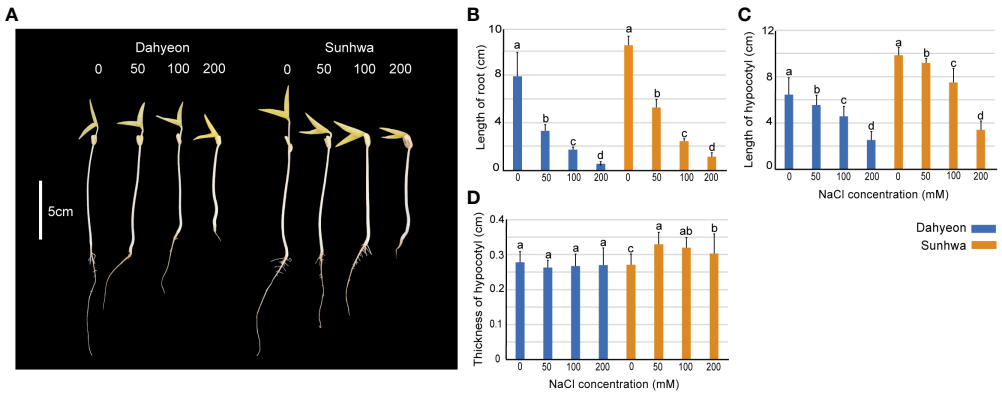


FIGURE 1
Morphological changes in mungbean sprouts of dahyeon and sunhwa cultivars under different concentrations of NaCl treatment (0, 50, 100, and 200 mM). **(A)** Picture of sprouts, **(B)** length of root, **(C)** length of hypocotyl, and **(D)** thickness of hypocotyl. The color bar indicates the mungbean cultivars, blue represents dahyeon and orange represents sunhwa. The results were expressed with a standard error bar, and different lowercase letters indicate statistical significance ($P < 0.05$).

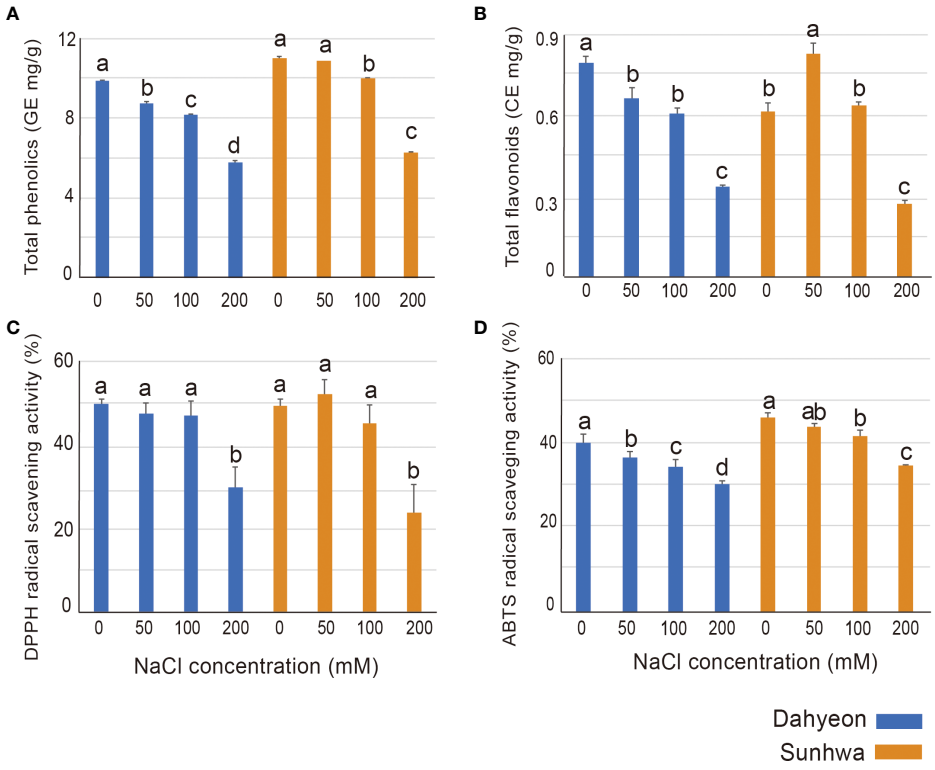


FIGURE 2
Comparison of **(A)** total phenolic contents, **(B)** total flavonoid contents, **(C)** antioxidant scavenging activity of DPPH, and **(D)** antioxidant scavenging activity of ABTS in the two mungbean cultivars cultivated on different concentrations of NaCl treatment (0, 50, 100, and 200 mM). The color bar indicates the mungbean cultivars, blue represents dahyeon and orange represents sunhwa. The results were expressed with a standard error bar, and different lowercase letters indicate statistical significance ($P < 0.05$).

decreased under higher concentrations of NaCl (100 and 200 mM) (Figure 2B).

DPPH and ABTS assays were conducted to evaluate the antioxidant activity of the mungbean sprouts (Figures 2C, D). In the DPPH assay, no difference in the antioxidant activities was observed in both cultivars treated with 0, 50, and 100 mM NaCl, and the antioxidant activity significantly declined under 200 mM NaCl treatment (Figure 2C). The antioxidant activity detected through the ABTS assay was gradually reduced in both cultivars as NaCl concentration increased (Figure 2D).

Profile of polyphenolic compounds using UHPLC analysis

Metabolic profiling was conducted on sunhwa sprouts treated with Na50, where flavonoid contents significantly increased (Table 1). To better understand the effects of salinity on polyphenolic compounds, we quantified and qualified 14 compounds in mungbean sprouts, including six flavonoids (catechin, isovitexin, kaempferol, myricetin, quercetin, and vitexin) and eight phenolic acids (caffeic acid, chlorogenic acid gallic acid, *p*-coumaric acid, neochlorogenic acid, resveratrol, syringic acid, and t-ferulic acid) using UHPLC.

Among the 14 compounds, 13 compounds excluding kaempferol were detected in both the control and NaCl treatment samples (Table 1). Chlorogenic acid was the most abundant in both the control and NaCl treatments, with values of 1144.1 mg/100g and 1444.2 mg/100g, respectively (Table 1). Myricetin was the second most abundant compound, with values of 328.6 mg/100g (control) and 233.9 mg/100g (Na50), followed by gallic acid (143.3 and 146.8 mg/g), catechin (114.2 and 160.4 mg/100g), quercetin (132.8 and 113.6 mg/100g), and vitexin (75.5 and 123.8 mg/100g), respectively (Table 1).

The contents of seven compounds (catechin, chlorogenic acid, glycerin, isovitexin, *p*-coumaric acid, syringic acid, t-ferulic acid, and vitexin) were significantly increased after Na50 treatment (Table 1). In contrast, the contents of four compounds (myricetin, neochlorogenic acid, quercetin, and resveratrol) were significantly reduced under NaCl treatment (Table 1). In the other two compounds (caffeic acid and gallic acid), no significant differences were observed in their contents between the control and Na50 (Table 1).

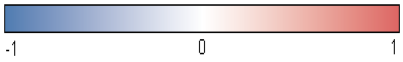
Transcriptome analysis

To investigate the effects of salinity stress on transcriptional changes, we conducted RNA sequencing on sunhwa sprouts treated with Na50. In total, 96.8 Gb of filtered reads (16.1 Gb per library, on average) were obtained from the control and Na50 samples, with three replications (Table S1). On average, 98.08% of the reads were mapped properly against the mungbean reference genome (Ha et al., 2021) (Table S1). DEGs were

TABLE 1 The effects of 50mM of NaCl treatment on the polyphenol composition of mungbean sprouts of sunhwa cultivar.

Chemical group	Compound	Content (mg/100g)		
		Control	NaCl treatment(50 mM)	Log_2FC
Flavonoid	Catechin	114.2±6.8^{b}	160.4±1.7^{a}	
	Isovitexin	16.6±0.1^{b}	25.8±0.4^{a}	
	Kaempferol	ND	ND	
	Myricetin	328.6±2.1^{a}	233.9±0.9^{b}	
	Quercetin	132.8±1.1^{a}	113.6±0.3^{b}	
	Vitexin	75.5±0.4^{b}	123.8±1^{a}	
Phenolic acid	Caffeic acid	57.6±1.6 ^{ns}	57.4±0.2 ^{ns}	
	Chlorogenic acid	1144.1±5.7^{b}	1444.2±7.1^{a}	
	Gallic acid	143.3±2.2 ^{ns}	146.8±0.7 ^{ns}	
	Neochlorogenic acid	35.7±0.3^{a}	30.4±0.9^{b}	
	p-Coumaric acid	16.4±0.7^{b}	26.2±0.1^{a}	
	Resveratrol	15.6±0.1^{a}	14.6±0.2^{b}	
	Syringic acid	4.1±0.5^{b}	6.5±1^{a}	
	t-Ferulic acid	13.4±0.2^{b}	18.1±0.2^{a}	

The color box indicates log2-fold change values of each metabolite (Na 50 over control), blue and red represent lower and higher values, respectively. Different lowercase letters indicate statistical significance ($P < 0.05$).
ND and NS mean non-detected and non-significance, respectively.



identified between the NaCl treatment and control groups. A total of 728 DEGs were identified, and 151 and 577 genes were downregulated and upregulated after NaCl treatment, respectively (Figures 3A, B).

Functional annotation of DEGs through enrichment analysis

The KEGG pathway and GO enrichment analyses were conducted to understand DEG's molecular functions. In KEGG pathway analysis, 133 DEGs were assigned to 11 pathways and the top three pathways were "phenylpropanoid biosynthesis" (15), "starch and sucrose metabolisms" (11), and "biosynthesis of various plant secondary metabolites" (8) (Figure 4A). In GO analysis, 464 DEGs were clustered in 104 categories, and the major clusters were "response to stimulus" (137), followed by "response to stress" (100) and "response to chemical stimulus" (83) (Figure 4B). Within the categories involved in secondary metabolites, "phenylpropanoid metabolic process" (18) and "flavonoid metabolic process" (6) were identified as enriched categories. Based on the enrichment analyses, 46 DEGs were identified as genes related to the biosynthesis of phenylpropanoid metabolites. The 46 DEGs were mapped into the phenylpropanoid pathway in KEGG, and 11 DEGs were selected as key genes encoding the enzymes directly involved in the biosynthesis of the secondary metabolites that increased significantly after NaCl treatment (Figure 5; Table S2).

The 11 DEGs encode seven enzymes: catechol O-methyltransferase (COMT: Vradi03g00001673.1, Vradi02g00004009.1), caffeoyl-CoA O-methyltransferase (CCOMT: Vradi09g00001754.1), chalcone synthase (CHS: Vradi09g00003301.1, Vradi09g00003302.1, Vradi09g00003303.1), chalcone isomerase (CHI: Vradi03g00001100.1), shikimate O-hydroxycinnamoyl transferase (HCT: Vradi09g00002363.1), coniferyl-aldehyde dehydrogenase (REF1: Vradi09g00002363.1, Vradi02g00003639.1), and dihydroflavonol 4-reductase (DFR: Vradi07g00001339.1) (Table S3 and Figure 5). In addition, the 11 DEGs were significantly upregulated under salinity conditions, and their Log_2 FC values were COMT (1.75, 1.31), CCOMT (1.38), CHI (1.18), CHS (1.14, 1.13, 1.1), HCT (2.35), REF1 (1.85, 1), and DFR (1.46) (Figure 5). For validation of RNA-seq data, we conducted qRT-PCR and found that the result of qRT-PCR well agree with that of RNA-seq, indicating the reliability of the RNA-seq data (Figure 6).

Discussion

Mungbean sprouts are a good source of natural antioxidants, including catechin, chlorogenic acid, vitexin, isovitexin, and *p*-coumaric acid (Kim et al., 2009b; Guo et al., 2012; Gan et al., 2016; Singh et al., 2017). These secondary metabolites provide various health benefits for the human body, such as anticancer, antioxidant, antimicrobial, anti-inflammatory, and antitumor effects (Gorzynik-Debicka et al., 2018; Stagos, 2020; Teodor et al., 2020). In plants, abiotic stresses, such as salinity

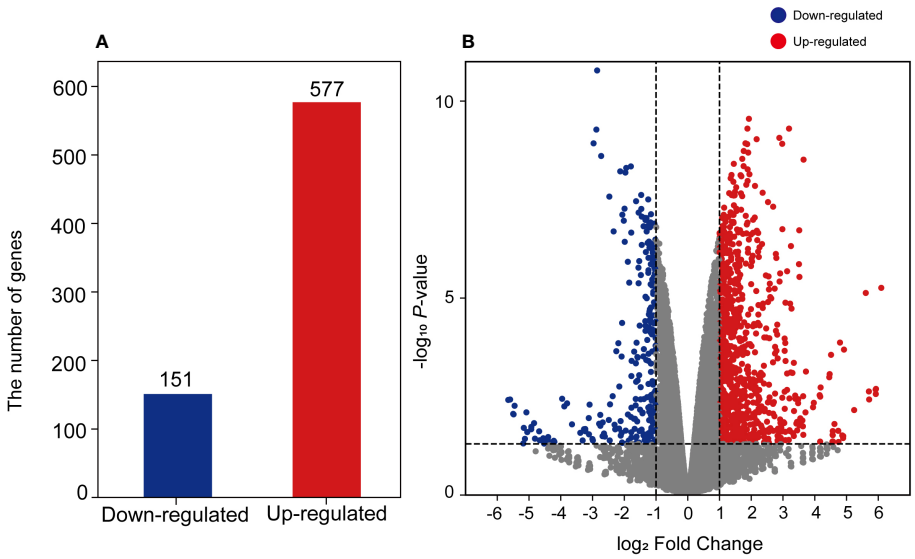


FIGURE 3
Visualization of identified DEGs using **(A)** bar plot and **(B)** volcano plot. The y-axis of the volcano plot indicates log2-fold change values of each DEG, and the x-axis indicates −log10 p-value. The color bar indicates gene regulation, blue represents downregulated genes and red represents upregulated genes.

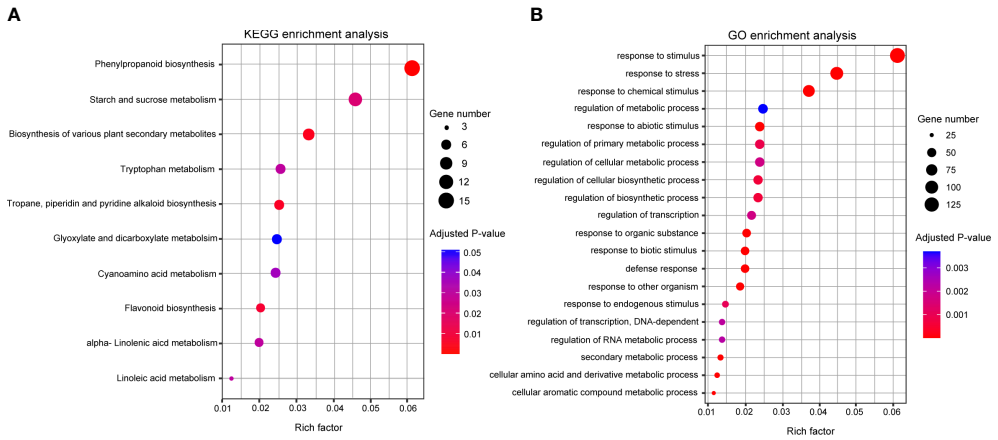


FIGURE 4
(A) KEGG and **(B)** GO pathway enrichment analyses of DEGs. The y-axis indicates the pathway or term name, and the x-axis indicates the enriched factor in each pathway and term. The circle size indicates the number of genes. The color bar indicates the adjusted p-value, blue and red represent higher and lower values, respectively. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

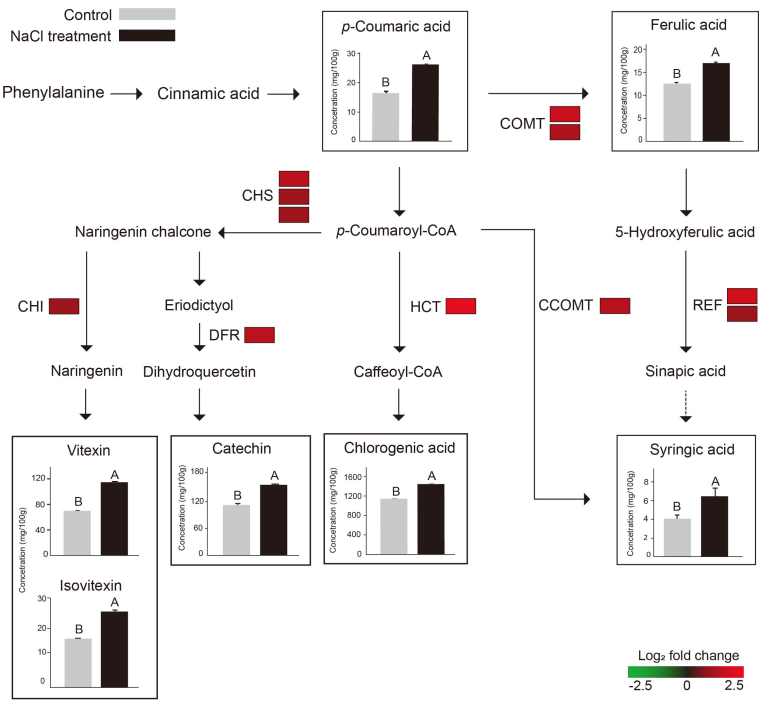


FIGURE 5
Schematic illustration of phenylpropanoid pathway with the expression levels of 11 DEGs (COMT, CCOMT, CHS, CHI, HCT, REF1, and DFR) and the contents of 7 metabolites that increased after NaCl treatment (catechin, chlorogenic acid, glycerin, isovitexin, *p*-coumaric acid, syringic acid, t-ferulic acid, and vitexin). The color scale indicates the log2-fold change (FC) value of each DEG, green and red represent lower and higher expression levels, respectively. The color blocks indicate FC of expression levels of paralogs of each DEG. The bar plots in the square show the comparison of metabolite contents between the control (gray) and NaCl treatment (black). The different uppercase letters indicate statistical significance ($P < 0.05$). The dotted line indicates an enzymic reaction that has not been fully identified in plants. COMT, catechol O-methyltransferase; CCOMT, caffeoyl-CoA O-methyltransferase; CHS, chalcone synthase; CHI, chalcone isomerase; HCT, shikimate O-hydroxycinnamoyl transferase; REF1, coniferyl-aldehyde dehydrogenase; DFR, Dihydroflavonol 4-reductase;.

conditions, significantly influence secondary metabolite accumulation (Ksouri et al., 2008; Yang et al., 2018; Sharma et al., 2019; Chrysargyris et al., 2020). Many studies have demonstrated that mild NaCl treatment can improve secondary metabolite contents in various vegetables (Lim et al., 2012; Hassini et al., 2017; Sarker and Oba, 2018; Banik and Bhattacharjee, 2020; Benincasa et al., 2021). In the present study, we investigated the interactive effects of salinity stress on transcriptomic and metabolomic changes in mungbean sprouts.

NaCl treatment significantly suppressed mungbean sprout development (Figure 1). The root and hypocotyl length was reduced in both cultivars as NaCl concentration increased from 50 to 200 mM (Figures 2A, B). These results agree with previous findings that osmotic imbalance caused by salinity results in a significant reduction in root and hypocotyl length in maize (Khodarahmpour et al., 2012), wheat (Akbari et al., 2007), cowpea (Dantas et al., 2005), weed species (Hakim et al., 2011), tomato (Abdel-Farid et al., 2020), and mungbean (Saha et al., 2010). This inhibition of root growth by salinity stress is obviously a drawback in plants (Yin et al., 2022). However, in bean sprout production, long roots are considered a totally undesirable feature to reduce sprout texture, whereas a certain length and thickness of hypocotyl are required (Price, 1988). A previous study suggests that the most suitable features of mungbean sprouts are short roots and hypocotyl longer and thicker than 5 cm in length and 2 mm in diameter (Buescher and Chang, 1982). In our observation, sunhwa treated with Na50 showed an increase in hypocotyl thickness from 2.7 to 3.3 mm, maintaining a hypocotyl length over 8 cm (Figures 1B, C), indicating that mild NaCl treatment can promote the morphological quality of mungbean sprouts in consumers' favor.

The total flavonoid contents increased in sunhwa at Na50 (Figure 2B), which corroborates previous findings that NaCl treatment can enhance the total flavonoid contents in wheat (Kiani et al., 2021), tomato (Abdel-Farid et al., 2020), and rapeseed sprout (Falcinelli et al., 2017). However, the total flavonoid contents gradually decreased in dahyeon as NaCl concentration increased (Figure 3B). This opposite response might be caused by sensitivity/insensitivity of the two cultivars to salt stress. Sunhwa (VC1973A) is a reference mungbean cultivar developed by the Asian Vegetable Research and Development Center ensuring high yield and insensitivity against diverse stresses (Habibzadeh and Yagoob, 2014; Kang et al., 2014; Yoseph Ganta et al., 2021). Secondary metabolites can be differentially regulated in salt-tolerant or sensitive cultivars under salinity condition (Borghesi et al., 2011; Linić et al., 2019). In the Brassicaceae family, the contents of phenolic acid were higher in more salt-tolerant varieties, such as white cabbage and kale, than in salt-sensitive varieties (Linić et al., 2019). In tomatoes, the accumulation of total anthocyanin contents differed depending on genotypes under NaCl treatment (Borghesi et al., 2011). Our results, along with those from other studies, reveal that the genotype with high salinity tolerance might be a valuable breeding material for high contents of antioxidant compounds in mungbean sprouts.

In this study, the antioxidant activity of mungbean sprouts was measured using two colorimetric methods (Figures 2C, D). DPPH and ABTS are both widely used for measuring the antioxidant activity of plant extraction (Floegel et al., 2011). While ABTS assay is suitable to measure both hydrophilic and lipophilic antioxidants, DPPH assay is more suited to measure hydrophobic systems (Floegel et al., 2011). We found that the antioxidant activities obtained by ABTS and DPPH assays presented slightly different trends (Figures 2C, D). In the ABTS assay, the antioxidant activity gradually declined as NaCl concentration increased, whereas no significant differences were observed in the DPPH assay under mild concentrations of NaCl (Figures 2C, D). This difference between the two assays indicates that the antioxidants mainly affected by mild NaCl treatment (0-100 mM) might be hydrophilic phytochemicals in mungbean sprouts.

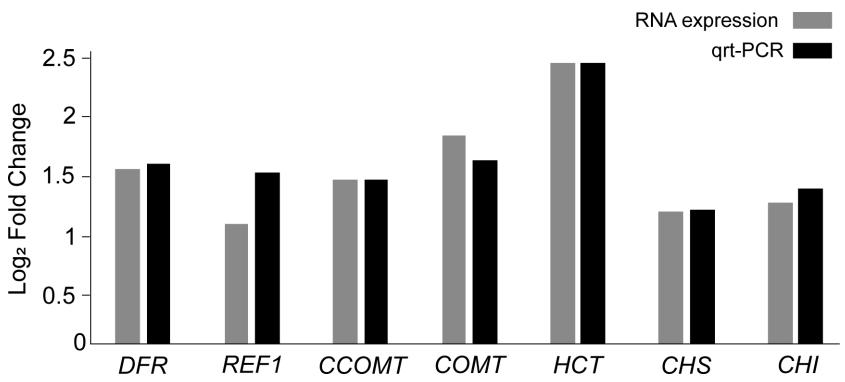


FIGURE 6
Validation of the RNA expression of DEGs by qRT-PCR analysis.

Although a few studies have reported the effects of NaCl treatment on the total phenol and flavonoid contents in mungbean sprouts, the changes in the contents of individual phytochemicals along with transcriptomic changes remain unexplored (Koodkaew, 2019; Mankar et al., 2021). We conducted UHPLC analysis using 14 phenylpropanoid compounds in different groups (flavonoids and phenolic acid), which have been reported to be responsible for antioxidant activity in mungbean seeds and sprouts (Kim et al., 2009b; Guo et al., 2012; Gan et al., 2016; Singh et al., 2017). We found that most flavonoid and phenolic acid compounds significantly increased under salinity conditions, including catechin, chlorogenic acid, isovitexin, vitexin, *p*-coumaric acid, syringic acid, and t-ferulic acid (Table 1). In buckwheat sprouts, the contents of rutin and vitexin were enhanced under 50 mM NaCl treatment (Lim et al., 2012). The contents of catechin were increased in rice treated with 250 mM NaCl, and chlorogenic, *p*-coumaric, syringic, and t-ferulic acids were also increased by NaCl treatments in the Amaranthus, *brassica* family, rapeseed, and rice (Hassini et al., 2017; Sarker and Oba, 2018; Banik and Bhattacharjee, 2020; Benincasa et al., 2021). In contrast, salinity treatment can decrease some polyphenols in broccoli and romaine lettuce (Kim et al., 2008; López-Berenguer et al., 2009). In rice, while gallic acid, syringic acid, and catechin increased with 250 mM NaCl treatment, myricetin and quercetin decreased (Banik and Bhattacharjee, 2020). We also observed decreased contents of myricetin and quercetin under salinity conditions in the mungbean sprout (Table 1). Our findings suggest that the biosynthetic pathways of flavonoids and phenolic acids respond differently to salinity stress, which might be regulated by the key enzymes involved in each pathway.

According to the transcriptomic results, some genes (151 downregulated and 528 upregulated) were differentially regulated by NaCl treatment (Figure 3). We found that the expression of the genes related to the biosynthesis of the secondary metabolites was significantly increased by NaCl treatment (Table S2). *DFR*, *REF1*, *CCOMT*, *COMT*, *HCT*, *CHS*, and *CHI* genes were significantly highly expressed in mungbean sprouts under 50 mM NaCl treatment compared to the control, which corresponded to the changes in the secondary metabolite contents (Figure 5). In a previous study, the expression of *TaDFR* and *TaCHI* genes increased following 50 mM NaCl treatment in wheat sprouts (Cuong et al., 2020). In pak choi, *DFR*, *CHS*, and *CHI* genes were upregulated under 250 mM NaCl treatment (Yun et al., 2019). The expression of *CHS* was enhanced in *S. nigrum* as NaCl concentrations increased from 50 to 150 mM (Ben Abdallah et al., 2016). The *CCOMT* and *COMT* genes showed increased expression under salinity in various plants, including basil leaves, barley root, and tomato root (Sugimoto and Takeda, 2009; Manaa et al., 2011; Rastogi et al., 2019). In addition, many plant TF families were identified as upregulated DEGs, including ERFs, WRKYs, and MYBs (Figure 7). These TF families often stimulate the metabolic pathway under environmental stress, such as drought and salinity, resulting in more accumulated secondary metabolites in plants (Espley et al., 2007; Ilk et al., 2015; Qiu et al., 2016; Wei et al., 2017; Upadhyay et al., 2018; Li et al., 2019). Our TF DEGs were enriched to "response to stress," "response to osmotic stress," "response to salt stress," "response to abiotic stimulus," among other responses, and the expressions of these TF DEGs were increased up to $\log_2$FC 4.5 ($\log_2$FC 2.0 on average) (Table S4). In Arabidopsis, the overexpression of MYB111, MYB1D, and Mdmyb10 TFs significantly increased the expression of *COMT*, *CHS*, and *CHI* genes under salinity stress (Gao et al., 2011; Li et al., 2019). The silencing of OscWRKY1 resulted in reduced expression of *PAL*, *COMT*, and *4CL* transcripts in Arabidopsis (Joshi et al., 2022). In potatoes, *CHS*, *DFR*, and

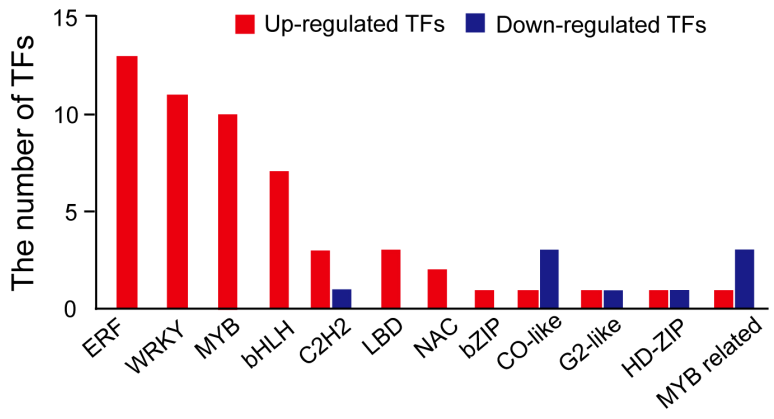


FIGURE 7
The number of transcription factors identified in differentially expressed genes. Red represents upregulated TF DEGs, and blue represents downregulated TF DEG.

ANS genes were upregulated by IbMYB1 overexpression under osmotic stress (Cheng et al., 2013). In this study, four paralogous genes encoding MYB14 were detected as up-regulated DEGs, that had been reported to be positively correlated with key flavonoid synthetic genes, including *ANS*, *ANR*, *DFR*, *CHI*, and *CHS* (Liu et al., 2018). *WRKY72* that had been reported to be associated with the expression of *ANS* and *DFR* was also identified as upregulated TF DEG (Hu et al., 2020). Our transcriptomic results show that the biosynthetic pathways of phenylpropanoids, especially flavonoids, are regulated at the transcription level under salinity stress.

Conclusion

We found that salinity stress can enhance the accumulation of phenylpropanoids and the expression of their responsive genes in mungbean sprouts. Our results indicate that mild NaCl treatment can be a practical approach to increasing the nutritional value of mungbean sprouts. Our comprehensive metabolomic and transcriptomic data provide insights into the complex regulatory mechanisms for salinity tolerance, as well as valuable information for improving the nutritional quality in legume crops.

Data availability statement

The data presented in the study are deposited in the NCBI SRA repository, accession number: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA853140.

Author contributions

IL and MK designed the experiments and wrote the manuscript. IL and MK collected the material data and conducted experiments. BK performed the qualifying and quantifying of the metabolites. IL analyzed the transcriptomic results and contributed to the illustration. JH supervised and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a grant from the National Research Foundation of Korea (NRF), funded by the Korean government (MSIT; no. 2021R1C1C1004233).

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.2022.1030677/full#supplementary-material

References

Abdel-Farid, I. B., Marghany, M. R., Rowezek, M. M., and Sheded, M. G. (2020). Effect of salinity stress on growth and metabolomic profiling of cucumis sativus and solanum lycopersicum. *Plants* 9, 1626. doi: 10.3390/plants9111626

Akbari, G., Sanavy, S. A. M. M., and Yousefzadeh, S. (2007). Effect of auxin and salt stress (NaCl) on seed germination of wheat cultivars (*Triticum aestivum* l.). *Pak J. Biol. Sci.* 10, 2557–2561. doi: 10.3923/pjbs.2007.2557.2561

Banik, N., and Bhattacharjee, S. (2020). Complementation of ROS scavenging secondary metabolites with enzymatic antioxidant defense system augments redox-regulation property under salinity stress in rice. *Physiol. Mol. Biol. Plants* 26, 1623–1633. doi: 10.1007/s12298-020-00844-9

Ben Abdallah, S., Aung, B., Amyot, L., Lalin, I., Lachâal, M., Karray-Bouraoui, N., et al. (2016). Salt stress (NaCl) affects plant growth and branch pathways of carotenoid and flavonoid biosyntheses in solanum nigrum. *Acta Physiol. Plant* 38, 72. doi: 10.1007/s11738-016-2096-8

Benincasa, P., Bravi, E., Marconi, O., Lutts, S., Tosti, G., and Falcinelli, B. (2021). Transgenerational effects of salt stress imposed to rapeseed (*Brassica napus* var. oleifera del.) plants involve greater phenolic content and antioxidant activity in the edible sprouts obtained from offspring seeds. *Plants* 10, 932. doi: 10.3390/plants10050932

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Borghesi, E., González-Miret, M. L., Escudero-Gilete, M. L., Malorgio, F., Heredia, F. J., and Meléndez-Martínez, A. J. (2011). Effects of salinity stress on carotenoids, anthocyanins, and color of diverse tomato genotypes. *J. Agric. Food Chem.* 59, 11676–11682. doi: 10.1021/jf2021623

Buescher, R., and Chang, J. (1982). Production of mung bean sprouts. *Arkansas farm Res. Arkansas Agric. Experiment Station* 31, 13.

Chandran, A. K. N., Kim, J.-W., Yoo, Y.-H., Park, H. L., Kim, Y.-J., Cho, M.-H., et al. (2019). Transcriptome analysis of rice-seedling roots under soil–salt stress using RNA-seq method. *Plant Biotechnol. Rep.* 13, 567–578. doi: 10.1007/s11816-019-00550-3

Cheng, Y.-J., Kim, M.-D., Deng, X.-P., Kwak, S.-S., and Chen, W. (2013). Enhanced salt stress tolerance in transgenic potato plants expressing IbMYB1, a sweet potato transcription factor. *J. Microbiol. Biotechnol.* 23, 1737–1746. doi: 10.4014/jmb.1307.07024

Chrysargyris, A., Mikallou, M., Petropoulos, S., and Tzortzakis, N. (2020). Profiling of essential oils components and polyphenols for their antioxidant activity of medicinal and aromatic plants grown in different environmental conditions. *Agronomy* 10, 727. doi: 10.3390/agronomy10050727

Cui, M. H., Yoo, K. S., Hyoung, S., Nguyen, H. T. K., Kim, Y. Y., Kim, H. J., et al. (2013). An arabidopsis R2R3-MYB transcription factor, AtMYB20, negatively regulates type 2C serine/threonine protein phosphatases to enhance salt tolerance. *FEBS Lett.* 587, 1773–1778. doi: 10.1016/j.febslet.2013.04.028

Cuong, D. M., Kwon, S.-J., Nguyen, B. V., Chun, S. W., Kim, J. K., and Park, S. U. (2020). Effect of salinity stress on phenylpropanoid genes expression and related gene expression in wheat sprout. *Agronomy* 10, 390. doi: 10.3390/agronomy10030390

Dai, X., Xu, Y., Ma, Q., Xu, W., Wang, T., Xue, Y., et al. (2007). Overexpression of an R1R2R3 MYB gene, OsMYB3R-2, increases tolerance to freezing, drought, and salt stress in transgenic arabidopsis. *Plant Physiol.* 143, 1739–1751. doi: 10.1104/pp.106.094532

Dantas, B. F., Ribeiro, L., de, S., and Aragão, C. A. (2005). Physiological response of cowpea seeds to salinity stress. *Rev. Bras. sementes* 27, 144–148. doi: 10.1590/S0101-31222005000100018

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

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. *Plant J.* 49, 414–427. doi: 10.1111/j.1365-313X.2006.02964.x

Eulgem, T., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr. Opin. Plant Biol.* 10, 366–371. doi: 10.1016/j.pbi.2007.04.020

Falcinelli, B., Sileoni, V., Marconi, O., Perretti, G., Quinet, M., Lutts, S., et al. (2017). Germination under moderate salinity increases phenolic content and antioxidant activity in rapeseed (Brassica napus var oleifera del.) sprouts. *Molecules* 22, E1377. doi: 10.3390/molecules22081377

Fan, X.-D., Wang, J.-Q., Yang, N., Dong, Y.-Y., Liu, L., Wang, F.-W., et al. (2013). Gene expression profiling of soybean leaves and roots under salt, saline–alkali and drought stress by high-throughput illumina sequencing. *Gene* 512, 392–402. doi: 10.1016/j.gene.2012.09.100

Floegel, A., Kim, D.-O., Chung, S.-J., Koo, S. I., and Chun, O. K. (2011). Comparison of ABTS/DPPH assays to measure antioxidant capacity in popular antioxidant-rich US foods. *J. Food Composition Anal.* 24, 1043–1048. doi: 10.1016/j.jfca.2011.01.008

Gan, R.-Y., Lui, W.-Y., Chan, C.-L., and Corke, H. (2017). Hot air drying induces browning and enhances phenolic content and antioxidant capacity in mung bean (*Vigna radiata* l.) sprouts. *J. Food Process. Preservation* 41, e12846. doi: 10.1111/jfpp.12846

Gan, R.-Y., Wang, M.-F., Lui, W.-Y., Wu, K., and Corke, H. (2016). Dynamic changes in phytochemical composition and antioxidant capacity in green and black mung bean (*Vigna radiata*) sprouts. *Int. J. Food Sci. Technol.* 51, 2090–2098. doi: 10.1111/ijfs.13185

Gao, F., Yao, H., Zhao, H., Zhou, J., Luo, X., Huang, Y., et al. (2016). Tartary buckwheat FtMYB10 encodes an R2R3-MYB transcription factor that acts as a novel negative regulator of salt and drought response in transgenic arabidopsis. *Plant Physiol. Biochem.* 109, 387–396. doi: 10.1016/j.plaphy.2016.10.022

Gao, J.-J., Zhang, Z., Peng, R.-H., Xiong, A.-S., Xu, J., Zhu, B., et al. (2011). Forced expression of Mdmyb10, a myb transcription factor gene from apple, enhances tolerance to osmotic stress in transgenic arabidopsis. *Mol. Biol. Rep.* 38, 205–211. doi: 10.1007/s11033-010-0096-0

Gayacharan., Archak, S., Gupta, K., Gupta, V., Tyagi, V., and Singh, K. (2020). "Mungbean genetic resources and utilization," in *The mungbean genome*, 9–25. (New York:Springer, Cham). doi: 10.1007/978-3-030-20008-4_2

Gorzynik-Debicka, M., Przychodzen, P., Cappello, F., Kuban-Jankowska, A., Marino Gammazza, A., Knap, N., et al. (2018). Potential health benefits of olive oil and plant polyphenols. *Int. J. Mol. Sci.* 19, 686. doi: 10.3390/ijms19030686

Guo, X., Li, T., Tang, K., and Liu, R. H. (2012). Effect of germination on phytochemical profiles and antioxidant activity of mung bean sprouts (*Vigna radiata*). *J. Agric. Food Chem.* 60, 11050–11055. doi: 10.1021/jf304443u

Guo, L., Yang, R., Wang, Z., Guo, Q., and Gu, Z. (2014). Effect of NaCl stress on health-promoting compounds and antioxidant activity in the sprouts of three broccoli cultivars. *Int. J. Food Sci. Nutr.* 65, 476–481. doi: 10.3109/09637486.2013.860583

Habibzadeh, Y., and Yagoob, M. (2014). The effects of water deficit stress on protein yield of mung bean genotypes. *Peak J. Agri. Sci.* 2, 30–35.

Hakim, M. A., Juraimi, A. S., Hanafi, M. M., Selamat, A., Ismail, M. R., and Karim, S. M. R. (2011). Studies on seed germination and growth in weed species of rice field under salinity stress. *J. Environ. Biol.* 32, 529–536.

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

Hassini, I., Baenas, N., Moreno, D. A., Carvajal, M., Boughanmi, N., and Martinez Ballesta, M. D. C. (2017). Effects of seed priming, salinity and methyl jasmonate treatment on bioactive composition of brassica oleracea var. capitata (white and red varieties) sprouts. *J. Sci. Food Agric.* 97, 2291–2299. doi: 10.1002/jsfa.8037

Hu, J., Fang, H., Wang, J., Yue, X., Su, M., Mao, Z., et al. (2020). Ultraviolet b-induced MdWRKY72 expression promotes anthocyanin synthesis in apple. *Plant Sci.* 292, 110377. doi: 10.1016/j.plantsci.2019.110377

Ilk, N., Ding, J., Ihnatowicz, A., Koornneef, M., and Reymond, M. (2015). Natural variation for anthocyanin accumulation under high-light and low-temperature stress is attributable to the ENHANCER OF AG-4 2 (HUA2) locus in combination with PRODUCTION OF ANTHOCYANIN PIGMENT1 (PAP1) and PAP2. *New Phytol.* 206, 422–435. doi: 10.1111/nph.13177

Joshi, A., Jeena, G. S., Shikha,, Kumar, R. S., Pandey, A., and Shukla, R. K. (2022). Ocimum sanctum, OscWRKY1, regulates phenylpropanoid pathway genes and promotes resistance to pathogen infection in arabidopsis. *Plant Mol. Biol* 2 (3), 30–35. doi: 10.1007/s11103-022-01297-2

Kang, Y. J., Kim, S. K., Kim, M. Y., Lestari, P., Kim, K. H., Ha, B.-K., et al. (2014). Genome sequence of mungbean and insights into evolution within vigna species. *Nat. Commun.* 5, 1–9. doi: 10.1038/ncomms6443

Khodarahmpour, Z., Ifar, M., and Motamedi, M. (2012). Effects of NaCl salinity on maize (*Zea mays l.*) at germination and early seedling stage. *Afr. J. Biotechnol.* 11, 298–304. doi: 10.4314/ajb.v11i2

Kiani, R., Arzani, A., and Mirmohammady Maibody, S. (2021). Polyphenols, flavonoids, and antioxidant activity involved in salt tolerance in wheat, aegilops cylindrica and their amphidiploids. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.646221

Kim, D., Choi, J., Lee, Y., Son, D., Moon, J., Oh, Y., et al. (2009a). A new mungbean cultivar, "Dahyeon" with many pod and high yielding. *Korean J. Breed. Sci.* 41, 36–39.

Kim, H.-J., Fonseca, J. M., Choi, J.-H., Kubota, C., and Kwon, D. Y. (2008). Salt in irrigation water affects the nutritional and visual properties of romaine lettuce (*Lactuca sativa* l.). *J. Agric. Food Chem.* 56, 3772–3776. doi: 10.1021/jf0733719

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, 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

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36. doi: 10.1186/gb-2013-14-4-r36

Kim, D.-K., Son, D.-M., Chon, S.-U., Lee, K.-D., Kim, K.-H., and Rim, Y.-S. (2009b). Phenolic compounds content and DPPH, ADH, ALDH activities of mungbean sprout based on growth temperature. *Korean J. Of Crop Sci.* 54, 1–6.

Koodkaew, I. (2019). NaCl And glucose improve health-promoting properties in mung bean sprouts. *Scientia Hortic.* 247, 235–241. doi: 10.1016/j.scienta.2018.12.022

Ksouri, R., Megdiche, W., Falleh, H., Trabelsi, N., Boulaaba, M., Smaoui, A., et al. (2008). Influence of biological, environmental and technical factors on phenolic content and antioxidant activities of Tunisian halophytes. *Comptes Rendus Biologies* 331, 865–873. doi: 10.1016/j.crvi.2008.07.024

Kumar, S., Ayachit, G., and Sahoo, L. (2020). Screening of mungbean for drought tolerance and transcriptome profiling between drought-tolerant and susceptible genotype in response to drought stress. *Plant Physiol. Biochem.* 157, 229–238. doi: 10.1016/j.plaphy.2020.10.021

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

Li, B., Fan, R., Guo, S., Wang, P., Zhu, X., Fan, Y., et al. (2019). The arabidopsis MYB transcription factor, MYB111 modulates salt responses by regulating flavonoid biosynthesis. *Environ. Exp. Bot.* 166, 103807. doi: 10.1016/j.envexpbot.2019.103807

Lim, I., Kim, B. C., Park, Y., Park, N. I., and Ha, J. (2022). 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

Lim, J.-H., Park, K.-J., Kim, B.-K., Jeong, J.-W., and Kim, H.-J. (2012). Effect of salinity stress on phenolic compounds and carotenoids in buckwheat (*Fagopyrum esculentum* m.) sprout. *Food Chem.* 135, 1065–1070. doi: 10.1016/j.foodchem.2012.05.068

Linić, I., Šamec, D., Grúz, J., Vujčić Bok, V., Strnad, M., and Salopek-Sondi, B. (2019). Involvement of phenolic acids in short-term adaptation to salinity stress is species-specific among brassicaceae. *Plants (Basel)* 8, 155. doi: 10.3390/plants8060155

Liu, L., Li, Y., She, G., Zhang, X., Jordan, B., Chen, Q., et al. (2018). Metabolite profiling and transcriptomic analyses reveal an essential role of UVR8-mediated signal transduction pathway in regulating flavonoid biosynthesis in tea plants (*Camellia sinensis*) in response to shading. *BMC Plant Biol.* 18, 233. doi: 10.1186/s12870-018-1440-0

López-Berenguer, C., Martínez-Ballesta, M., del, C., Moreno, D. A., Carvajal, M., and Garcia-Viguera, C. (2009). Growing hardier crops for better health: Salinity tolerance and the nutritional value of broccoli. *J. Agric. Food Chem.* 57, 572–578. doi: 10.1021/jf802994p

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21, 3448–3449. doi: 10.1093/bioinformatics/bti551

Ma, W., Kim, J. K., Jia, C., Yin, F., Kim, H. J., Akram, W., et al. (2019). Comparative transcriptome and metabolic profiling analysis of buckwheat (*Fagopyrum tataricum* (L.) gaertn.) under salinity stress. *Metabolites* 9, 225. doi: 10.3390/metabo9100225

Manaa, A., Ben Ahmed, H., Valot, B., Bouchet, J.-P., Aschi-Smiti, S., Causse, M., et al. (2011). Salt and genotype impact on plant physiology and root proteome variations in tomato. *J. Exp. Bot.* 62, 2797–2813. doi: 10.1093/jxb/erq460

Mankar, G. D., Wayase, U. R., Shelke, D. B., Nikam, T. D., and Barmukh, R. B. (2021). Morphological, physiological, and biochemical responses to NaCl-induced salt stress in mungbean (*Vigna radiata* l.) varieties. *Notulae Scientia Biologicae* 13, 10936–10936. doi: 10.15835/nsb13210936

Nair, R. M., Yang, R.-Y., Easdown, W. J., Thavarajah, D., Thavarajah, P., Hughes, J.D'A, et al. (2013). Biofortification of mungbean (*Vigna radiata*) as a whole food to enhance human health. *J. Sci. Food Agric.* 93, 1805–1813. doi: 10.1002/jsfa.6110

Nakashima, K., Tran, L.-S. P., Van Nguyen, D., Fujita, M., Maruyama, K., Todaka, D., et al. (2007). Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. *Plant J.* 51, 617–630. doi: 10.1111/j.1365-313X.2007.03168.x

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

Price, T. V. (1988). Seed sprout production for human consumption — a review. *Can. Inst Food Sci. Technol. J.* 21, 57–65. doi: 10.1016/S0315-5463(88)70718-X

Pudenz, M., Roth, K., and Gerhauser, C. (2014). Impact of soy isoflavones on the epigenome in cancer prevention. *Nutrients* 6, 4218–4272. doi: 10.3390/nu6104218

Qiu, Z., Wang, X., Gao, J., Guo, Y., Huang, Z., and Du, Y. (2016). The tomato hoffman's anthocyaninless gene encodes a bHLH transcription factor involved in anthocyanin biosynthesis that is developmentally regulated and induced by low temperatures. *PloS One* 11, e0151067. doi: 10.1371/journal.pone.0151067

Rastogi, S., Shah, S., Kumar, R., Vashisth, D., Akhtar, M. Q., Kumar, A., et al. (2019). Ocimum metabolomics in response to abiotic stresses: Cold, flood, drought and salinity. *PloS One* 14, e0210903. doi: 10.1371/journal.pone.0210903

R Core Team (2017). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/.

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

Saha, P., Chatterjee, P., and Biswas, A. K. (2010). NaCl Pretreatment alleviates salt stress by enhancement of antioxidant defense system and osmolyte accumulation in mungbean (*Vigna radiata* l. wilczek). *Indian J. Exp. Biol.* 48, 593–600.

Sarker, U., and Oba, S. (2018). Augmentation of leaf color parameters, pigments, vitamins, phenolic acids, flavonoids and antioxidant activity in selected amaranthus tricolor under salinity stress. *Sci. Rep.* 8, 12349. doi: 10.1038/s41598-018-30897-6

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Sharma, A., Shahzad, B., Rehman, A., Bhardwaj, R., Landi, M., and Zheng, B. (2019). Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. *Molecules* 24, 2452. doi: 10.3390/molecules24132452

Shen, X., Guo, X., Guo, X., Zhao, D., Zhao, W., Chen, J., et al. (2017). PacMYBA, a sweet cherry R2R3-MYB transcription factor, is a positive regulator of salt stress tolerance and pathogen resistance. *Plant Physiol. Biochem.* 112, 302–311. doi: 10.1016/j.plaphy.2017.01.015

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., et al. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists, (2021 update). *Nucleic Acids Res.* 10, gkac194. doi: 10.1093/nar/gkac194

Singh, B., Singh, J. P., Kaur, A., and Singh, N. (2017). Phenolic composition and antioxidant potential of grain legume seeds: A review. *Food Res. Int.* 101, 1–16. doi: 10.1016/j.foodres.2017.09.026

Singleton, V. L., and Rossi, J. A. (1965). Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol Vitic.* 16, 144–158.

Stagos, D. (2020). Antioxidant activity of polyphenolic plant extracts. *Antioxidants* 9, 19. doi: 10.3390/antiox9010019

Sugimoto, M., and Takeda, K. (2009). Proteomic analysis of specific proteins in the root of salt-tolerant barley. *Biosci Biotechnol Biochem.* 73, 2762–2765. doi: 10.1271/bbb.90456

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 Med Chemistry- Anti-Cancer Agents)* 20, 2216–2227. doi: 10.2174/1871520620666200516150829

Tian, F., Yang, D.-C., Meng, Y.-Q., Jin, J., and Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.* 48, D1104–D1113. doi: 10.1093/nar/gkz1020

Upadhyay, A., Gaonkar, T., Upadhyay, A. K., Jogaiah, S., Shinde, M. P., Kadoo, N. Y., et al. (2018). Global transcriptome analysis of grapevine (*Vitis vinifera* l.) leaves under salt stress reveals differential response at early and late stages of stress in table grape cv. Thompson seedless. *Plant Physiol. Biochem.* 129, 168–179. doi: 10.1016/j.plaphy.2018.05.032

Wang, G., Zhu, Q., Meng, Q., and Wu, C. (2012). Transcript profiling during salt stress of young cotton (Gossypium hirsutum) seedlings *via* solexa sequencing. *Acta Physiol. Plant* 34, 107–115. doi: 10.1007/s11738-011-0809-6

Wei, Q., Zhang, F., Sun, F., Luo, Q., Wang, R., Hu, R., et al. (2017). A wheat MYB transcriptional repressor TaMyb1D regulates phenylpropanoid metabolism and enhances tolerance to drought and oxidative stresses in transgenic tobacco plants. *Plant Sci.* 265, 112–123. doi: 10.1016/j.plantsci.2017.09.020

Xu, Z., Gongbuzhaxi,, Wang, C., Xue, F., Zhang, H., and Ji, W. (2015). Wheat NAC transcription factor TaNAC29 is involved in response to salt stress. *Plant Physiol. Biochem.* 96, 356–363. doi: 10.1016/j.plaphy.2015.08.013

Yang, Y., and Guo, Y. (2018). Elucidating the molecular mechanisms mediating plant salt-stress responses. *New Phytol.* 217, 523–539. doi: 10.1111/nph.14920

Yang, L., Wen, K.-S., Ruan, X., Zhao, Y.-X., Wei, F., and Wang, Q. (2018). Response of plant secondary metabolites to environmental factors. *Molecules* 23, 762. doi: 10.3390/molecules23040762

Yen, G.-C., and Chen, H.-Y. (1995). Antioxidant activity of various tea extracts in relation to their antimutagenicity. *J. Agric. Food Chem.* 43, 27–32. doi: 10.1021/jf00049a007

Yin, Y., Xu, J., He, X., Yang, Z., Fang, W., and Tao, J. (2022). Role of exogenous melatonin involved in phenolic acid metabolism of germinated hulless barley under NaCl stress. *Plant Physiol. Biochem.* 170, 14–22. doi: 10.1016/j.plaphy.2021.11.036

Yoseph Ganta, T., Mekbib, D. F., and Tadele, P. Z. (2021). *Genetic diversity, drought tolerance, and genotype by environment interaction of mung bean [vigna radiata (l.) wilczek] genotypes in ethiopia* (Ethiopia: University of Haramaya).

Yuan, G., Wang, X., Guo, R., and Wang, Q. (2010). Effect of salt stress on phenolic compounds, glucosinolates, myrosinase and antioxidant activity in radish sprouts. *Food Chem.* 121, 1014–1019. doi: 10.1016/j.foodchem.2010.01.040

Yun, Y. B., Jung, H.-J., Rahim, M. A., Park, J.-I., and Kuk, Y. I. (2019). Molecular analysis of genes related to phenylpropanoid and ascorbate biosynthesis in salt and UV-b treated pak choi grown under LEDs. *Botany* 97, 513–519. doi: 10.1139/cjb-2018-0183

Zhishen, J., Mengcheng, T., and Jianming, W. (1999). The determination of flavonoid contents in mulberry and their scavenging effects on superoxide radicals. *Food Chem.* 64, 555–559. doi: 10.1016/S0308-8146(98)00102-2

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 25 November 2022
DOI 10.3389/fpls.2022.1043489

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Xiaoming Song,
North China University of Science and Technology, China
Xiaohua Li,
Wuhan Polytechnic University, China

*CORRESPONDENCE
Fengde Wang
wfengde@163.com
Jianwei Gao
jwg_738@163.com

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 13 September 2022
ACCEPTED 08 November 2022
PUBLISHED 25 November 2022

CITATION
Wang L, Zhang S, Li J, Zhang Y, Zhou D, Li C, He L, Li H, Wang F and Gao J (2022) Identification of key genes controlling soluble sugar and glucosinolate biosynthesis in Chinese cabbage by integrating metabolome and genome-wide transcriptome analysis.
Front. Plant Sci. 13:1043489.
doi: 10.3389/fpls.2022.1043489

Identification of key genes controlling soluble sugar and glucosinolate biosynthesis in Chinese cabbage by integrating metabolome and genome-wide transcriptome analysis

Lixia Wang[1], Shu Zhang[1], Jingjuan Li[1], Yihui Zhang[1], Dandan Zhou[1,2], Cheng Li[1], Lilong He[1], Huayin Li[1], Fengde Wang[1*] and Jianwei Gao[1*]

[1]Institute of Vegetables, Shandong Academy of Agricultural Sciences, Jinan, China, [2]College of Life Sciences, Shandong Normal University, Jinan, China

Introduction: Soluble sugar and glucosinolate are essential components that determine the flavor of Chinese cabbage and consumer preferences. However, the underlying regulatory networks that modulate the biosynthesis of soluble sugar and glucosinolate in Chinese cabbage remain largely unknown.

Methods: The glucosinolate and carotene content in yellow inner-leaf Chinese cabbage were observed, followed by the combination of metabolome and transcriptome analysis to explore the metabolic basis of glucosinolate and soluble sugar.

Results: This study observed high glucosinolate and carotene content in yellow inner-leaf Chinese cabbage, which showed a lower soluble sugar content. The differences between the yellow and the white inner-leaf Chinese cabbage were compared using the untargeted metabonomic and transcriptomic analyses in six cultivars of Chinese cabbage to explore the metabolic basis of glucosinolate and soluble sugar. Aliphatic glucosinolate and two soluble sugars (fructose and glucose) were the key metabolites that caused the difference in Chinese cabbage's glucosinolate and soluble sugar. By integrating soluble sugar and glucosinolate-associated metabolism and transcriptome data, we indicated *BraA05gAOP1* and *BraA04gAOP4*, *BraA03gHT7* and *BraA01gHT4* were the glucosinolates and soluble sugar biosynthesis structural genes. Moreover, *BraA01gCHR11* and *BraA07gSCL1* were two vital transcription factors that regulate soluble sugar and glucosinolate biosynthesis.

Discussion: These findings provide novel insights into glucosinolate and soluble sugar biosynthesis and a possible explanation for the significant difference in nutrients between yellow and white inner-leaf Chinese cabbage. Moreover, it will facilitate genetic modification to improve the Chinese cabbage's nutritional and health values.

KEYWORDS

soluble sugar, glucosinolate, metabolome analysis, transcriptome analysis, Chinese cabbage

Introduction

Chinese cabbage is one of the essential vegetables in the world. They are the most widely grown vegetables in China and northern areas and account for over one-quarter of the total annual vegetable consumption. The leafy head comprising numerous incurved leaves is the main edible organ. Soluble sugar, carotenoid and glucosinolate are the most common nutrient compounds affecting the flavor of Chinese cabbage (Ishida et al., 2014; Cao et al., 2021; Cheng et al., 2022). However, over the past few decades, intensive breeding has been mainly focused on yield and disease resistance rather than the quality of vegetables. With the increase in citizens' living standards, the market-driven orientation artificial selection strategy becomes more critical. Furthermore, functional Chinese cabbage has gained popularity within health and wellness circles in recent years with the increasing diversity. Therefore, Chinese cabbage has been the subject of much research to evaluate its nutrient compounds and characteristic flavors (Higdon et al., 2007).

Glucosinolate is sulfur- and nitrogen-containing plant secondary metabolite found in cruciferous vegetables, whose hydrolysis products contribute to the unique flavours and tastes of Brassica species (Shim et al., 2016; Bischoff, 2021). Glucosinolate is known for its beneficial effect on human health, such as strong anti-cancer effects and regulatory functions in inflammation and dampening the stress responses and antimicrobial properties (Hu et al., 2022; Jo et al., 2022; Tandayu et al., 2022). Aliphatic glucosinolate has been reported as the primary type of glucosinolate in *Brassica rapa* (Liao, 2011). Interestingly, aliphatic glucosinolate, positively correlated with carotenoid content in Chinese cabbage, indicated the synergistic regulation between these nutrient components (Baek et al., 2016). Content and biosynthesis of glucosinolate that affects fruit flavor have been widely studied in model *Arabidopsis* (Harun et al., 2020), which paved the way for elucidating the molecular mechanism of glucosinolate biosynthesis in Chinese cabbage. The biosynthesis of aliphatic glucosinolate originates from the elongation of methionine (Met). *AOPs* (2-oxoglutarate-dependent dioxygenases), *SUR1* (C-S lyase), *UGT74* (UDP-glycosyltransferases), *BCATs* (branched-chain amino acid aminotransferase), *CYP79s* (cytochrome P450 monooxygenases) and *CYP83s* (Kang et al., 2018; Zuluaga et al., 2019), *MAM* (Methylthioalkylmalate synthase) (Das, 2021), *SOT* (sulfotransferases) (Klein and Papenbrock, 2008) play a vital role in glucosinolate biosynthesis. Overexpression of *CYP79s* and *CYP83s*, which are involved in forming the core glucosinolate structure, increased the aliphatic glucosinolate level in Chinese cabbage (Zang et al., 2008). In addition, transcription factors, such as MYB, ERF and bHLH, are all involved in the regulation of glucosinolate biosynthesis (Chhajed et al., 2020; Mitreiter and Gigolashvili, 2021). In the previous report, *MYB122* (transcription factor), *CYP79B2*, *UGT74B1*, *SUR1*, *SOT16*, *SOT17*, *SOT18* (core structure biosynthesis genes), and IGMT1 (indole glucosinolate side chain modification gene) play a differentiated role in the transcriptional response pattern of JA (jasmonic acid) in two broccoli cultivars (Ku et al., 2016). Therefore, these metabolite and transcript biomarkers could be helpful in an effective marker-assisted breeding strategy for improving the resistance in Brassica vegetables. Another study indicated three *MYB28* homologs regulating glucosinolate biosynthesis in Chinese kale sprouts (Guo et al., 2016).

Soluble sugar, including sucrose, glucose and fructose, is one of the essential qualitative traits to evaluate the shelf-life of Chinese cabbage and is also the vital element that provides energy and bare carbon skeletons for various metabolic pathways (Shen et al., 2018; Liu et al., 2020). In addition, soluble sugar can act as signal transduction molecules to regulate the development and adaptation to environmental challenges (Durán-Soria et al., 2020). Even though soluble sugar significantly influences the sensory quality, few studies on the sweetness of heading leaves in Chinese cabbage have yet been reported. The sweetness of vegetables is mainly related to the composition and type of soluble sugar. The biosynthesis and metabolism of soluble sugar were involved in many biological processes regulated by the expression of many genes of enzymes. INV (Invertase), SUS (sucrose synthase) and SPS (sucrose

phosphate synthase) have been recognized as the essential enzymes affecting the metabolism and accumulation of soluble sugar (Nookaraju et al., 2010; Matsukura, 2016). The biosynthesis of sucrose is probably regulated by PFK6 and SUS1/SuSy1 in Chinese cabbage, and it is also regulated by SPS4, a crucial enzyme in sucrose synthesis (Li et al., 2019). During the soluble sugar biosynthesis and metabolic process of Chinese cabbage, HxK (hexokinase) and FRK (fructokinase) were involved in the further utilization of sucrose degradation products, including glucose and fructose (Moscatello et al., 2011; Hu et al., 2016).

Even though the soluble sugar and glucosinolate have the most significant influence on the flavor and quality (Wang et al., 2011; Vallone et al., 2013), their regulatory mechanism in Chinese cabbage is still unclear. Compared to white inner-leaf Chinese cabbage, yellow inner-leaf Chinese cabbage showed high glucosinolate content and low soluble sugar. Therefore, it provided an ideal experimental material to study the biosynthesis mechanism of soluble sugar and glucosinolate in Chinese cabbage. Integration of transcriptome and metabolome information offers unique insights into pathways associated with agronomy traits while identifying potential targets for genetic modification. To clarify the regulatory pathways of glucosinolate and soluble sugar in Chinese cabbage, we first investigated the differences in the accumulation of glucosinolate and soluble sugar between the yellow and the white inner-leaf Chinese cabbage, and the expression of biosynthesis-related genes of glucosinolate and soluble sugar was quantified. Then, to explain nutrient differences between the yellow and the white inner-leaf Chinese cabbage, we used conjoint transcriptome and metabolome analysis to identify different metabolites and DEGs (differentially expressed genes) in glucosinolate and soluble sugar biosynthesis. In addition, a series of physiological tests were carried out to evaluate the commercial value of the varieties. Besides, the qRT-PCR (Quantitative Real-time Polymerase Chain Reaction) was used to verify the transcriptome analysis results. Our study can provide new insights into understanding the nutrient formation of Chinese cabbage.

Material and methods

Plant materials and treatments

Three varieties of yellow inner-leaf Chinese cabbage, Wawahuang (WWH), Xiqing (XQ), Gaochunhuanag (GCH) and three types of white inner-leaf Chinese cabbage, Rewangzi (RWZ), Xiayangwang (XYW), Rejiangjun (RJJ) were grown and collected from an experiment field of Shandong Academy of Agricultural Sciences (Jinan city, Shandong Province) in 2021. The middle third of the edible part was cut as selected material for measuring of total soluble solids, total carotenoid, total glucosinolate, the metabolomics and RNA-Seq analysis, and qRT-PCR.

Seeds of "XQ" and "XYW" were germinated in a plastic pot (10 ×10 cm) containing soil and vermiculite at a volume ratio of 1:3 and grown under controlled conditions, 20°C ± 2, in a plant culture room in the institute of Vegetables, Shandong Academy of Agricultural Sciences. One-month-old seedlings with 5-7 leaves were selected, and the whole plants were cut up for the qRT-PCR analysis.

We used three biological replicates for one experiment, and each replicate consisted of three Chinese cabbage. Therefore, three Chinese cabbage were cut up, mixed as one sample and three biological replicates for one experiment. The samples were frozen in liquid nitrogen and stored at -80°C for further research.

Total carotenoid content measurement

1.0 g fresh Chinese cabbage was ground in 50 ml 100% acetone and extracted in the dark at 4°C in the dark. The absorbance of the extraction solution was measured in A663, A645, and A470 (Cao et al., 2007). We calculated the total carotenoid by the following formula:

$$\text{Chlorophyll a } (\text{mg/L}) = 12.21 \times A_{663} - 2.81 \times A_{646}$$

$$\text{Chlorophyll b}(\text{mg/L}) = 20.13 \times A_{646} - 5.03 \times A_{663}$$

$$\text{Carotenoids}(\text{mg/L}) = \frac{(1000 \times A_{470} - 3.27 \times \text{Chlorophyll a} - 104 \times \text{Chlorophyll b})}{229}$$

Total soluble sugar content measurement

1.0 g fresh Chinese cabbage was ground and rinsed with 5 to 10 ml ddH_2O, then the extraction was filtered and boiled for 30 min, and the tube and residue were flushed to the final volume (100 ml). In the next step, we added 0.5 ml extract solution and 1.5 ml distilled water to a 25 ml glass tube with 0.5 ml anthrone-ethyl acetate and 5 ml concentrated sulfuric acid. Later, we boiled the sample for 1 min. Finally, we measured the absorbance of the reaction in A630 (Cao et al., 2007).

Total glucosinolate content measurement

We dried 10 g of Chinese cabbage at 120°C for 15 minutes and then incubated it at 80°C for 2 days. Then, the samples were ground into powder and filtered by 100 mesh to remove the large debris. First, we added 0.1 g powder to 1.5 ml of 90% ethanol and

incubated at it 70°C for 45 min. Then we filtered and washed using ddH_2O to 10 ml. Later, we added 1.0 ml solution with 2.0 ml 4 mmol/L PdCl and kept it at room temperature for 2 hours. Finally, we measured the absorbance of the reaction in A540 (Chelimoge et al., 2014).

Metabolomics analysis of soluble sugar in Chinese cabbage

Soluble sugar content and composition were detected by MetWare (Wuhan, China, http://www.metware.cn/). An electronic mill ground the freeze-dried samples at 30 Hz for 1.5 min. 500 µL mixture with the volume ratio of methanol, isopropanol and water (3:3:2 V/V/V) was used to dissolve 20 mg sample powder and vortexed for 3 min and ultrasound for 30 min. After centrifugation at 14,000 rpm under 4°C for 3 min. The supernatant was evaporated under an N_2 stream, and then lyophilized. The residue was used for further derivatization. A gas chromatograph and mass spectrometer (GC-MS) were used to analyse Chinese cabbage's soluble sugar. 1.0 mL/min of pure Helium was used as a carrier gas. 1 µL gas was injected with a split ratio of 5:1. The heating program of the column oven was conducted as follows: starting at 170°C for 2 min, and then raised to 240°C at 10°C/min, the temperature increased to 280°C at 5°C/min, the temperature increased at 25°C/min to 310°C, and maintained at 310°C for 4 min. We used a selective ion monitoring mode to analyze all of the samples. The ion source and transfer line temperatures were 230°C and 240°C, respectively.

Metabolomics analysis of glucosinolate in Chinese cabbage

Glucosinolate content and composition were detected by MetWare (Wuhan, China, http://www.metware.cn/). The sample preparation was the same as the soluble sugar measurement. 1.2 mL 70% methanol (70:30, v/v) was added into 100 mg lyophilized powder and blended. Extraction was carried out in a refrigerator at 4°C overnight. The samples were centrifuged at 12000 rpm for 10 min to remove the undissolved residue, and then were filtrated *via* SCAA-104 with 0.22 µm pore size (ANPEL, Shanghai) before being done UPLC-MS/MS analysis. Glucosinolate metabolites were measured by a UPLC-ESI-MS/MS system (UPLC, SHIMADZU Nexera X2; MS, Applied Biosystems 4500 Q TRAP) equipped with SB-C18 (1.8 µm, 2.1 mm × 100 mm). Solvent A (pure water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid) were used as the UPLC mobile phase. Sample measurements were carried out with a gradient program with 0.35 mL/min flow rate and 40°C of the column oven temperature. The injection volume was 4 µL. The gradient program was started with 95% A:5% B, followed by 5% A: 95% B at 0-10 min, and at 11.1-14 min, 95% A: 5% B. The effluent was connected to an ESI-triple quadrupole-linear ion trap (QTRAP)-MS. We carried linear ion trap (LIT), and triple quadrupole (QQQ) scans on an AB4500 Q TRAP UPLC/MS/MS System equipped with an ESI Turbo Ion-Spray interface, operating in both positive and negative ion mode and controlled by the Analyst 1.6.3 software (AB Sciex). The ESI conditions were set as follows: a turbo spray ion source, 5500 V (positive ion mode)/-4500 V (negative ion mode) and 550°C; ion source gas I (GSI), gas II (GSII) and curtain gas (CUR) was set at 50, 60, and 25.0 psi, respectively; the collision-activated dissociation (CAD) was high. 10 and 100 µmol/L polypropylene glycol solutions in QQQ and LIT were conducted for Instrument tuning and mass calibration, respectively. The nitrogen was placed in the medium for the MRM experiment. Declustering potential (DP) and collision energy (CE) was done for each MRM transition.

RNA-Seq analysis

The RNA-Seq was performed by MetWare (Wuhan, China). Total RNA was extracted from six Chinese cabbage. A total amount of 1 µg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, USA). After the raw data were filtered, the sequencing error rate and the GC content distribution were checked, clean reads were obtained for subsequent analysis, and the mapped data were obtained by sequence alignment with the Chinese cabbage reference genome (http://brassicadb.cn). FPKM (fragments per kilobase of transcript per million fragments) values were used to indicate transcript or gene expression levels. The original count data were analyzed by using DESeq2 v1.22.1 software. After difference analysis, the hypothesis test probability (P-value) was corrected using the Benjamini-Hochberg method, and the FDR (the false discovery rate) was obtained. DEG screening criteria were a $|\log 2\text{fold change}| \geq 1$ and an FDR< 0.05. The enrichment analysis is based on the hypergeometric test using KEGG (Kyoto Encyclopedia of Genes and Genomes), and the hypergeometric distribution test is performed with the unit of the pathway. In addition, the genes showing expression values (an averaged NRPKM from three replicates) were higher than one was selected, and WGCNA (weighted gene co-expression network analysis) was performed using the WGCNA v1.69 software package.

qRT-PCR analysis

The total RNA of six Chinese cabbage flesh frozen samples were extracted and used as a template, and a Takara Kit (PrimeScript 1st strand cDNA Synthesis Kit) was used to

reverse-transcribe RNA into cDNA. Reactions were carried out on a Roche LightCycler 96 qRT-PCR detection system. The analysis of each sample was repeated three times, and the $2^{-\Delta CT}$ method was used for quantitative data analysis. We used *Actin* as an internal reference gene. In this study, all the primers (primers synthesized by Qingdao WeiLai Biotechnology Co., Ltd.) are shown in Supplementary Material 2- Supplementary Table S1.

Statistical analysis

KEGG annotation and enrichment analysis were used to test the statistical enrichment of the DEGs in KEGG pathways. Three biological replicates were performed in all the experiments in this study. Statistical significance (Student's t-tests) and Pearson correlation coefficients were analyzed by using SPSS v24.0 software (SPSS Inc., Chicago, IL, USA), and a difference was considered to be statistically significant when $P \leq 0.05$.

Results

Differences in carotenoid, soluble sugar and glucosinolate content between yellow and white inner-leaf Chinese cabbage

Six representative Chinese cabbage cultivars with yellow or white inner-leaf color, WWH, XQ, GCH, RWZ, XYW and RJJ, were selected for this study. The inner-leaf color of WWH, XQ and GCH are yellow, while those of RWZ, XYW and RJJ are white (Figure 1A). The content of total carotenoid, total soluble

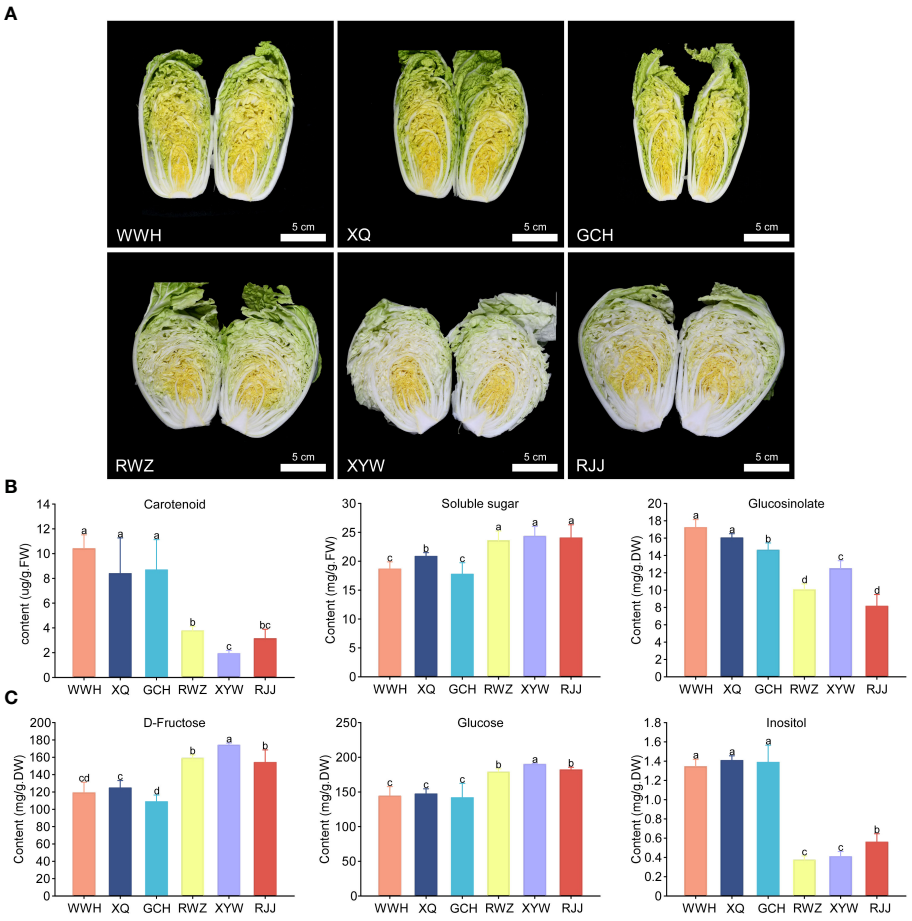


FIGURE 1
The contents of carotenoid, soluble sugar and glucosinolate in yellow and white inner-leaf Chinese cabbage. **(A)** the phenotype of yellow (WWH, XQ, GCH) and white (RWZ, XYW, RJJ) inner-leaf Chinese cabbage. Scale bars are 5 cm. **(B)** Carotenoid, Soluble sugar and Glucosinolate content. **(C)** D-Fructose, Glucose and Inositol content. Data are means ± SD, N = 3. Different letters indicate significant differences among six Chinese cabbage by t-tests ($p \leq 0.05$).

sugar and total glucosinolate in these six cultivars of Chinese cabbage were measured (Figure 1B). The yellow inner-leaf Chinese cabbage observed a significantly high total carotenoid content. Similarly, the glucosinolate content in the yellow inner-leaf Chinese cabbage was generally higher than that in the white inner-leaf Chinese cabbage. Meantime, we measured soluble sugar content that could affect the taste and sweetness. We found that the rule of total soluble sugar content in yellow and white inner-leaf Chinese cabbage was opposite to carotenoid and glucosinolate content, showing a significantly high soluble sugar content in white inner-leaf Chinese cabbage.

Qualitative and quantitative analysis of the soluble sugar and glucosinolate

To investigate the soluble sugar content and type in Chinese cabbage, we measured 13 soluble sugars using HPLC-MS/MS. As a result, there were 10 soluble sugars, including Maltose, Sucrose, Trehalose, D-Arabinose, L-Fucose, Glucose, Inositol, L-Rhamnose, and Xylitol were identified in white and yellow inner-leaf Chinese cabbage (Supplementary Material 1). In particular, D-fructose and glucose were the major components of soluble sugar in Chinese cabbage, accounting for more than 96% of the total soluble sugar content. Therefore, it is speculated that the significant difference between D-fructose and glucose content leads to the substantial difference in total soluble sugar content between the yellow and the white inner-leaf Chinese cabbage ($p \leq 0.05$). In addition, inositol is a kind of hygienic component. However, the total content was relatively low, significantly higher in yellow inner-leaf Chinese cabbage than in white inner-leaf Chinese cabbage (Figure 1C).

Based on fold change ≥ 2 or ≤ 0.5 and VIP ≥ 1, the glucosinolate metabolites in the yellow and the white inner-leaf Chinese cabbage were quantitatively and qualitatively analyzed. Fifty-three glucosinolate metabolites were identified and divided into four classes, including thirty-eight aliphatic glucosinolates, nine aromatic glucosinolates, four indole glucosinolate, and one thiocyanate (Supplementary Material 1). Among the 53 glucosinolates, 5-Hexenyl Glucosinolate, 3-Methylamylthioglycoside, N-Hexyl glucoside, and 2-Hydroxy-4-Pentenylglucosinolate in yellow inner-leaf Chinese cabbage was generally much higher than that in white inner-leaf Chinese cabbage (Table 1). Furthermore, the qualitative analysis showed that all four glucosinolates belong to aliphatic glucosinolate. Therefore, we speculated that the difference in total glucosinolate content between yellow and white inner-leaf Chinese cabbage was mainly caused by the significant difference in the content of these four aliphatic glucosinolates.

TABLE 1 Compares glucosinolate content between yellow inner-leaf Chinese cabbage and white inner-leaf Chinese cabbage.

Compounds	Fold change								
	WWH/ RWZ	XQ/ RWZ	GCH/ RWZ	WWH/ XYW	XQ/ XYW	GCH/ XYW	WWH/ RJJ	XQ/ RJJ	GCH/ RJJ
5-Hexenyl Glucosinolate	4.15	4.43	5.09	6.37	6.81	7.81	3.83	4.09	4.69
N-Hexyl glucoside	88500*	104000*	112000*	4.34	5.08	5.49	2.98	3.49	3.77
3-Methylamylthioglycoside	3.09	3.52	3.93	4.65	5.30	5.91	3.14	3.58	3.99
2-Hydroxy-4-Pentenylglucosinolate	2.67	2.44	2.59	2.84	2.59	2.74	3.35	3.06	3.24
3-Hydroxy-5-(methylthio)pentyl Glucosinolate	0.18	0.28	0.15	0.29	0.44	0.24	0.23	0.35	0.19
3-(Methylthio)propyl glucosinolate	0.27	0.39	0.31	0.25	0.36	0.29	0.20	0.29	0.23
2(R)-Hydroxy-2-Phenylethyl Glucosinolate	0.23	0.24	0.23	0.30	0.31	0.30	0.19	0.19	0.19
5-(Methylsulfinyl)amyl glucoside	0.17	0.24	0.14	0.29	0.42	0.23	0.18	0.26	0.15
4-Methylthiobutyl glucosinolate (Glucoerucin)	0.06	0.11	0.03	0.04	0.06	0.02	0.04	0.06	0.02
5-Methylthiopentyl glucosinolate (Glucoberteroin)	0.25	0.32	0.14	0.29	0.36	0.16	0.26	0.32	0.14
Sulforaphane (4-methylsulphinylbutyl glucosinolate)	0.15	0.19	0.11	0.73	0.09	0.06	0.10	0.13	0.08
4-Methylsulfinylbutyl glucosinolate (Glucoraphanin)	0.04	0.06	0.03	0.04	0.07	0.03	0.03	0.05	0.02
Glucocheirolin	0.15	0.19	0.15	0.17	0.21	0.17	0.10	0.13	0.10
3-Phenylpropyl Glucosinolate	0.23	0.29	0.12	0.25	0.31	0.13	0.19	0.24	0.10

*means not detected in RWZ.

DEGs between yellow and white inner-leaf Chinese cabbage

Based on the metabolomic analysis and the content of soluble sugar and glucosinolate, the significant differences focused on soluble sugar and glucosinolate biosynthesis pathway between yellow and white inner-leaf Chinese cabbage. Therefore, we performed RNA-Seq on yellow and white inner-leaf Chinese cabbage to study the molecular regulatory mechanisms of soluble sugar and glucosinolate biosynthesis. The 18 transcriptome samples produced 125.2 Gb Clean Data, more than 6.0 Gb per sample, with a percentage of Q30 bases above 93%. The clean reads were compared and annotated. To better elucidate the difference between yellow and white inner-leaf Chinese cabbage, the Pearson correlation coefficient (r^2, PCC) of RNA-Seq datasets was conducted to analyze the correlation of samples (Figure 2A). The closer the r^2 is to 1, the stronger the correlation between the replications of samples. The results showed a good correlation between each sample in this study was good. The PCA (principal component analysis) score plot showed that yellow inner-leaf Chinese cabbage exhibited an apparent separation from white inner-leaf Chinese cabbage, and three biological replicates of each variety were compactly gathered together (Supplementary Material 2-Supplementary Figure 1A), indicating that the experiment was reproducible and reliable. This comparison showed significant differences between yellow and white inner-leaf Chinese cabbage ($p \leq 0.05$). In addition, 18 samples were divided into four groups in the cluster and correlation analysis on the heatmap (Supplementary Material 2-Supplementary Figure 1B), indicating significant differences in the contents of metabolites in yellow and white inner-leaf Chinese cabbage. These results suggested enormously different metabolite profiles in different inner-leaf color Chinese cabbage.

FPKM values were calculated for each sample to normalize the expression and further investigate the gene expression patterns. A pairwise comparison was made between yellow inner-leaf Chinese cabbage and white inner-leaf Chinese cabbage. The results showed that the most DEGs were observed in the comparison of WWH/RWZ, and the fewest DEGs were found in comparing GCH/XYW (Figure 2B). A total of 1163 DEGs (padj-value< 0.05, |log2Fold Change| ≥ 1), including 571 up-regulated genes and 592 down-regulated genes, were detected between yellow and white inner-leaf Chinese cabbage (Figure 2C). To analyze the function of common DEGs during soluble sugar and glucosinolate biosynthesis, we carried KEGG enrichment analysis between white and yellow inner-leaf Chinese cabbage (Figure 2D). These 1163 DEGs were enriched mainly in the metabolic pathway (ko01100) related to 203 DEGs, followed by the biosynthesis of secondary metabolites (ko01110), which underlies the soluble

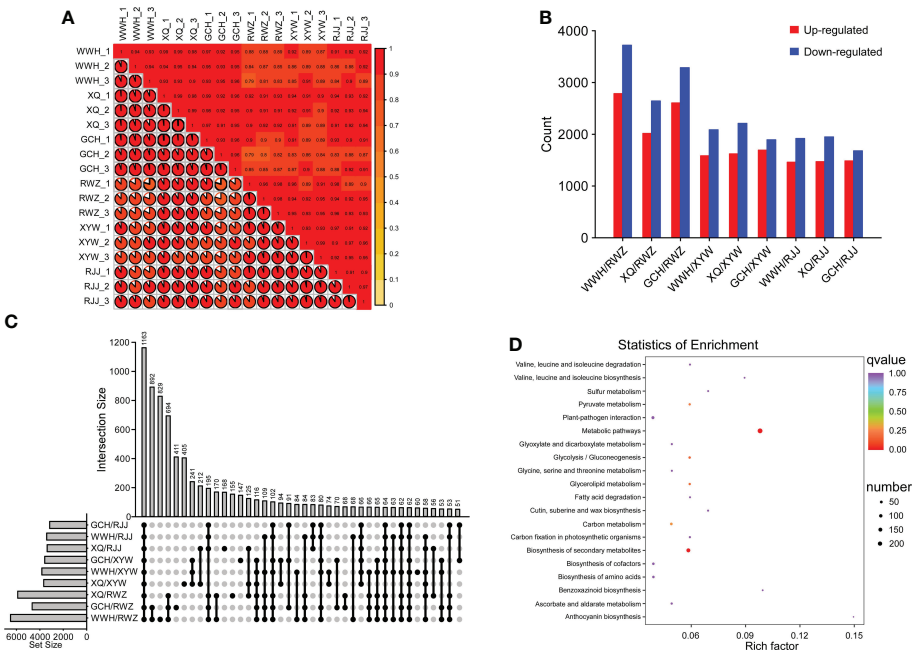


FIGURE 2
Preliminary analysis of transcriptome data. **(A)** heat map of correlation in different comparison groups. **(B)** The number of DEGs in each group. **(C)** The number of common DEGs in white and yellow inner-leaf Chinese cabbage (|log2Fold Change| ≥ 1, FDR< 0.05). **(D)** KEGG enrichment analysis of DEGs between the comparison groups (Yellow vs White inner-leaf Chinese cabbage). Each bubble in the plot represents a metabolic pathway. A larger bubble size indicates a larger impact factor. Darker bubble colors represent a higher degree of enrichment.

sugar and glucosinolate difference between white and yellow inner-leaf Chinese cabbage.

To gain further insight into the regulation of soluble sugar and glucosinolate biosynthesis, we carried out WGCNA to investigate the co-expression gene modules and the critical modules involved in soluble sugar and glucosinolate biosynthesis. A total of 18 co-expression modules were identified according to their expression patterns (Figure 3A). The correlation between the gene matrix of different modules and samples of Chinese cabbage was analyzed, and the correlation and corresponding e-value were presented in a digital form in the grid where each module and trait intersect (Figure 3A). According to the 'module character' correlation analysis, the blue module showed a significant positive correlation with all yellow inner-leaf Chinese cabbage and a negative correlation with all white inner-leaf Chinese cabbage. In addition, we carried out KEGG enrichment analyses of the blue module (Figure 3B). These DEGs in the blue module were enriched mainly in the metabolic pathway, followed by the biosynthesis of secondary metabolites. We found that more than 60% of DEGs between the yellow and the white inner-leaf Chinese cabbage enriched in the metabolic pathway and secondary metabolite synthesis pathway in Figure 3 existed in the blue module, signifying blue module should be strongly correlated to the metabolism of the Chinese cabbage, including soluble sugar and glucosinolate metabolism. These results indicated that the blue module is the critical module regulating soluble sugar and glucosinolate biosynthesis in Chinese cabbage.

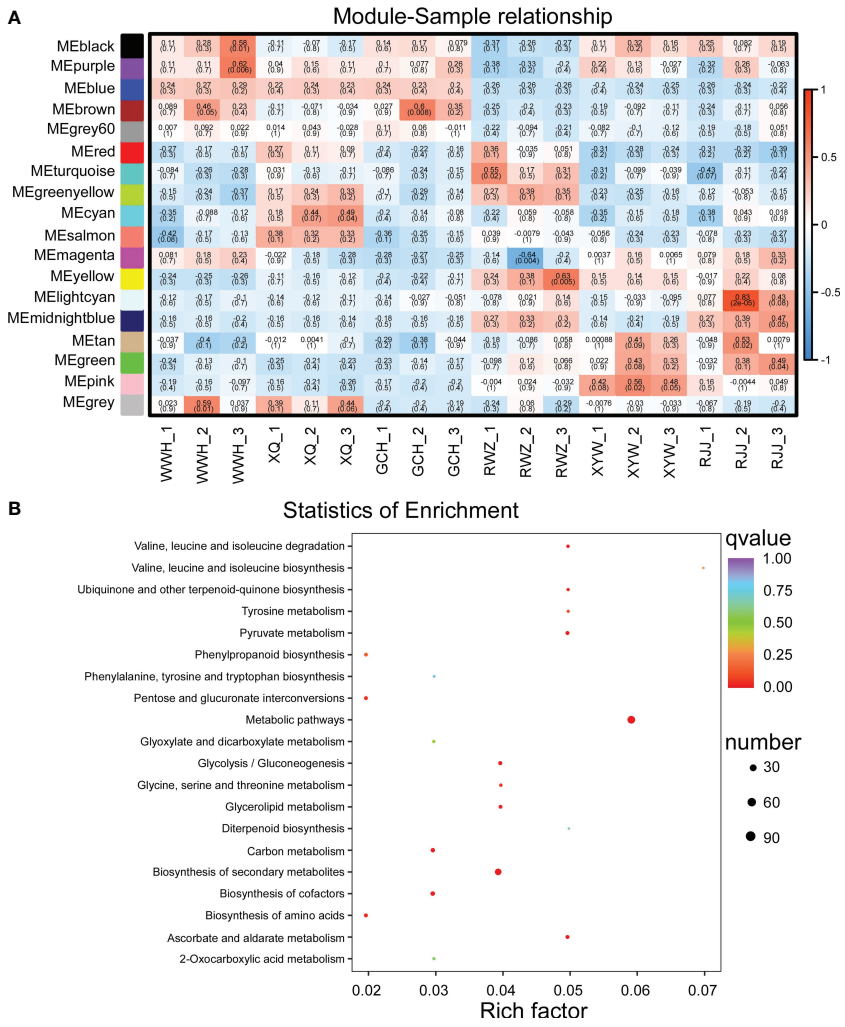


FIGURE 3
WGCNA of RNA-Seq data. **(A)** Eigen gene adjacency heatmap. **(B)** KEGG enrichment of DEGs in the blue module. Each bubble represents a metabolic pathway. A larger bubble size indicates a more considerable impact factor. The bubble colors represent the p-values of the enrichment analysis.

Analysis of soluble sugar biosynthetic pathway in Chinese cabbage

To generate the regulatory network associated with soluble sugar biosynthesis, we constructed the soluble sugar metabolic pathway in Figure 4A, and examined the structural genes involved in soluble sugar biosynthesis. The structural genes and transcription factors were organized into a connection network using Cytoscape software (Figures 4B, C). Among 1163 DEGs, six structural genes strongly correlated to the biosynthesis of soluble sugar were identified (Pearson correlation coefficients > 0.8, p-value< 0.05), including *BraA06gSPS2*, *BraA03gFRK1*, *BraA09gFRK3*, *BraA01gHT4* (hexose transporter), *BraA03gHT7*, *BraA03gHT3*. According to the analysis of gene function and RNA-Seq, *BraA01gHT4* and *BraA03gHT7* were positively correlated to the soluble sugar content, while the negative correlations were shown with *BraA03gFRK1*, *BraA09gFRK3*, *BraA06gSPS2* and *BraA03gHT3* in Chinese cabbage (Figure 4D). According to the soluble sugar metabolic pathway, HT mainly transported fructose and glucose from the cytoplasm to the vacuole. *BraA01gHT4* and *BraA03gHT7* encode the HT enzymes, regulating fructose and glucose metabolism. *BraA03gHT7* is in the blue module, and *BraA01gHT4* is clustered in the green module. According to weight value, the top ten transcription factors were selected from 1163 DEGs in the blue and green modules, respectively. In the blue modules, these ten transcription factors include *BraA01gCHR11* (CHROMATIN-REMODELING PROTEIN 11), *BraA05gRAP2.2* (RELATED TO AP2.2), *BraA07gSCL1* (SCARECROW-LIKE 1), *BraA09gNLP6* (NIN-LIKE PROTEIN 6), *BraA05gWRKY23*, *BraA03gSPL3* (SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 3), *BraA04gFRS5* (FAR1-RELATED SEQUENCE 5), *BraA06gAGL97* (AGAMOUS-LIKE 97), *BraA09gTGA1* (TGACG SEQUENCE-SPECIFIC BINDING PROTEIN 1), and *BraA09gMYB1R1* and include *BraA01gNF-YC11* (NUCLEAR FACTOR Y, SUBUNIT C11), *BraA02gARR18* (RESPONSE REGULATOR 18), *BraA02gMYB34*, *BraA03gNAC41* (NAC DOMAIN CONTAINING PROTEIN 41), *BraA06gC3H6* (Zinc finger C-x8-C-x5-C-x3-H type family protein), *BraA06gNF-YC4*, *BraA07gGPRI1* (GBF'S PRO-RICH REGION-INTERACTING FACTOR 1), *BraA07gbHLH77* (Basic helix-loop-HELIX 77), *BraA07gZFP1* (ZINC-FINGER PROTEIN 1), and *BraA08gSNF2* (Helicase protein with RING/U-box domain-containing protein) in the green module. The structural genes and transcription factors were organized into a connection network using Cytoscape software. Interestingly, four common transcription factors, including *BraA01gCHR11*, *BraA05gRAP2.2*, *BraA07gSCL1*, and *BraA09gNLP6*, were recognized both in the biosynthesis of glucosinolate and soluble sugar in the blue module, whose expression was highly correlated with the glucosinolate and soluble sugar content (Figures 4E), suggested that these four transcription factors not only correspond to the regulation of glucosinolate biosynthesis but also the regulated soluble sugar biosynthesis.

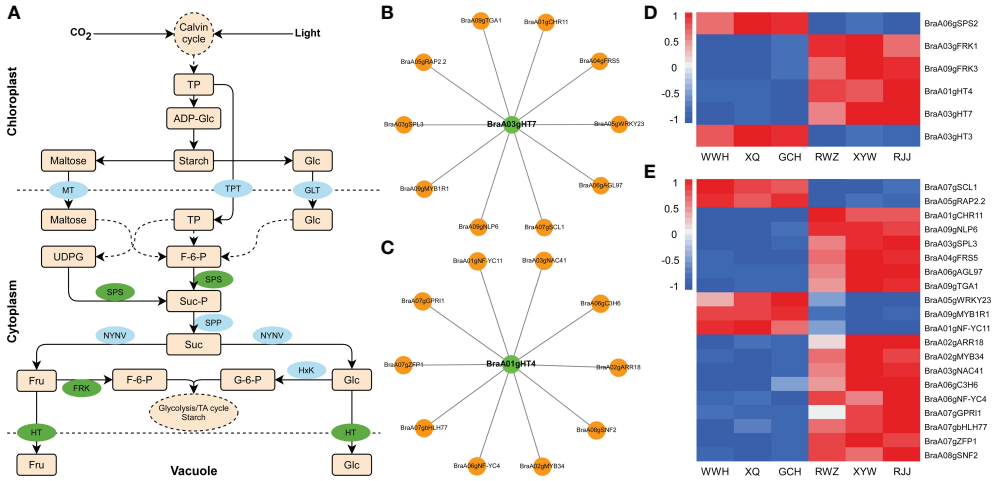


FIGURE 4
Pathway related to soluble sugar biosynthesis in Chinese cabbage. **(A)** Reconstruction of the soluble sugar biosynthetic pathway with the differentially expressed structural genes. Green ellipses represent essential genes we identified by RNA-Seq, and blue ellipses represent the genes involved in the glucosinolate biosynthesis. **(B)** Heatmap of the DEGs involved in soluble sugar biosynthetic pathway. **(C)** Correlation network of essential gene and transcription factors in the blue module. **(D)** Correlation network of structural genes and transcription factors in the green module. Green circles represent structural genes, and orange circles represent TFs. **(E)** The expression of transcription factors involved in soluble sugar biosynthesis in RNA-Seq result.

Analysis of the glucosinolate biosynthetic pathway in Chinese cabbage

To further investigate the regulation mechanism underlying glucosinolate between yellow and white inner-leaf Chinese cabbage, we analyzed the aliphatic glucosinolate biosynthetic pathway to perform the critical genes in the metabolism of yellow and white inner-leaf Chinese cabbage. The biosynthesis of aliphatic glucosinolate is derived from Met. *BCAT*, *MAM*, *CYP79/CYP83*, *AOP*, *SOT*, and *UGT74* regulate aliphatic glucosinolate biosynthesis (Figure 5A). Among 1163 DEGs, seven structural genes involved in aliphatic glucosinolate biosynthesis, including *BraA05gUGT74F1*, *BraA03gAOP1.2*, *BraA03gAOP1.1*, *BraA06gBCAT2*, *BraA02gSOT12*, *BraA04gAOP4*, and *BraA05gAOP1* were identified (Figure 5B). Correlation analysis was conducted between the expression of these structural genes and aliphatic glucosinolates, significantly different in yellow and white inner-leaf Chinese cabbage (PCC were calculated using the cor function in R). *BraA05gAOP1*, *BraA02gSOT12*, *BraA06gBCAT2*, and *BraA04gAOP4* located in the blue module were positively correlated to the glucosinolate content in yellow inner-leaf Chinese cabbage (PPC>0.8, p-value<0.05). While *BraA05gUGT74F1*, *BraA03gAOP1.2*, and *BraA03gAOP1.1* were negatively correlated to the glucosinolate content in yellow inner-leaf Chinese cabbage.

The top ten transcription factors in the blue modules were selected based on weight values from the 1163 DEGs. The visualization network between structural genes and transcription factors (Cytoscape software) was shown in Figure 5C. Among these ten transcriptions factors, eight of them, including *BraA08gGATA28* (BraA08g002520.3C), *BraA08gSOP1* (SUPPRESSOR OF PAS2 1), *BraA09gAREB3* (ABA-RESPONSIVE ELEMENT BINDING PROTEIN 3), *BraA10gAGL25*, *BraA02gAGL70*, *BraA05gCHR10*, *BraA05gRAP2.2*, and *BraA07gSCL1* were up-regulated (Figure 5D). At the same time, *BraA01gCHR11* and *BraA09gNLP6* were down-regulated in yellow inner-leaf Chinese cabbage, suggesting positive and negative regulation of glucosinolate biosynthesis, respectively (Figure 5D). These results indicated that these ten transcription factors correspond to the putative regulators controlling glucosinolate biosynthesis in Chinese cabbage.

Confirmation of the transcriptome data using qRT-PCR

To validate the transcriptome data, the expression pattern of the structural genes which we identified related to the biosynthesis of soluble sugar (*BraA03gHT7* and *BraA01gHT4*)

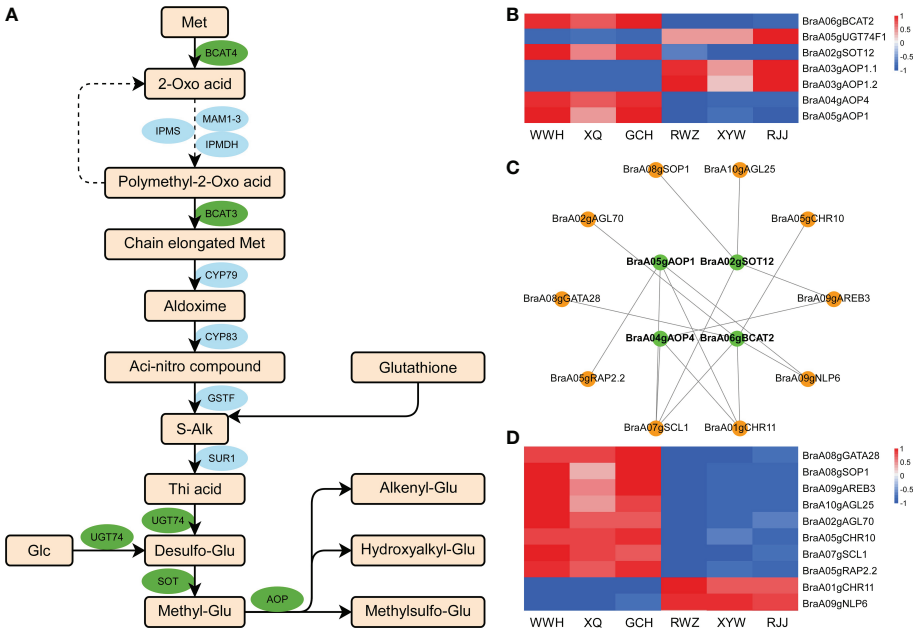


FIGURE 5
Aliphatic glucosinolate biosynthesis in Chinese cabbage. **(A)** Reconstruction of the aliphatic glucosinolate biosynthetic pathway with the differentially expressed structural genes. Green ellipses represent essential genes we identified by RNA-Seq; blue ellipses represent the genes involved in the glucosinolate biosynthesis. **(B)** Heatmap of the DEGs involved in aliphatic glucosinolate biosynthesis. **(C)** Correlation network of crucial genes involved in aliphatic glucosinolate biosynthesis in yellow and white inner-leaf Chinese cabbage. Green circles represent structural genes; orange circles represent TFs. **(D)** The expression of transcription factors involved in glucosinolate biosynthesis in RNA-Seq result.

and glucosinolate (*BraA06gBCAT2*, *BraA02gSOT12*, *BraA04gAOP4*, and *BraA05gAOP1*) was investigated in the yellow and the white inner-leaf Chinese cabbage by using the qRT-PCR (Figure 6). The relative expression of these structural genes agrees with the RNA-seq results, indicating consistency in the RNA-seq data and the qRT-PCR results. In addition, six transcription factors, including *BraA03gNAC41*, *BraA03gSPL3*, *BraA05gRAP2.2*, *BraA06gNF-YC4*, *BraA07gSCL1*, and *BraA08gSOP1* were selected randomly from the candidate genes for the regulation of glucosinolate and soluble sugar biosynthesis, and their expression was analyzed using qRT-PCR to validate the transcriptome data sets from RNA-Seq. The results showed that the relative expression trends of these transcription factors are in close agreement with the corresponding relative transcript abundances, validating the reproducibility and credibility of the transcriptome data in this study.

Expression analysis of genes related to the biosynthesis of soluble sugar and glucosinolate

The seedling stage of yellow inner-leaf (XQ) and white inner-leaf (XYW) Chinese cabbage were used to verify the above results and confirm the molecular mechanism of Chinese cabbage soluble sugar and glucosinolate biosynthesis. We measured the total soluble sugar content and total glucosinolate content of the seeding stage in 'XQ' and 'XYW' (Figures 7A, B). The results showed a high glucosinolate in yellow inner-leaf Chinese cabbage. And, a significant high soluble sugar content is observed in white inner-leaf Chinese cabbage, which is consistent to with their content at the heading stage.

The expression patterns of all structural genes and ten transcription factors regulating the biosynthesis of glucosinolate and soluble sugar were analyzed using the qRT-PCR in the seedlings of 'XQ' and 'XYW' (Figures 7C, D). The relative expression of four structural genes regulating glucosinolate biosynthesis was significantly higher in 'XQ' than in 'XYW'. *BraA05gAOP1*, *BraA02gSOT12*, and *BraA04gAOP4* as the positive regulators of glucosinolate biosynthesis. It agrees with the glucosinolate content (Figure 7A) and the RNA-Seq data (Figure 7C). While the expression of the negative regulator of *BraA06gBCAT2* is inconsistent with the glucosinolate content.

Meanwhile, the relative expression level of two structural genes, *BraA01gHT4* and *BraA03gHT7*, confirmed the soluble sugar content, with the significantly high expression level in white inner-leaf Chinese cabbage. The verification of ten transcription factors correlated to *BraA01gHT4* and *BraA03gHT7* showed that *BraA03gSPL3* and *BraA04gFRS5* in the blue module are inconsistent RNA-Seq results with higher expression in white inner-leaf Chinese cabbage in RNA-Seq data but lower expression at seedling stage of white inner-leaf Chinese cabbage (Figure 7D). It agrees with the soluble sugar content at the heading and seedling stages. In the green module, the expression of *BraA03gNAC41*, *BraA07gGPRI1*, and *BraA07gbHLH77* were high in white inner-leaf Chinese

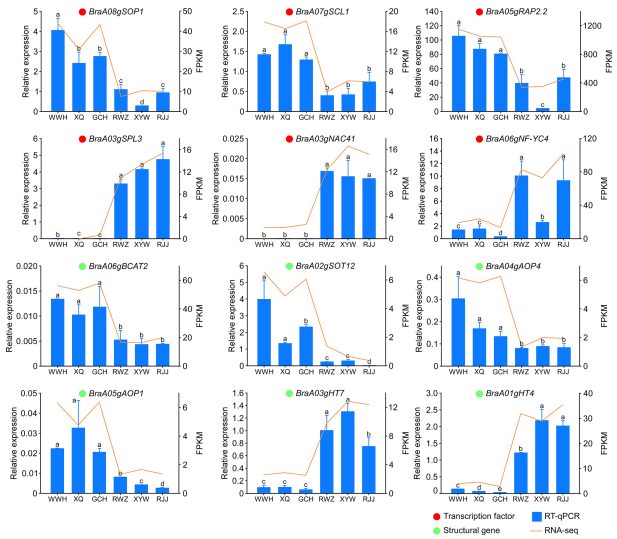


FIGURE 6
qRT- PCR verification of expression levels of key genes identified by RNA sequencing. The relative expression levels analyzed by qRT-PCR and calculated by $2^{-\triangle Ct}$ and bars with different lowercase letters are significantly different ($P \leq 0.05$). The red circle represents transcription factors, and the green circle represents structural genes.

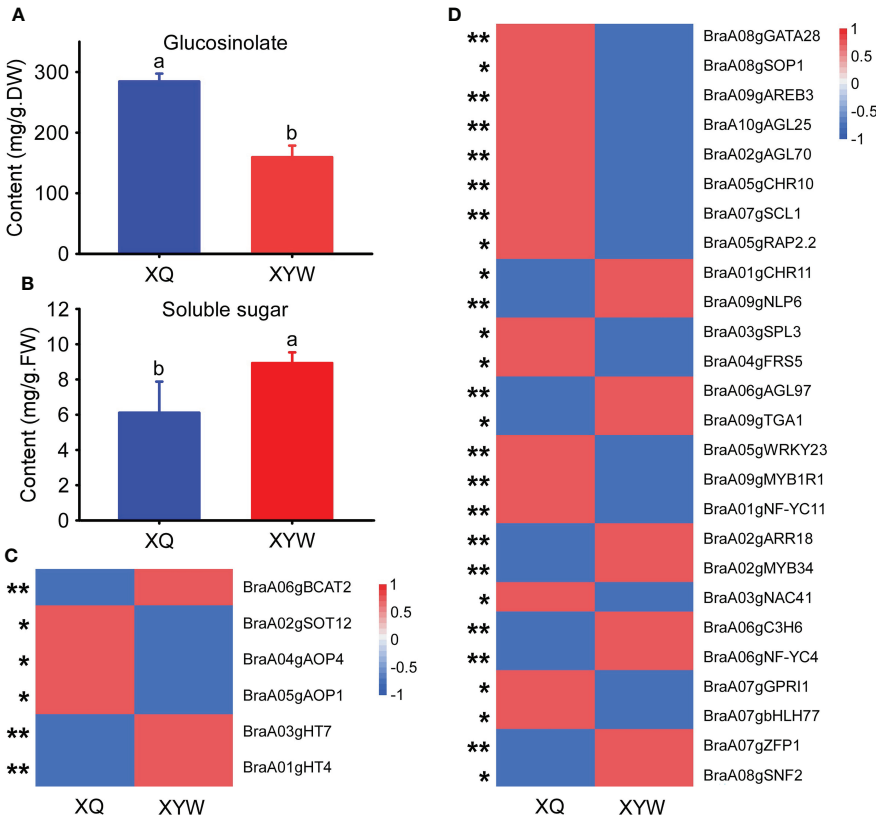


FIGURE 7
Expression analysis of key genes related to soluble sugar and glucosinolate biosynthesis in yellow and white inner-leaf Chinese cabbage. **(A)** Glucosinolate content at the seedling stage. **(B)** Soluble sugar content at the seedling stage. Data are means ± SD, N = 3. Lowercase letters indicate that there are significant differences between 'XQ' and 'XYW' ($P \leq 0.05$). **(C)** Relative expression of structural genes involved in glucosinolate and soluble sugar biosynthesis at the seedling stage. **(D)** The expression of transcription factors involved in glucosinolate and soluble sugar biosynthesis at the seedling stage. Asterisks indicate statistical significance using Student's t-test: *,$P \leq 0.05$, **,$P \leq 0.01$.

cabbage. It is contrary to the soluble sugar content and RNA-Seq at the seedling stage.

In addition, the four common transcription factors regulating glucosinolate and soluble sugar biosynthesis, *BraA07gSCL1*, *BraA05gRAP2.2*, *BraA01gCHR11*, *BraA09gNLP6* exhibited a consistent result to RNA-Seq and glucosinolate content. However, the RNA-Seq results and the expression pattern of *BraA07gSCL1* and *BraA05gRAP2.2* showed a contrary trend to the soluble sugar content.

Discussion

Quantitative and qualitative analysis of soluble sugar and glucosinolate in Chinese cabbage

Flavor and leaf color are two essential quality indices of Chinese cabbage (Huang et al., 2016). The content of glucosinolate and soluble sugar determines the flavor of Chinese cabbage. In a previous study, the total glucosinolate content in Chinese cabbage was much lower than that in broccoli, kale and other cruciferous vegetables (Bhandari et al., 2015). The metabolic pathways and regulators governing the biosynthesis of soluble sugar and glucosinolate were elucidated to improve the quality of Chinese cabbage by increasing the soluble sugar and glucosinolate content. After pairwise comparison, a higher content of carotenoid and glucosinolate was found in yellow inner-leaf Chinese cabbage (Figure 1B). The qualitative and quantitative analysis showed that aliphatic glucosinolate was the primary type of glucosinolate in Chinese cabbage (Supplementary Material 1). It was consistent with the report that the content of aliphatic glucosinolate was higher than other kinds of glucosinolate and positively correlated with the carotenoid in Chinese cabbage (Baek et al., 2016). The content and type of soluble sugar are essential factors that affect the flavor and sweetness of Chinese cabbage. In our study, the soluble sugar content in white inner-leaf Chinese cabbage was

generally higher than that of yellow inner-leaf Chinese cabbage. D-fructose and glucose were the main ingredients of soluble sugar in Chinese cabbage. Sucrose, fructose and glucose show significant differences in sweetness; fructose's sweetness is 1.73 times that of sucrose and 2.34 times that of glucose (Pangborn, 1963). This result signified that fructose and glucose could be mainly responsible for the high sweetness of white inner-leaf Chinese cabbage. Although inositol was detected in both yellow inner-leaf and white inner-leaf Chinese cabbage, the content in yellow inner-leaf Chinese cabbages had a much higher range than in white inner-leaf Chinese cabbage. Inositol is not only involved in plant sugar transport and resistance but also inhibits the growth of tumor cells while also having the same function as vitamin B1 (Loewus and Loewus, 1983). This also further explains the traits of the higher nutritional value of yellow inner-leaf Chinese cabbage.

Molecular mechanism of soluble sugar biosynthesis and accumulation in Chinese cabbage

Elucidating the metabolic pathways and regulators governing the biosynthesis and accumulation of soluble sugar is essential to provide new leads for improving Chinese cabbage quality. Fructose and glucose biosynthesis and metabolism are regulated by many essential structural genes, for instance, *SUS*, *INV*, *HxK*, *SPS*, and *FRK* (Rosa et al., 2009; Liu et al., 2020; Cheng et al., 2022). In our study, two critical structural genes encoding HT were identified by the conjugation of metabolomic and transcriptomic analyses. HT is a class of transporter proteins which directly regulates hexose content by transporting the hexose into a vacuole from the plasma membrane. As reported, the high HT activity transfers hexose into parenchyma cells against the concentration gradient and promotes sugar accumulation in fruit (Li et al., 2018). In apples with overexpression of *HT*, fructose and glucose content were significantly increased, while the sucrose content was decreased (Wang, 2020). And the *LeHT1* and *LeHT2* exhibited transport properties consistent with a high-affinity glucose/H^+ and low-affinity fructose/H^+ symporters (McCurdy et al., 2010) in tomatoes. Moreover, the decrease of LeHT1 caused a decline in hexoses concentrations, including glucose, fructose, and mannose (McCurdy et al., 2010). According to metabolic data analysis, fructose and glucose were the main component of soluble sugar in Chinese cabbage. Therefore, we suspected that HT probably is a key enzyme regulating the soluble sugar content in Chinese cabbage and also leads to the difference in soluble sugar between the yellow and the white inner-leaf Chinese cabbage.

Furthermore, transcription factors are the key factors in the biosynthesis of soluble sugar in Chinese cabbage. So far, *MYB*, *WARKY*, and *bZIP* have been confirmed to regulate soluble sugar biosynthesis. It has been reported that CmMYB113 could interact with CmSPS1 and CmACO1, regulating ethylene-dependent sucrose accumulation (Gao et al., 2021). *AtbZIP1* was vital in sugar-mediated gene expression and maintained sucrose homeostasis in plants (Wiese et al., 2005). Sucrose-specific signaling pathways showed to be responsible for the repression of ATB2/AtbZIP (Wiese et al., 2005). The expression of *WRKY42* is also closely related to the sucrose content in Pomegranate (Feng et al., 2022). In our study, we identified the critical transcription factors modulating the transcription of the vital structural genes, such as *WRKY*, *bZIP*, *GRAS*, *MYB*, *bHLH*, *RWP-RK*, *AP2/ERF-ERF*, and *B3*. Among them, transcription factors/transcription regulation factors *BraA07gSCL1* (GRAS), *BraA01gCHR11* (SNF2), *BraA05gRAP2.2* (AP2/ERF-ERF), and BraA09gNLP6 (RWP-RK) are involved in the regulation of both glucosinolate and soluble sugar biosynthesis. Whereas their expression levels of them at the seedling stage suggested that the *BraA09gNLP6* and *BraA01gCHR11* are not the critical transcription factors regulating soluble sugar biosynthesis in Chinese cabbage.

Molecular mechanism of glucosinolate in Chinese cabbage

In previous studies, the biosynthesis of aliphatic glucosinolate goes through the chain elongation of selected precursor amino acids, the formation of core glucosinolates structure and the secondary modification of the amino acid side chains in three steps (Sønderby et al., 2010). The vital roles of *BCAT*, *MAM*, *CYP79/CYP83*, *AOP*, *SOT*, and *UGT74* in this process have been reported (Sanchez-Pujante et al., 2017). In previous study, the *AOP* is probably the principal regulator of aliphatic glucosinolate accumulation in the biosynthetic pathway (Kliebenstein et al., 2001; Wang et al., 2021), consistent with our results. BCAT, as an aminotransferase enzyme, can act on branched-chain amino acids to regulate sulfur amino acid metabolism (Klein et al., 2006). The function of *AtSOT16*, *AtSOT17* and *AtSOT18* sulfonated desulphoglucosinolates in the last step of glucosinolate biosynthesis also were confirmed (Hashiguchi et al., 2014). In addition, *AtSOT12* was shown to be sulphurated brassinosteroids and flavanone (Marsolais et al., 2006; Gigolashvili et al., 2007). Even though these structural genes of glucosinolate biosynthesis have been studied a lot, previous studies mainly focused on Arabidopsis, and the relevant studies on Chinese cabbage are still lacking.

In our study, at the heading stage, the expression level of structural genes, *BraA06gBCAT2*, *BraA02gSOT12*, *BraA05gAOP1*, and *BraA04gAOP4* was significantly higher in the yellow inner-leaf Chinese cabbage than that in the white inner-leaf Chinese cabbage (Figure 7). Nevertheless, the expression of four structural genes was significantly higher. Therefore, considering the higher expression of *BraA06gBCAT2* in white inner-leaf Chinese cabbage at the seedling stage, we proposed that *BraA06gBCAT2* might not be an essential gene regulating glucosinolates accumulation at the seedling stage.

In past reports, MYB28, MYB29, and MYB76 participated in aliphatic glucosinolate biosynthesis by regulating *MAM1*, *MAM3*, *CYP79F1*, *CYP79F2*, *CYP83A1*, *AtST5b* and *AtST5c* (Zang et al., 2008). The gain of *MYB28* function increased the content of glucosinolate while decreasing in HAG1 RNAi knock-down mutants (Schön et al., 2013). Previous studies proved that bHLH could form protein complexes with MYB and allow the coordinated and rapid regulation of glucosinolate genes in cruciferous plants (Mitreiter and Gigolashvili, 2021). In addition, WRKY18 and WRKY40 are synergistic with CYP81F2 to negatively regulate the biosynthesis of indole glucosinolate (Schweizer et al., 2013). The crucial role of the MYC-MYB interaction system has been reported in the regulation of glucosinolate biosynthesis (Zang et al., 2021). To validate the previous studies, our report can also be used to verify the regulation of transcription factors for glucosinolate accumulation in Chinese cabbage. Our dataset showed that *bZIP* (Wiese et al., 2005), *SNF2*, *GRAS*, *MYB* (Zang et al., 2021), *bHLH* (Mitreiter and Gigolashvili, 2021) et al. play an essential role in regulating structural genes of glucosinolate biosynthesis. Of all the ten transcription factors, *BraA07gSCL1* belonged to the *GRAS* transcription factor family and was noted for its strong relationship with all four structural genes. The critical role of *SCL* on plant growth and development, signal transduction, disease resistance and stress resistance has been reported (Chen et al., 2020; Wang et al., 2020; Zang et al., 2021), while a lack of direct evidence on glucosinolate accumulation require the further research and verification.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA867427.

Author contributions

FW and JG supervised and conceived this project. LW wrote the paper. SZ and FW revised the paper. LW and SZ performed the transcript analysis and metabolic profiling. LW carried out qRT-PCR. JL, YZ, CL, LH, and HL performed sampling and preprocessing for transcriptome and metabolism analyses.

Funding

This research was funded by the Taishan Scholars Program of Shandong Province, China (tsqn201909167); the National Natural Science Foundation, China (32172591); the Modern Agricultural Industrial Technology System Funding of Shandong Province, China (SDAIT-05); the Agricultural Science and Technology Innovation Project of SAAS, China (CXGC2022D01); the Key R&D Program of Shandong Province, China (2019GHZ014); the Shandong Upgraded Project of "Bohai Granary" Science and Technology Demonstration Engineering (2019BHLC005); the China Agriculture Research System (CARS-23-G13); the Agricultural Science and Technology Innovation Project of SAAS (CXGC2022E08); the Jinan Agricultural Application Technology Innovation Program (202009).

Acknowledgments

Great appreciation is given to the editor and reviewers' critical comments on the improvement of the 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.2022.1043489/full#supplementary-material

References

Baek, S. A., Jung, Y. H., Lim, S. H., Park, S. U., and Kim, J. K. (2016). Metabolic profiling in Chinese cabbage (Brassica rapa l. subsp. pekinensis) cultivars reveals that glucosinolate content is correlated with carotenoid content. *J. Agric. Food Chem.* 64, 4426–4434. doi: 10.1021/acs.jafc.6b01323

Bhandari, S. R., Jo, J. S., and Lee, J. G. (2015). Comparison of glucosinolate profiles in different tissues of nine brassica crops. *Molecules* 20, 15827–15841. doi: 10.3390/molecules200915827

Bischoff, K. L. (2021). Chapter 53 -Glucosinolates. In R. C. Gupta*Nutraceuticals* (Boston: Academic Press), 551–4.

Cao, J. K., Jiang, W. B., and Zhao, Y. M. (2007). *Physiological and biochemical guidance of fruits and vegetables after harvest*. (Beijing: China Light Industry Press).

Cao, W., Wang, P., Yang, L., Fang, Z., Zhang, Y., Zhuang, M., et al. (2021). Carotenoid biosynthetic genes in cabbage: Genome-wide identification, evolution, and expression analysis. *Genes (Basel)* 12 (12), 2027–40. doi: 10.3390/genes12122027

Chelimoge,, Yuan, Q., Yang, R. J., Cao, X. L., and Zhu, S. H. (2014). 'Screening of conditions for determination of glucosinolates in rapeseed by PdCl colorimetry. *Heilongjiang Anim. Sci. Vet. Med.* 11, 216–217. doi: 10.13881/j.cnki.hljxmsy.2014.0450

Chen, Q. J., Deng, B. H., Gao, J., Zhao, Z. Y., Chen, Z. L., Song, S. R., et al. (2020). A miRNA-encoded small peptide, vvi-miPEP171d1, regulates adventitious root formation. *Plant Physiol.* 183, 656–670. doi: 10.1104/pp.20.00197

Cheng, H., Kong, W., Tang, T., Ren, K., Zhang, K., Wei, H., et al. (2022). Identification of key gene networks controlling soluble sugar and organic acid metabolism during oriental melon fruit development by integrated analysis of metabolic and transcriptomic analyses. *Front. Plant Sci.* 13, 830517. doi: 10.3389/fpls.2022.830517

Chhajed, S., Mostafa, I., He, Y., Abou-Hashem, M., El-Domiaty, M. , and Chen, S. (2020). Glucosinolate biosynthesis and the glucosinolate–myrosinase system in plant defense. *Agronomy* 10, 1786–1810. doi: 10.3390/agronomy10111786

Das, B. (2021). Glucosinolate biosynthesis: role of MAM synthase and its perspectives. *Biosci. Rep.* 41, BSR20211634. doi: 10.1042/BSR20211634

Durán-Soria, S., Pott, D. M., Osorio, S., and Vallarino, J. G. (2020). Sugar signaling during fruit ripening. *Front. Plant Sci.* 11, 564917. doi: 10.3389/fpls.2020.564917

Feng, L., Wang, C., Yang, X., Jiao, Q., and Yin., Y. (2022). Transcriptomics and metabolomics analyses identified key genes associated with sugar and acid metabolism in sweet and sour pomegranate cultivars during the developmental period. *Plant Physiol. Biochem.* 181, 12–22. doi: 10.1016/j.plaphy.2022.04.007

Gao, G., Duan, X., Jiang, H., Yang, F., and Qi., H. (2021). CmMYB113 regulates ethylene-dependent sucrose accumulation in postharvest climacteric melon fruit. *Postharvest Biol. Technol.* 181, 111682. doi: 10.1016/j.postharvbio.2021.111682

Gigolashvili, T., Yatusevich, R., Berger, B., Muller, C., and Flugge, U. I. (2007). The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in arabidopsis thaliana. *Plant J.* 51, 247–261. doi: 10.1111/j.1365-313X.2007.03133.x

Guo, R., Huang, Z., Deng, Y., Chen, X., XuHan, X., and Lai., Z. (2016). Comparative transcriptome analyses reveal a special glucosinolate metabolism mechanism in brassica alboglabra sprouts. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01497

Harun, S., Abdullah-Zawawi, M. R., Goh, H. H., and Mohamed-Hussein, Z. A. (2020). A comprehensive gene inventory for glucosinolate biosynthetic pathway in arabidopsis thaliana. *J. Agric. Food Chem.* 68 (28), 7281–7297. doi: 10.1021/acs.jafc.0c01916

Hashiguchi, T., Sakakibara, Y., Shimohira, T., Kurogi, K., Yamasaki, M., Nishiyama, K., et al. (2014). Identification of a novel flavonoid glycoside sulfotransferase in arabidopsis thaliana. *J. Biochem.* 155, 91–97. doi: 10.1093/jb/mvt102

Higdon, J. V., Delage, B., Williams, D. E., and Dashwood, R. H. (2007). Cruciferous vegetables and human cancer risk: Epidemiologic evidence and mechanistic basis. *Pharmacol. Res.* 55, 224–236. doi: 10.1016/j.phrs.2007.01.009

Huang, S., Liu, Z., Li, D., Yao, R., Hou, L., Li, X., et al. (2016). Physiological characterization and comparative transcriptome analysis of a slow-growing reduced-thylakoid mutant of Chinese cabbage (Brassica campestris ssp. pekinensis). *Front. Plant Sci.* 7, 3. doi: 10.3389/fpls.2016.00003

Hu, J. H., Chen, B. Y., Zhao, J., Zhang, F., Xie, T., Xu, K., et al. (2022). Genomic selection and genetic architecture of agronomic traits during modern rapeseed breeding. *Nat. Genet.* 54, 694–704. doi: 10.1038/s41588-022-01055-6

Hu, L., Wu, G., Hao, C., Yu, H., and Tan, L. (2016). Transcriptome and selected metabolite analyses reveal points of sugar metabolism in jackfruit (Artocarpus heterophyllus lam.). *Plant Sci.* 248, 45–56. doi: 10.1016/j.plantsci.2016.04.009

Ishida, M., Hara, M., Fukino, N., Kakizaki, T., and Morimitsu, Y. (2014). Glucosinolate metabolism, functionality and breeding for the improvement of brassicaceae vegetables. *Breed Sci.* 64, 48–59. doi: 10.1270/jsbbs.64.48

Jo, J. S., Bhandari, S. R., Kang, G. H., Shin, Y. K., and Lee., J. G. (2022). Selection of broccoli (Brassica oleracea var. italica) on composition and content of glucosinolates and hydrolysates. *Sci. Hortic.* 298, 110984. doi: 10.1016/j.scienta.2022.110984

Kang, K. B., Jayakodi, M., Lee, Y. S., Nguyen, V. B., Park, H. S., Koo, H. J., et al. (2018). Identification of candidate UDP-glycosyltransferases involved in protopanaxadiol-type ginsenoside biosynthesis in panax ginseng. *Sci. Rep.* 8, 11744. doi: 10.1038/s41598-018-30262-7

Klein, M., and Papenbrock, J. (2008). "Sulfotransferases and their role in glucosinolate biosynthesis," in *Sulfur assimilation and abiotic stress in plants*. Eds. N. A. Khan, S. Singh and S. Umar (Berlin, Heidelberg: Springer Berlin Heidelberg).

Klein, M., Reichelt, M., Gershenzon, J., and Papenbrock., J. (2006). The three desulfoglucosinolate sulfotransferase proteins in arabidopsis have different substrate specificities and are differentially expressed. *FEBS J.* 273, 122–136. doi: 10.1111/j.1742-4658.2005.05048.x

Kliebenstein, D. J., Gershenzon, J., and Mitchell-Olds, T. (2001). Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in arabidopsis thaliana leaves and seeds. *Genetics* 159, 359–370. doi: 10.1093/genetics/159.1.359

Ku, K. M., Becker, T. M., and Juvik, J. A. (2016). Transcriptome and metabolome analyses of glucosinolates in two broccoli cultivars following jasmonate treatment for the induction of glucosinolate defense to trichoplusia ni (Hübner). *Int. J. Mol. Sci.* 17 (7), 1135–52. doi: 10.3390/ijms17071135

Liao, Y. C. (2011). *'Analysis of Structure and Content of Glucosinolates and QTL Location in Brassica RaPa* (Chongqing: Southwese University).

Li, Y., Fan, Y., Jiao, Y., Wu, J., Zhang, Z., Yu, X., et al. (2019). Transcriptome profiling of yellow leafy head development during the heading stage in Chinese cabbage (Brassica rapa subsp. pekinensis). *Physiol. Plant.* 165, 800–813. doi: 10.1111/ppl.12784

Li, M., Li, P., Ma, F., Dandekar, A. M., and Cheng, L. (2018). Sugar metabolism and accumulation in the fruit of transgenic apple trees with decreased sorbitol synthesis. *Hortic. Res.* 5, 60. doi: 10.1038/s41438-018-0064-8

Liu, Q., Li, J., and Liu., W. (2020). Sugar accumulation and characterization of metabolizing enzyme genes in leafy head of Chinese cabbage (Brassica campestris l. ssp. pekinensis). *Hortic. Environ. Biotechnol.* 62, 17–29. doi: 10.1007/s13580-020-00294-y

Loewus, F. A., and Loewus, M. W. (1983). Myo-inositol: Its biosynthesis and metabolism. *Annual Review of Plant Physiology* 34, 137–161.

Marsolais, F., Boyd, J., Paredes, Y., Schinas, A. M., Garcia, M., Elzein, S., et al. (2006). Molecular and biochemical characterization of two brassinosteroid sulfotransferases from arabidopsis, AtST4a (At2g14920) and AtST1 (At2g03760). *Planta* 225 (5), 1233–1244. doi: 10.1007/s00425-006-0413-y

Matsukura, C. (2016). "Sugar accumulation in tomato fruit and its modification using molecular breeding techniques," in *Functional genomics and biotechnology in solanaceae and cucurbitaceae crops*. Eds. H. Ezura, T. Ariizumi, J. Garcia-Mas and J. Rose (Berlin, Heidelberg: Springer Berlin Heidelberg).

McCurdy, D. W., Dibley, S., Cahyanegara, R., Martin, A., and Patrick, J. W. (2010). Functional characterization and RNAi-mediated suppression reveals roles for hexose transporters in sugar accumulation by tomato fruit. *Mol. Plant* 3, 1049–1063. doi: 10.1093/mp/ssq050

Mitreiter, S., and Gigolashvili, T. (2021). 'Regulation of glucosinolate biosynthesis. *J. Exp. Bot.* 72, 70–91. doi: 10.1093/jxb/eraa479

Moscatello, S., Famiani, F., Proietti, S., Farinelli, D., and Battistelli., A. (2011). Sucrose synthase dominates carbohydrate metabolism and relative growth rate in growing kiwifruit (Actinidia deliciosa, cv Hayward). *Sci. Hortic.* 128, 197–205. doi: 10.1016/j.scienta.2011.01.013

Nookaraju, A., Upadhyaya, C., Pandey, S., Young, K., Hong, S., Park, S., et al. (2010). Molecular approaches for enhancing sweetness in fruits and vegetables. *Sci. Hortic.* 127, 1–15. doi: 10.1016/j.scienta.2010.09.014

Pangborn, R. M. (1963). Relative taste intensities of selected sugars and organic acids. *J. Food Sci.* 28, 726–733. doi: 10.1111/j.1365-2621.1963.tb01680.x

Rosa, M., Prado, C., Podazza, G., Interdonato, R., Gonzalez, J. A., Hilal, M., et al. (2009). Soluble sugars–metabolism, sensing and abiotic stress: a complex network in the life of plants. *Plant Signaling Behav.* 4, 388–393. doi: 10.4161/psb.4.5.8294

Sønderby, I. E., Geu-Flores, F., and Halkier., B. A. (2010). Biosynthesis of glucosinolates – gene discovery and beyond. *Trends Plant Sci.* 15, 283–290. doi: 10.1016/j.tplants.2010.02.005

Sanchez-Pujante, P. J., Borja-Martinez, M., Pedreno, M. A., and Almagro, L. (2017). Biosynthesis and bioactivity of glucosinolates and their production in plant *in vitro* cultures. *Planta* 246, 19–32. doi: 10.1007/s00425-017-2705-9

Schön, M., Töller, A., Diezel, C., Roth, C., Westphal, L., Wiermer, M., et al. (2013). Analyses of wrky18 wrky40 plants reveal critical roles of SA/EDS1 signaling and indole-glucosinolate biosynthesis for golovinomyces orontii resistance and a loss-of resistance towards pseudomonas syringae pv. tomato AvrRPS4. *Mol. Plant Microbe Interact.* 26, 758–767. doi: 10.1094/MPMI-11-12-0265-R

Schweizer, F., Fernández-Calvo, P., Zander, M., Diez-Diaz, M., Fonseca, S., Glauser, G., et al. (2013). Arabidopsis basic helix-loop-helix transcription factors MYC2, MYC3, and MYC4 regulate glucosinolate biosynthesis, insect performance, and feeding behavior. *Plant Cell.* 25, 3117–3132. doi: 10.1105/tpc.113.115139

Shen, J., Zou, Z., Zhang, X., Zhou, L., Wang, Y., Fang, W., et al. (2018). Metabolic analyses reveal different mechanisms of leaf color change in two purple-leaf tea plant (Camellia sinensis l.) cultivars. *Hortic. Res.* 5, 7. doi: 10.1038/s41438-017-0010-1

Shim, J. Y., Kim, D. G., Park, J. T., Kandpal, L., Hong, S. J., Cho, B. K., et al. (2016). Physicochemical quality changes in Chinese cabbage with storage period and temperature: A review. *J. Biosyst. Eng.* 41, 373–388. doi: 10.5307/JBE.2016.41.4.373

Tandayu, E., Borpatragohain, P., Mauleon, R., and Kretzschmar, T. (2022). Genome-wide association reveals trait loci for seed glucosinolate accumulation in Indian mustard (Brassica juncea l.). *Plants (Basel)* 11 (3), 364–375. doi: 10.3390/plants11030364

Vallone, S., Sivertsen, H., Anthon, G. E., Barrett, D. M., Mitcham, E. J., Ebeler, S. E., et al. (2013). An integrated approach for flavour quality evaluation in muskmelon (Cucumis melo l. reticulatus group) during ripening. *Food Chem.* 139, 171–183. doi: 10.1016/j.foodchem.2012.12.042

Wang, Z. Y. (2020). *QTL Mapping of Fructose Content and Functional Analysis of Hexose Transporter MdHT2.2 in Apple* (Shanxi: Northwest Agriculture & Forestry University).

Wang, Y., Feng, C., Zhai, Z., Peng, X., Wang, Y., Sun, Y., et al. (2020). The apple microR171i-SCARECROW-LIKE PROTEINS26.1 module enhances drought stress tolerance by integrating ascorbic acid metabolism. *Plant Physiol.* 184, 194–211. doi: 10.1104/pp.20.00476

Wang, H., Wu, J., Sun, S. L., Liu, B., Cheng, F., et al. (2011). 'Glucosinolate biosynthetic genes in brassica rapa. *Gene* 487, 135–142. doi: 10.1016/j.gene.2011.07.021

Wang, K., Zhang, Z. Y., Tsai, H. I., Liu, Y. F., Gao, J., Wang, M., et al. (2021). Branched-chain amino acid aminotransferase 2 regulates ferroptotic cell death in cancer cells. *Cell Death Differ.* 28, 1222–1236. doi: 10.1038/s41418-020-00644-4

Wiese, A., Elzinga, N., Wobbes, B., and Smeekens, S. (2005). Sucrose-induced translational repression of plant bZIP-type transcription factors. *Biochem. Soc. Trans.* 33, 272–275. doi: 10.1042/BST0330272

Zang, Y. X., Kim, J. H., Park, Y. D., Kim, D. H., and Hong, S. B. (2008). Metabolic engineering of aliphatic glucosinolates in Chinese cabbage plants expressing arabidopsis MAM1, CYP79F1, and CYP83A1. *Bmb Rep.* 41, 472–478. doi: 10.5483/BMBRep.2008.41.6.472

Zang, Q. L., Zhang, Y., Han, S. Y., Li, W. F., and Qi, L. W. (2021). Transcriptional and post-transcriptional regulation of the miR171-LaSCL6 module during somatic embryogenesis in larix kaempferi. *Trees* 35, 145–154. doi: 10.1007/s00468-020-02026-2

Zuluaga, D. L., Graham, N. S., Klinder, A., van Ommen Kloeke, A. E. E., Marcotrigiano, A. R., Wagstaff, C., et al. (2019). Overexpression of the MYB29 transcription factor affects aliphatic glucosinolate synthesis in brassica oleracea. *Plant Mol. Biol.* 101, 65–79. doi: 10.1007/s11103-019-00890-2

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 December 2022
DOI 10.3389/fpls.2022.1033316

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Wei-Wei Deng,
Anhui Agricultural University, China
Yanxia Xu,
Zhejiang Agriculture and Forestry University, China
Xin-Qiang Zheng,
Zhejiang University, China
Chengwen Shen,
Hunan Agricultural University, China

*CORRESPONDENCE
Changsong Chen
ccs6536597@163.com
Liyuan Wang
wangly@tricaas.com

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 31 August 2022
ACCEPTED 29 November 2022
PUBLISHED 15 December 2022

CITATION
Zhang Y, Wei K, Guo L, Lei Y, Cheng H, Chen C and Wang L (2022) Functional identification of purine permeases reveals their roles in caffeine transport in tea plants (*Camellia sinensis*). *Front. Plant Sci.* 13:1033316. doi: 10.3389/fpls.2022.1033316

Functional identification of purine permeases reveals their roles in caffeine transport in tea plants (*Camellia sinensis*)

Yazhen Zhang[1,2], Kang Wei[1], Lingling Guo[1], Yuping Lei[1], Hao Cheng[1], Changsong Chen[2*] and Liyuan Wang[1*]

[1]Key Laboratory of Tea Biology and Resources Utilization, Ministry of Agriculture, National Center for Tea Improvement, Tea Research Institute Chinese Academy of Agricultural Sciences (TRICAAS), Hangzhou, China, [2]Tea Research Institute, Fujian Academy of Agricultural Sciences, Fuzhou, China

Caffeine is a characteristic secondary metabolite in tea plants. It confers tea beverage with unique flavor and excitation effect on human body. The pathway of caffeine biosynthesis has been generally established, but the mechanism of caffeine transport remains unclear. Here, eight members of purine permeases (PUPs) were identified in tea plants. They had diverse expression patterns in different tissues, suggesting their broad roles in caffeine metabolism. In this study, F1 strains of "Longjing43" ♂ × "Baihaozao" ♀ and different tea cultivars were used as materials to explore the correlation between caffeine content and gene expression. The heterologous expression systems of yeast and Arabidopsis were applied to explore the function of CsPUPs. Correlation analysis showed that the expressions of CsPUP1, CsPUP3.1, and CsPUP10.1 were significantly negatively correlated with caffeine content in tea leaves of eight strains and six cultivars. Furthermore, subcellular localization revealed that the three CsPUPs were not only located in plasma membrane but also widely distributed as circular organelles in cells. Functional complementation assays in yeast showed that the three CsPUPs could partly or completely rescue the defective function of *fcy2* mutant in caffeine transport. Among them, transgenic yeast of CsPUP10.1 exhibited the strongest transport capacity for caffeine. Consistent phenotypes and functions were further identified in the CsPUP10.1-over-expression Arabidopsis lines. Taken together, it suggested that CsPUPs were involved in caffeine transport in tea plants. Potential roles of CsPUPs in the intracellular transport of caffeine among different subcellular organelles were proposed. This study provides a theoretical basis for further research on the PUP genes and new insights for caffeine metabolism in tea plants.

KEYWORDS

caffeine, transport, purine permease, functional identification, tea plant

1 Introduction

Caffeine, namely, 1,3,7-trimethylxanthine, is an important functional metabolite in plants and has effects of excitement and diuresis on human body (Shilo et al., 2002; Celik et al., 2010; Gramza-Michalowska, 2014). It is a common component in tea (*Camellia sinensis*), coffee (*Coffea arabica* and *Coffea canephora*), and cocoa (*Theobroma cacao*), which were widely consumed as the three major soft drink in the world. However, the distribution of caffeine in tea plants is greatly distinct from coffee and cocoa. Caffeine in tea plants mainly occurred and accumulated in leaves, rather than in seeds and fruits. Moreover, its content in young tea leaves is 2%–3% of dry weight and even higher than that in coffee (Suzuki et al., 1992; Ashihara and Suzuki, 2004). The subcellular distribution of its biosynthesis is suggested to be localized in the chloroplasts, where *de novo* and salvage synthesis pathway occur and most key biosynthetic enzymes exist (Ashihara et al., 2013). Whereas, caffeine in coffee appears to be synthesized in the cytoplasm (Ogawa et al., 2001; Kumar et al., 2007). The vacuole was supposed to be the caffeine storage site, which was similar to other water-soluble secondary metabolites in plants (Wink, 2010). For now, the pathway of caffeine biosynthesis has been generally established (Ashihara et al., 2013; Jin et al., 2016). However, studies on its transport process have not been reported yet.

The mechanisms of alkaloids transport in some plants have been revealed (Shitan et al., 2014). Alkaloid transporters play an essential role in maintaining the dynamic equilibrium of alkaloids in different types of organelles, cells, and tissues. Purine permeases (PUPs) were recently reported to be involved in the transportation of purine, pyridine, tropane, and benzylisoquinoline alkaloids, as well as their derivatives and analogs (Jelesko et al., 2012; Dastmalchi et al., 2019).

PUPs are found specifically in vascular plants and assigned to an ancient drug and metabolite transporter super family (Jelesko, 2012). For the past decade, PUPs were primarily found to uptake purine nucleobase and their derivatives. The function was first demonstrated in Arabidopsis (Gillissen et al., 2000). In total, 21 PUP members were found in Arabidopsis. They were predicted to contain typically 9–10 transmembrane spanning domains (Qi and Xiong, 2013). AtPUP1 and AtPUP2 showed high sequence identity and similar substrate specificity. They were proved to be energy-dependent and H^+-coupled transporters of adenine and cytokinin. Their different tissue-specific expression patterns indicated distinctive transport mechanisms in Arabidopsis. An expanding role of AtPUP1 and AtPUP2 in caffeine transport in yeast was also proved. However, none of transport activity was detected in AtPUP3 yeast transformants (Gillissen et al., 2000; Bürkle et al., 2003). Further study on AtPUP1 revealed that it also showed high transport activity for pyridoxine (vitamin B6) (Szydlowski et al., 2013). AtPUP14 was localized on the plasma membrane and could import bioactive cytokinins to the cytosol (Zürcher et al., 2016). Twelve PUP members were found in rice (*Oryza sativa*). OsPUP7 was mainly expressed in the vascular bundle, pistil, and stamens. It was identified as caffeine transporter in yeast and played an important role in the cytokinin transportation in rice (Qi and Xiong, 2013). Further studies showed that OsPUP7 was localized on endoplasmic reticulum (ER) and positively regulated the grain size. The overexpression (OE) of OsPUP4, the closest homolog gene to OsPUP7, resulted in highly similar phenotypes. However, different from OsPUP7, OsPUP4 was localized on the plasma membrane. It was inferred that OsPUP4 and OsPUP7 function synergistically as cytokinin transporters in its long-distance movement and local allocation (Xiao et al., 2018). OsPUP1 was also localized on ER and could import cytokinins from the vascular tissues by cell-to-cell transport (Xiao et al., 2020). NtNUP1 (nicotine uptake permease), which belongs to PUPs family in tobacco, was demonstrated to transport various compounds in yeast system. Compared with pyridine alkaloids, tropane alkaloids, kinetin, and adenine, NtPUP1 showed preferential transport activities to nicotine. It revealed that NtPUP1 functioned as nicotine transporters in root tips from its synthesis sites (apoplastic space) into the cytoplasm (Hildreth et al., 2011; Kato et al., 2015). There are nine PUP members in opium poppy (*Papaver somniferum*), and six of them were identified as benzylisoquinoline alkaloid transporters. Each member showed a unique substrate preferential profile (Dastmalchi et al., 2019). Fifteen PUPs were found in *Coffea canephora*. CcPUP1 and CcPUP5 were demonstrated to facilitate the transport of adenine, but not caffeine in yeast uptake experiments (Kakegawa et al., 2019).

However, studies on PUP members or caffeine transporters in tea plants have not been reported. In this study, we aimed to (i) systematically investigate the genome-wide PUP genes in tea plants, including their bioinformatics analysis and expression patterns, (ii) explore the relationship between caffeine contents and CsPUP gene expression, and (iii) identify the function of CsPUPs in heterologous expression system, including yeast and Arabidopsis. This study could fill the research gap of caffeine transporters and provide a new insight into the regulation of secondary metabolism in tea plants.

2 Materials and methods

2.1 Plant materials

One-year-old cutting seedlings of "Zhongcha108" were used for the analysis of gene expression characteristics in different tissues, which was planted in the tea garden of the Tea Research Institute Chinese Academy of Agricultural Sciences (TRICAAS) in Hangzhou, China. Young leaves (YL) and mature leaves (ML) of different tea strains and cultivars were used for correlation analysis between caffeine content and gene expression. They

grew for 4–5 years in the tea garden of TRICAAS in Shengzhou, China. Details of tea plant materials were as follows: F1 strains of "Longjing43" ♂× "Baihaozao" ♀, numbered as 502, 915, 1417, 1916, 1501, 1511, 1608, and 1912, were collected on 24 July 2019. Tea cultivars, including "Longjing43" (LJ), "Zhongcha108" (ZC), "Zhongming7" (ZM), "Zhongbai4" (ZB), "Zhonghuang1" (ZH), and "Enshi4" (ESN), were collected on 17 April 2020.

2.2 Bioinformatics analysis of CsPUPs

CsPUP members were screened by text search "purine permease" on the "Shuchazao" (*Camellis sinensis* var. sinensis) genome database (http://tpia.teaplant.org/index.html). Then, the sequences of CsPUPs were submitted for BLASTX analysis with Arabidopsis in NCBI website (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Transmembrane regions of CsPUPs proteins were predicted by the online software TMHMM server v.2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/). Amino acid sequence alignment of CsPUPs was performed by DNAMAN. The phylogenetic tree was constructed by MEGA 6.0 using the neighbor-joining method. The genomic information of CsPUPs was screened by local blast software (ncbi-blast-2.9.0+); then, the genomic locations of CsPUPs were determined by MapChart.

2.3 Determination of caffeine content

The measurement of caffeine content by HPLC was described by Zhang et al. (2020). Collected materials were dried and grinded into powders. Then, 0.2 g of samples were extracted with 5 ml of 70% (v/v) methanol at 70°C water bath for 10 min and centrifuged at 3,500×g for 10 min. The extraction was repeated to reach a final volume of 10ml supernatant. Then, they were filtered through 0.45-μm Millipore filters and subjected for high performance liquid chromatography (HPLC) analysis. Reverse-phase column (Phenomenex C12, 4.6 × 250 mm, 5 μm) was used for the component assay. The mobile phases were as follows: 1% formic acid (solvent A), 100% acetonitrile (solvent B), and ultrapure water (solvent D). The linear elution gradient was 0–42 min (4%–18.7% B) and 42–43 min (18.7%–4% B). The samples were measured at 280 nm and eluted at 1 ml min^{-1}.

2.4 Gene expression analysis

First, total RNA extraction and cDNA synthesis were performed using the RNAprep Pure Plant Kit and FastQuant RT Kit (Tiangen, Beijing, China), respectively. Then, quantitative real-time (qRT)-PCR analysis was performed on a Roche LightCycler 480 II Real-Time PCR system using the PrimeScript RT reagent qPCR Kit (Takara, Dalian, China). All primers were designed by online software Primer-BLAST and displayed in Additional File 1. glyceraldehyde-3-phosphate dehydrogenase (CsGAPDH) was used as housekeeping gene. The relative gene expression level was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

2.5 Subcellular localization analysis of CsPUPs

The method of subcellular localization in tobacco epidermis was similar to that by Wang et al. (2022). The coding sequences (CDS) of CsPUP1, CsPUP3.1 and CsPUP10.1 without stop codon were amplified and cloned into the pBWA(V)HS-GFP vector, respectively. The specific primers were listed in Additional File 1. The plasmid of recombinant vector was transformed into *Agrobacterium* GV3101. The OsMCA1 was used as a plasma membrane (PM) marker for co-localization of target protein. 35S-GFP with empty vector (EV) was used as a positive control. Then, the transformant cells were injected into the lower epidermis of tobacco leaves and cultured for 48h under low light. Finally, the GFP signals were detected and recorded by confocal lase microscope (Nikon C2-ER).

2.6 Functional identification of CsPUPs in yeast

The homologous protein of PUP in yeast was encoded by FCY2 gene. It was first proved to transport purine-cytosine (Ferreira et al., 1997). The knock-out mutant *fcy2* was used as heterologous expression system for identifying transporters of nucleic acid bases and their derivatives in plants, such as cytokinin and caffeine. (Gillissen et al., 2000). It was also regarded as deficiency in caffeine transport and used for functional identification of PUPs in Arabidopsis and rice (Bürkle et al., 2003; Qi and Xiong, 2013). Therefore, to verify the function of CsPUPs, the gene CDS was cloned into the yeast-expression vector pYES2 by double enzymes digestion. Then, the plasmids with CsPUP1, CsPUP3.1, CsPUP10.1 and EV (control) were transformed into the yeast *fcy2* mutant (BY4741, *Mata, his3D1, leu2D0, met15D0, ura3D0*, YER056c::kanMX4). The normal WT yeast BY4741 was used as a positive control. The detailed method used the Quick and Easy Yeast Transformation Mix (Takara, Dalian, China). According to the information of pYES2 vector, the minimal medium lacking uracil was used as a basic medium. The medium containing 2% galactose and 1% raffinose as the carbon source was used as an inducing medium. The yeast cells were first activated in the basic medium plate for 2–3 days at 30°C. Then, single colony was pre-incubated overnight in the basic medium broth at 30°C with shaking at 200 rpm. The value of OD_{600} was measured and normalized to 0.5 by saline solution. Then, the suspension cells were diluted

tenfold successively. Finally, 50 µl of a hundredfold dilution (10^{-2}) was used for the yeast growth assay under 0.3% caffeine treatment (Qi and Xiong, 2013; Kakegawa et al., 2019).

For caffeine uptake assays in yeast, the method was slightly modified according to Morita et al. (2009) and Shoji et al. (2009). The OD_{600} value of pre-incubated yeast cells was normalized to 0.4 according to the pYES2 manual (www.invitrogen.com). Then, yeast cells were centrifuged at 4,000 rpm for 10 min and washed by sterile water for three times. Then, cells were incubated in 40 ml of inducing medium broth for 24 h and continuously cultured under 0.3% caffeine treatment for another 24 h. Next, yeast cells were harvested, and their fresh weight was recorded. Two milliliters of 70% (v/v) methanol was added for disrupting the reaction and caffeine extraction. The cells were then disrupted by ultrasonication under 400–500 W for 5 min. The determination of the caffeine content was performed according to the method abovementioned (see Section 2.3).

2.7 Functional identification of CsPUP10.1 in Arabidopsis

The recombinant plasmid pBWA(V)HS-GFP-35S-CsPUP10.1 was transformed into Arabidopsis (WT, *Col-0*) by inflorescence infection method. Then, the homozygous transgenic plants were screened by hygromycin (25 µg/ml) on 1/2 Murashige and Skoog (MS) medium for three generations (Zhang et al., 2021). For the phenotypic comparison between transformants and control, Arabidopsis seeds were first sowed on normal 1/2 MS medium plate and grew for 7 days. Then, seedlings with consistent growth conditions were selected for caffeine treatment. They were transplanted to 1/2 MS medium containing 0.01%, 0.03%, and 0.05% caffeine and cultivated vertically for 14 days. For caffeine uptake assays in Arabidopsis, seeds were sowed on 1/2 MS medium containing 0.02% caffeine and grew for 21 days. Then, the aerial parts of plants were chosen for caffeine measurement. The method used was the same as the abovementioned (see Section 2.3).

2.8 Statistical analysis

The abovementioned samples or treatments were performed for three biological replicates. Significance analysis was conducted by a one-way analysis of variance using the Statistical Package for the Social Sciences Statistics 17.0. Different lowercase letters represent significant differences at a P-value < 0.05. A bivariate correlation analysis between gene expression and caffeine content was conducted using Pearson's parametric correlation test. The figures were generated using the plotting software Origin 9.

3 Results

3.1 Identification and molecular characterization of CsPUPs in tea plants

On the basis of the tea genome database of "Shuchazao", eight CsPUP members were identified. The length of CDS was 723–1352 bp. Then, BLASTX analysis with Arabidopsis was conducted. The eight CsPUPs showed high sequence similarity (54%–73%) with five AtPUPs and were named as CsPUP1, CsPUP3.1, CsPUP3.2, CsPUP3.3, CsPUP4, CsPUP5, CsPUP10.1, and CsPUP10.2, respectively (Table 1). All of the CsPUP members have consensus conserved domains (purine nucleobase transmembrane, CL0184). The similarity of their amino acid sequences was 43.09%. Transmembrane structure analysis by TMHMM software showed that most of the CsPUPs contained 10 transmembrane domains (TMDs), except for CsPUP1 and CsPUP10.2. Further multiple alignment analysis showed that the sequence of TMDs was more conserved, especially for the TM4-TM7 (Figure 1).

To elucidate the phylogenetic relationships among PUPs in *Camellia sinensis*, *Arabidopsis thaliana*, *Oryza sativa*, *Nicotiana tabacum*, *Coffea canephora*, and *Papaver somniferum*, a phylogenetic tree was constructed by aligning eight CsPUPs, 20 AtPUPs, 13 OsPUPs, two NtPUPs, 15 CcPUPs, and nine

TABLE 1 Identification of PUPs in tea plants.

Name	Gene ID	CDS length	Homology	Similarity
CsPUP1	TEA003596	837	AtPUP1	59.46%
CsPUP3.1	TEA029223	1074	AtPUP3	63.07%
CsPUP3.2	TEA003576	1065	AtPUP3	59.45%
CsPUP3.3	TEA013066	1053	AtPUP3	56.67%
CsPUP4	TEA015426	1299	AtPUP4	64.49%
CsPUP5	TEA031626	1134	AtPUP5	73.16%
CsPUP10.1	TEA023430	1161	AtPUP10	53.93%
CsPUP10.2	TEA002721	813	AtPUP10	70.13%

http://tpia.teaplant.org/download.html

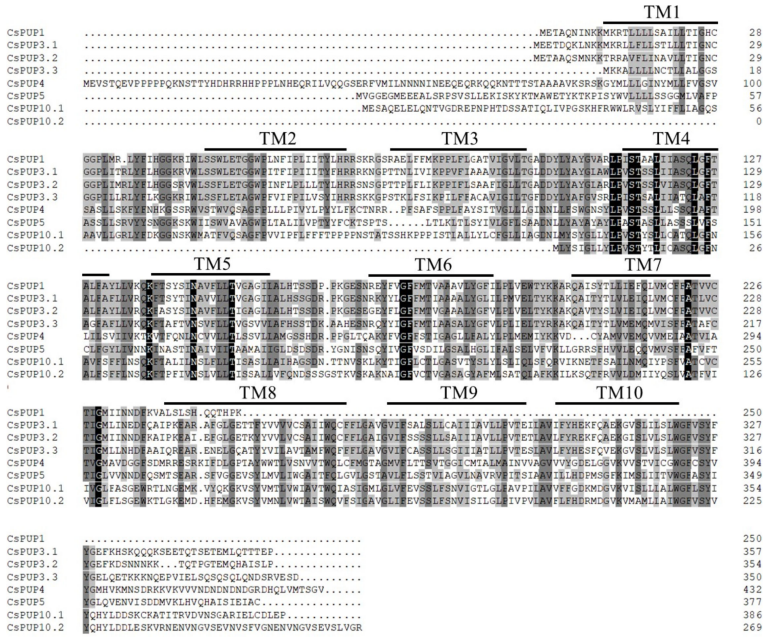


FIGURE 1
Amino acids sequences alignments of CsPUPs.

PsBUPs using MEGA6.0 (Figure 2). Phylogenetic analysis revealed that these 67 PUPs in the six species could be clustered into four subgroups. CsPUPs were distinctly distributed into three clades: CsPUP1, CsPUP3.1, CsPUP3.2, and CsPUP3.3 as clade I; CsPUP4 and CsPUP5 as clade II; and CsPUP10.1 and CsPUP10.2 as clade III.

Through the physical position identification, eight CsPUPs were unevenly distributed on six chromosomes (Figure 3). CsPUP1 and CsPUP3.2 were located on the chr1. CsPUP4 and CsPUP10.2 were located on the chr5. CsPUP3.1, CsPUP3.3, CsPUP5, and CsPUP10.1 were located on the chr10, chr6, chr8, and chr13, respectively.

3.2 Expression patterns of CsPUPs in different tissues

qRT-PCR was applied for the expression analysis of CsPUPs in different tissues, including young leaf (one bud with one leaf), mature leaf (the fifth leaf), stem, and root (Figure 4). CsPUP1 was mainly expressed in the mature leaf. CsPUP3.1, CsPUP3.3, CsPUP10.1, and CsPUP10.2 were expressed in the all tissues tested. CsPUP3.1 and CsPUP3.3 showed similar pattern and were primarily expressed in the aerial parts of tea plants. CsPUP10.1 and CsPUP10.2 were highly expressed in the mature leaf. Interestingly, CsPUP3.2 and CsPUP4 were mainly expressed in the stem, lower or no detectable expression in leaves. CsPUP5 was failed to be amplified in this study. The results revealed that these CsPUPs have tissue-specific or preferential expression patterns. In general, CsPUP1 showed the highest expression level in the aerial parts, followed by CsPUP3.1, CsPUP3.3, and CsPUP10.1 (Additional File 2). In addition, CsPUP3.1 and CsPUP3.3 were clustered into clade I in phylogenetic tree (Figure 2) and showed similar expression model in different tissues (Figure 4). Considering that caffeine is mainly distributed in leaves, CsPUP1, CsPUP3.1, and CsPUP10.1 were selected for further investigation in this study.

3.3 Caffeine contents in different tea strains and cultivars

Previous study proved that caffeine accumulation in tea leaves with different maturities varied greatly (Mohanpuria et al., 2009a). Gene expression levels of CsPUPs also showed obvious differences in YL and ML. Therefore, YL and ML were used for determining the correlation between gene expression and caffeine content. To investigate the caffeine accumulation in closely related and unrelated materials, eight strains of F1 population of plant hybrids from tea cultivars of “Longjing43” ♂× “Baihaozao” ♀, and six different cultivars were used in this study. The levels of caffeine ranged from 11.81 to 31.80 mg/g in the strains and from 0.92 to 43.18 mg/g in the cultivars, respectively (Figure 5). Similar caffeine accumulation patterns were found in different tea strains and cultivars. Caffeine content in YL was significantly higher than that in ML, which was consistent with the previous reports

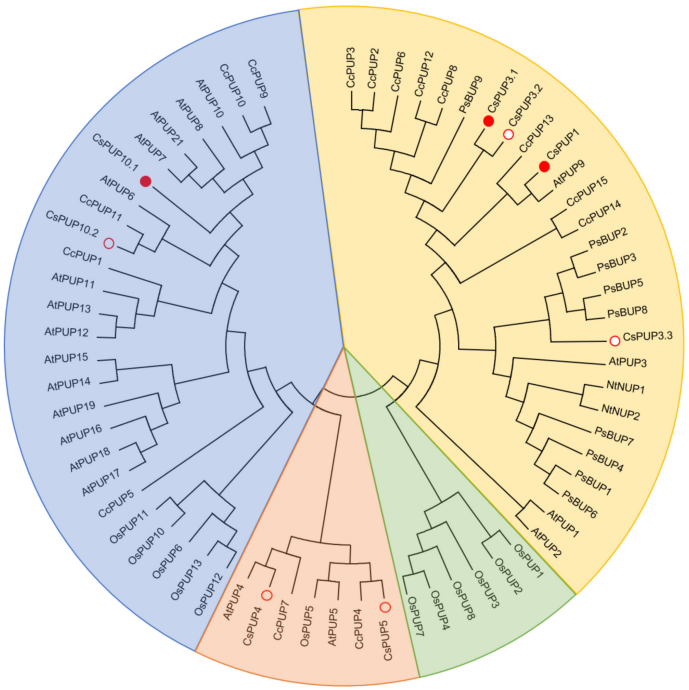


FIGURE 2
Phylogenetic tree analysis of PUP family members.

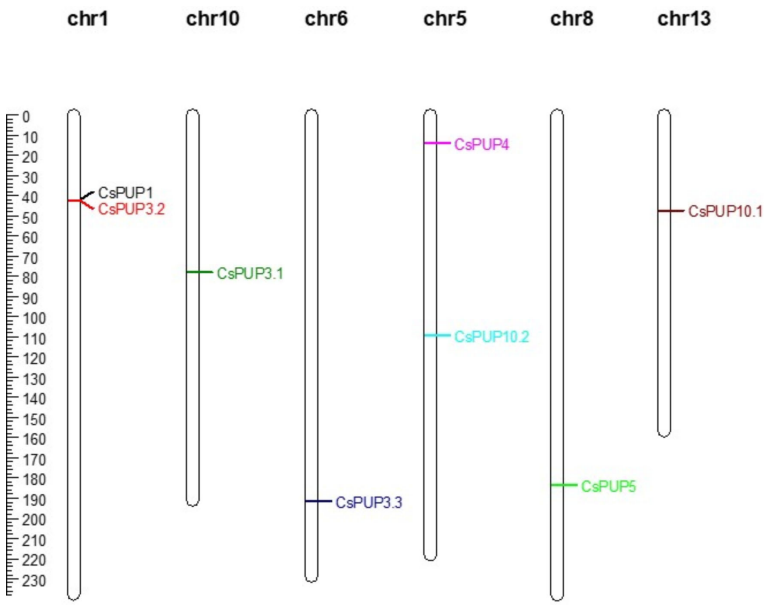
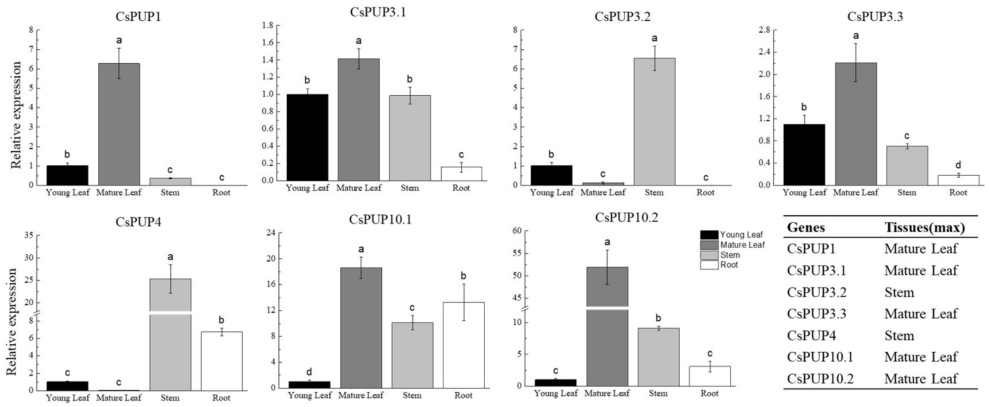


FIGURE 3
Distribution of the CsPUPs genes on chromosomes.



Genes	Tissues(max)
CsPUP1	Mature Leaf
CsPUP3.1	Mature Leaf
CsPUP3.2	Stem
CsPUP3.3	Mature Leaf
CsPUP4	Stem
CsPUP10.1	Mature Leaf
CsPUP10.2	Mature Leaf

FIGURE 4
Expression patterns of CsPUPs in different tissues.

(Mohanpuria et al., 2009a). In addition, compared with the eight strains, greater variations of caffeine accumulation between YL and ML were observed in the six cultivars, especially for ZB (22.56 folds).

3.4 CsPUP gene expression and correlation with caffeine contents

The expression levels of CsPUP1, CsPUP3.1, and CsPUP10.1 in the abovementioned materials were analyzed by qRT-PCR (Figure 6). Consistent expression patterns were identified among the three genes. Overall, CsPUPs were expressed higher in ML than that in YL, regardless of the tea strains or cultivars. Among of them, CsPUP1 showed greater variations, 4- 25-fold changes were found in ML and YL.

To further explore the relationship between CsPUPs and caffeine content, correlation analysis was performed. The results revealed that the three gene transcription levels were significantly negatively correlated with caffeine concentration (Table 2). CsPUP10.1 showed the highest correlation coefficient with caffeine in different cultivars (−0.834**). Whereas in the eight strains, CsPUP1 was found to have the closest correlation with caffeine (−0.710**). These results suggested that these CsPUPs might play important roles in caffeine metabolism.

3.5 Subcellular localization of CsPUPs

The function of these transporters strongly depends on their distributions in cells. To confirm subcellular localization of CsPUP1, CsPUP3.1, and CsPUP10.1, full length of each gene was translationally fused with GFP. As shown in Figure 7, CsPUP1, CsPUP3.1, and CsPUP10.1 showed similar distribution patterns in the epidermic cells. The green fluorescence signals could coincide with the red fluorescence probe of PM marker protein, indicating that these CsPUPs are PM-associated protein. In addition, strong signals emerged as

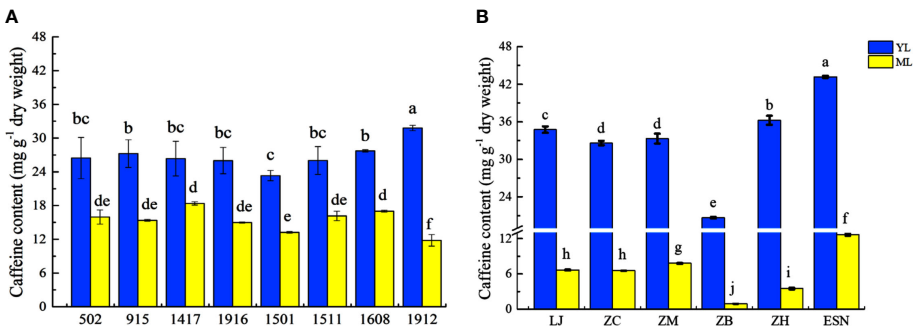


FIGURE 5
Caffeine content in different tea strains **(A)** and cultivars **(B)**.

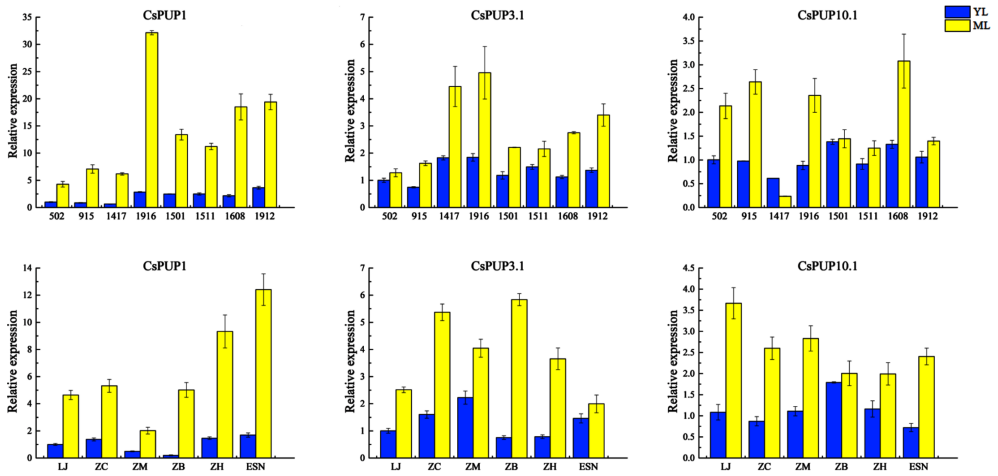


FIGURE 6
Expression levels of CsPUP1, CsPUP3.1 and CsPUP10.1 in different tea strains and cultivars.

solid or hollow green dots were also detected. The dots irregularly distributed throughout the cell (white arrows in Figure 7). They are presumed to be secreted vesicles, peroxisomes, or lysosomes, which may contribute to performing the transport function of CsPUPs.

3.6 Functional identification of CsPUPs in yeast

The yeast mutant *fcy2* was defective in caffeine transport (Gillissen et al., 2000) and usually used for the functional identification of alkaloid transporters. For species without any caffeine produced, such as Arabidopsis and rice, the OE of PUPs in the yeast mutant *fcy2* was reported to facilitate the transport of caffeine. It provided initial and indirectly experimental evidence for gene function in cytokinins or other derivatives transport (Bürkle et al., 2003; Qi and Xiong, 2013; Kakegawa et al., 2019). For tea plants with abundant caffeine accumulation, the *fcy2* mutant was ideal system for the functional identification of CsPUPs. To determine whether CsPUPs have similar function, CsPUP1, CsPUP3.1, and CsPUP10.1 were individually cloned into the pYES2 vector. Then, the recombinant plasmids were transformed into the yeast mutant *fcy2*. On the medium without caffeine treatment, there was no significant difference in yeast growth between EV and transgenic yeast. Whereas on the medium with 0.3% caffeine treatment, the *fcy2* yeast transformed with CsPUPs grew better than EV (Figure 8A). It suggested that CsPUP1, CsPUP3.1, and CsPUP10.1 mediated caffeine uptake into yeast cells.

To further assess their caffeine transport capacities, yeasts were incubated on inducing medium for 24 h and then treated with 0.3% caffeine for another 24 h. The contents of caffeine in these yeast cells were then determined by HPLC. As shown in Figure 8B, the caffeine accumulation in different yeast cells was ranked as follows: (BY4741 + EV) > CsPUP10.1 > CsPUP1 > CsPUP3.1 > (*fcy2* + EV). The uptake of caffeine was significantly higher in the transgenic cells than that in the negative control (*fcy2* + EV). Compared with *fcy2*, the caffeine transport capacity of CsPUP1, CsPUP3.1, and CsPUP10.1 transgenic cells increased for 2.7, 1.7, and 4.2 folds, respectively. In addition, caffeine content in normal wild-type (WT) yeast BY4741 was 4.6 folds higher than that in *fcy2* mutant, which confirmed the defective function of *fcy2* mutant. No significant difference was observed in caffeine uptake between BY4741 and CsPUP10.1. These results indicated that CsPUP1, CsPUP3.1, and CsPUP10.1 could partly or completely rescue the caffeine transport deficiency of *fcy2* mutant and function as caffeine influx transporters. Moreover, yeast transformed with CsPUP10.1 exhibited the strongest capacity of caffeine transport and was used for further investigation in this study.

TABLE 2 Correlation analysis of caffeine content and CsPUPs expression.

Correlation	Eight strains	Six cultivars
CsPUP1	−0.710**	−0.615*
CsPUP3.1	−0.593*	−0.776**
CsPUP10.1	−0.524*	−0.834**

"*" and "**" represent the significance level at $p < 0.05$ and $p < 0.01$, respectively.

3.7 Functional identification of CsPUP10.1 in Arabidopsis

To further clarify the function of CsPUP10.1 in caffeine transport and accumulation in plant, the gene was transformed

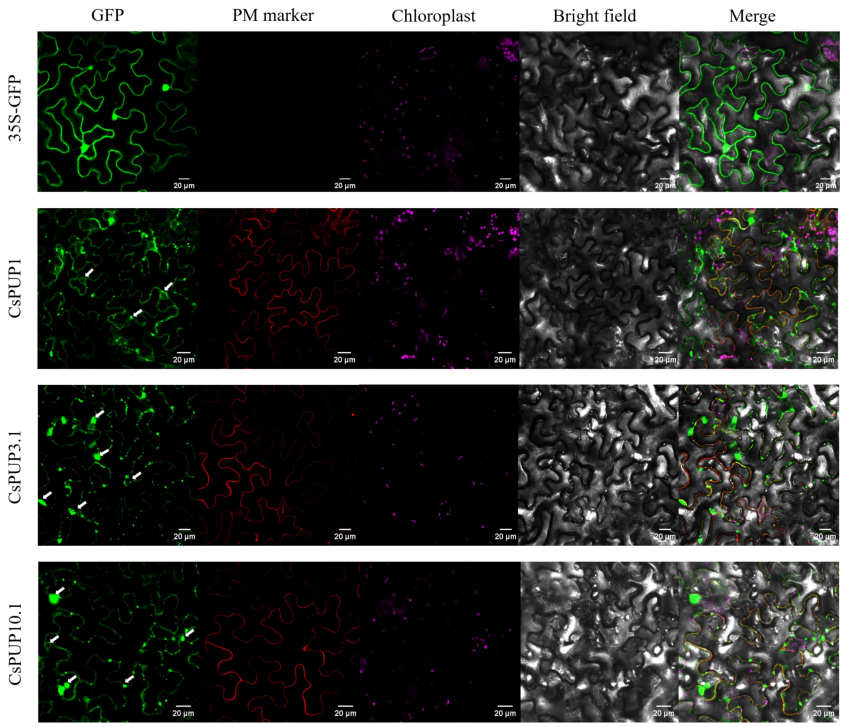


FIGURE 7
Subcellular localization of CsPUP1, CsPUP3.1 and CsPUP10.1.

into Arabidopsis for generating OE lines. Then, two transgenic lines (OE1 and OE2) and WT (control) were selected for exogenous caffeine treatment. On 1/2 MS medium without caffeine, no significant phenotypic difference was observed among the OE1, OE2, and WT (Additional File 3). Whereas on 1/2 MS plates containing caffeine, OE1 and OE2 grew better than WT, especially for the concentration of 0.03% and 0.05%. Transgenic plants showed longer taproots and more lateral roots. In addition, the toxic effect on Arabidopsis was observed in the presence of caffeine. All of the Arabidopsis lines were obviously inhibited under caffeine treatments, and the rosette leaves even showed visible symptoms of chlorosis in the plates containing 0.03% and 0.05% caffeine (Figure 9A).

To investigate the amounts of caffeine accumulated in the aerial parts of Arabidopsis, we cultivated plants on 1/2 MS medium with 0.02% caffeine for 21 days. Compared with WT, CsPUP10.1 transformants (OE1 and OE2) accumulated more caffeine (Figure 9B). The caffeine transport capacities in

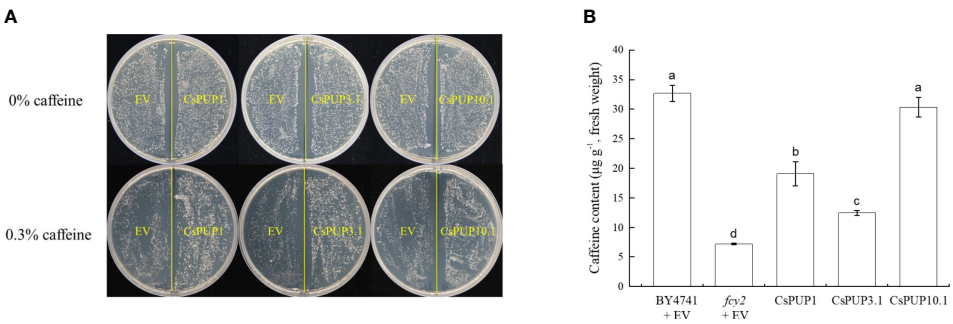


FIGURE 8
Functional identification of CsPUP1, CsPUP3.1 and CsPUP10.1 in yeast. **(A)** Phenotypic differences of yeast growth; **(B)** Caffeine accumulation in different yeast cells under 0.3% caffeine treatment.

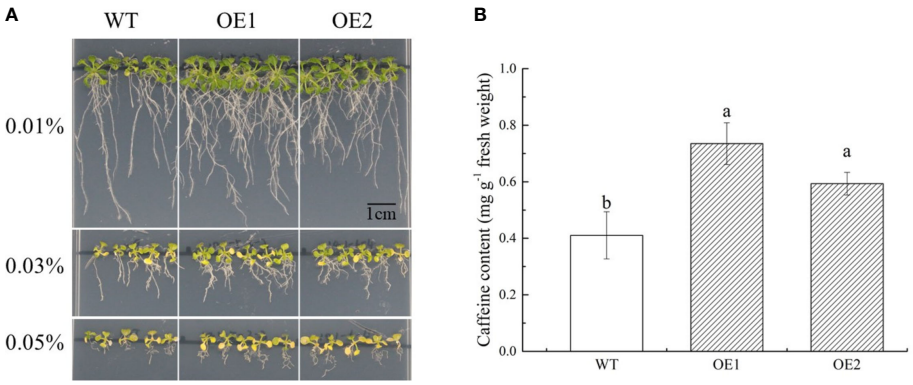


FIGURE 9
Functional identification of CsPUP10.1 in Arabidopsis. **(A)** Phenotypic differences of transgenic over-expression (OE) lines and wild type (WT) under caffeine treatments. **(B)** Caffeine accumulation in Arabidopsis under 0.01% caffeine treatment.

transgenic lines were improved for 79% and 45% in OE1 and OE2, respectively. These findings were consistent with the heterologous expression system in yeast and strongly demonstrated the function of CsPUP10.1 in caffeine transportation.

4 Discussion

PUPs are specifically found in vascular plants. Previous studies of PUPs in plants were mainly focused on their transport functions in adenine and cytokinin (Jelesko 2012) Because of their structural similarities to caffeine, a yeast mutant *fcy2* in caffeine transport deficiency was generally used as model cells for complementary assay. Generally, there is competitive inhibition between these purine-ring substrates. Therefore, more PUP members were proved to be cytokinin transporters in plants (Kang et al., 2017). AtPUP1 and AtPUP2 showed transport activity for adenine, cytokinin, and caffeine in *fcy2* yeast (Gillissen et al., 2000; Bürkle et al., 2003). OsPUP7 was proved to transport caffeine in *fcy2* yeast, which provides convincing evidence for its important role in cytokinin transport (Qi and Xiong, 2013). CcPUP1 and CcPUP5 were demonstrated to facilitate the transport of adenine, but not caffeine (Kakegawa et al., 2019). Recently, an expanding role of PUPs in the transportation of both purine-alkaloids and non-purine-alkaloids was also identified (Jelesko, 2012).

Here, eight CsPUP members were identified in tea plants and named by their homology with Arabidopsis (Table 1). The amino acid sequences of CsPUPs were highly conserved. Ten TMDs were detected in most members, which may have an important effect on protein function (Figure 1). Consistent results were also found in Arabidopsis, maize, and rice, implying similar transport function of CsPUPs in tea plants (Qi and Xiong, 2013). The CsPUP members could be divided into classes I, II, and III by phylogenetic analysis. PUPs with higher sequence similarity in different species were clustered into one group. Whereas, gene members from the same species showed further genetic distance on the evolutionary tree (Figure 2). The eight CsPUPs were randomly distributed on six chromosomes in the tea genome (Figure 3). These results indicated that both sub-functionalization and neo-functionalization occurred during the evolution of gene and species (Jelesko, 2012). The function of CsPUPs transporters in tea plants were also expected to diversified considerably. Analysis of gene expression in different tissues revealed that CsPUPs were mainly expressed in the aerial parts of tea plants. The expression levels of CsPUP1, CsPUP3.1, CsPUP3.3, CsPUP10.1, and CsPUP10.2 were higher in leaves, whereas CsPUP3.2 and CsPUP4 showed the highest expression level in stems (Figure 4). PUPs in other plants were also reported to have different expression characteristics. AtPUP1 was highly expressed in leaves, stems, and flowers (Gillissen et al., 2000). AtPUP2 was mainly expressed in the vascular tissue of leaves. The expression of AtPUP3 was restricted in pollen (Bürkle et al., 2003). AtPUP14 was widely expressed in all tissues and organs (Zürcher et al., 2016). The differences of gene expression patterns implied variations of protein function in another way. Previous studies revealed that gene expression patterns were closely correlated with the distribution of their substrates in plants. NtNUP1 showed the greatest expression level in root tips, which is the main organ nicotine biosynthesis. Further study proved that NtNUP1 mediated the uptake of nicotine in root cells and thus affected its accumulation in tobacco (Hildreth et al., 2011). There are abundant codeine and morphine in the latex of opium poppy. PsBUB1 was also highly expressed in the latex and functioned as benzylisoquinoline alkaloid transporters (Dastmalchi et al., 2019). Distinct from the alkaloid distributions in tobacco and opium poppy, caffeine in tea plants was found primarily in leaves, followed by stems and roots (Facchini, 2001).

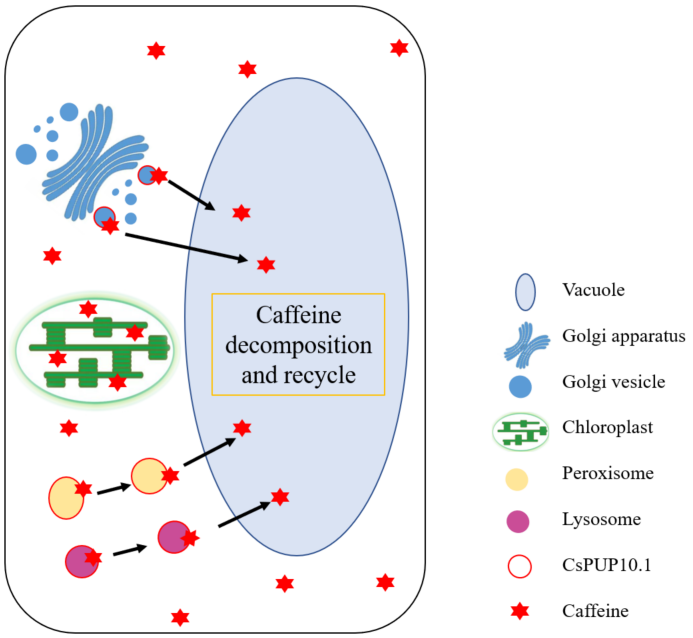


FIGURE 10
Putative mechanism of CsPUP10.1 in caffeine transport.

Therefore, CsPUP1, CsPUP3.1, and CsPUP10.1, which showed higher expression level in leaves, were selected for further investigation. The results of correlation analysis revealed that CsPUP1, CsPUP3.1, and CsPUP10.1 were significantly negatively correlated with caffeine content (Table 2). It suggested their potential roles in caffeine metabolism.

The complementation assay in the yeast and subcellular localization analysis demonstrated that CsPUP1, CsPUP3.1, and CsPUP10.1 functioned as caffeine influx transporters in plasma membrane (Figures 7, 8B). This functional mode of CsPUPs was coincident with their homologous genes in other plants. AtPUP1 was localized on plasma membrane. It was the first member to be reported to promote the caffeine uptake in the yeast expression system (Gillissen et al., 2000; Szydlowski et al., 2013). NtNUP1 was primarily localized on plasma membrane and showed nicotine-specific uptake activity (Hildreth et al., 2011). In opium poppy, six of the nine PsBUPs facilitated the transportation of alkaloids and increased their yields when expressed in host yeast. PsBUP1 was found to be a laticifer-specific plasma membrane transporter (Dastmalchi et al., 2019). OsPUP7 was proved to transport caffeine into cells through complementary experiment in the yeast *fcy2* mutant (Qi and Xiong, 2013).

For most of the species without any alkaloids produced, exogenous alkaloid has inhibitory and toxic effects on their growth and development. Whereas for the plants producing such metabolites, they develop characteristic self-tolerance mechanisms and usually are insensitive to alkaloids (Shitan, 2016). This phenomenon was demonstrated in Arabidopsis, tobacco, pea, and yeast (Bard et al., 1980; Mohanpuria and Yadav, 2009b; Curlango-Rivera et al., 2010; Alkhatib et al., 2016). Consistent phenotype was also observed in our study. Interestingly, compared with the blank group, repressed growth of transformants could be partly rescued by CsPUPs (Figures 8A, 9A). This phenomenon was distinct from studies on non-caffeine plants. Transgenic yeasts expressed AtPUP1 and OsPUP7 showed worse growth conditions than control under exogeneous caffeine treatment (Bürkle et al., 2003; Qi and Xiong, 2013). Transgenic *fcy2* yeast with CcPUP1 and CcPUP5 also showed more pronounced inhibited phenotype than control (Kakegawa et al., 2019). Whereas, yeast transformants with CsPUP1, CsPUP3.1, and CsPUP10.1 exhibited better growth conditions than control in media containing caffeine (Figure 8A). Similar phenotype was also confirmed in CsPUP10.1-over-expressed Arabidopsis lines (Figure 9A). It suggested that CsPUPs functioned distinctively in caffeine transport in tea plants.

Recently, functional identification of theanine transporters in tea plants was reported. The complementation assays of amino acid permeases showed that they could transport theanine into yeast cells as nitrogen source. Thus, transgenic yeast exhibited alleviated toxicity and better growth condition than control under theanine treatment (Dong et al., 2020). Moreover, transformant yeast with AtPUP1 imported exogenous adenine into cells as nitrogen source and grew better than control. Further studies revealed that AtPUP1

functioned as a cytokinin transporter and was involved in its retrieval from the xylem sap (Bürkle et al., 2003). Moderate exogenous caffeine was proved to positively regulate plant growth in sunflower and tobacco (Khursheed et al., 2009; Alkhatib et al., 2016). In addition, the absorption and degradation of exogenous caffeine were also revealed in wheat, barley, poplar, and lettuce (Pierattini et al., 2016; Yahyazadeh et al., 2017; Chuang et al., 2019; Broughton et al., 2020; Vannucchi et al., 2020). Caffeine in tea leaves could be slowly catabolized along with the growth and development of plants (Mohanpuria et al., 2009a; Ashihara et al., 1997; Zhu et al., 2019). Some microorganisms were also reported to degrade caffeine and its derivates as nitrogen source (Mazzafera, 2002; Mazzafera, 2004; Dash and Gummadi, 2006). These studies indicated that part of exogenous caffeine could be decomposed and recycled as nitrogen source in different organisms to maintain their survival and growth.

Furthermore, in addition to that in the plasma membrane, abundant and stronger GFP signals of CsPUPs were also observed in the form of hollow or solid dots and distributed widely in cells (Figure 7). According to the shape and allocation, it was speculated that they were whether vesicles, peroxisomes, or lysosomes (Uemura et al., 2004; Winters et al., 2020). The characteristic localization patterns of these CsPUPs implied that they might shuttle between the organelles and function distinctly in tea plants from other species.

Intracellular and extracellular vesicles play an essential role in the transportation and compartmentation of secondary metabolites (Shoji et al., 2014; Zhao et al., 2020). The transport vesicles could bound to membrane and carry substrates from the organelle (generally ER) where their biosynthesis takes place to another organelle (generally vacuoles) (Bonifacino and Glick, 2004). Vacuoles can be divided into two types: lytic vacuoles (LVs) and protein storage vacuoles (PSVs). Hydrolases in LVs degrade unwanted cellular substances. Whereas, PSVs accumulate abundant substances for storage (Tan et al., 2019). The final distribution patterns of metabolites vary with cell types, tissues, developing stages, and plant species (Shitan and Yazaki, 2007; Shitan and Yazaki, 2020; Gani et al., 2021). This transport mechanism was widely demonstrated in flavonoid accumulation in plants (Zhao and Dixon, 2009; Zhao, 2015; Pucker and Selmar, 2022). Hyper-accumulated anthocyanin in purple tea leaves was also proved to be closely related with active vesicular trafficking (Wei et al., 2019).

Peroxisome is a circular organelle and widely distributes in nearly all eukaryotes. The shape of peroxisome could be changed from sphere to ellipse along with plant growth and environmental conditions (Beevers, 1979). This process is closely associated with diverse metabolic pathways in plants, including catabolism of polyamines, biosynthetic pathway of hormones, and alleviation of the toxicity of reactive oxygen species and other harmful products (Oikawa et al., 2019).

Lysosome is another round organelle and displays as vesicular structure in eukaryotes. The lysosome system is one of the powerful hydrolytic mechanisms involved in the degradation of cellular components (Lloyd, 1996). Numerous metabolite transporters were discovered in the lysosome membrane. The degradation products of nutriments, such as amino acids, could be transported by these transmembrane proteins for recycle and reutilization (Tong et al., 2010).

To date, the exact subcellular organelle of caffeine synthesized, stored, and degraded remains unclear. The site of caffeine biosynthesis seemed to vary with species. Caffeine biosynthesis in tea plants was closely associated with chloroplasts (Ashihara et al., 2013). Whereas, caffeine in coffee appears to be synthesized in the cytoplasm (Ogawa et al., 2001; Kumar et al., 2007). Vacuole was a putative location for caffeine storage in plants (Wink, 2010).

Therefore, it can be inferred that CsPUPs functioned as caffeine transporters in two ways in heterologous expression systems. For one thing, CsPUPs on plasma membrane imported the external caffeine into cells, which caused more caffeine accumulation in the transgenic yeast and Arabidopsis than control. For another thing, deposited caffeine in cells could be comparted and transported by CsPUPs. Then, part of caffeine could be degraded and converted into a form of nitrogen source for further utilization in vacuole (Figure 10). Thus, transformants with CsPUPs showed better growth conditions and enhanced tolerance to caffeine. For tea plants with endogenous caffeine produced, caffeine was widely distributed in leaf cells. CsPUPs showed extended subcellular localization and distinct phenotypic differences in heterologous expression systems. These results implied characteristic mechanism of caffeine transport in tea plants. Compartmentalized caffeine in tea plant cells could be delivered by CsPUPs from the site of synthesis or storage to the other organelles for decomposition and further utilization. In conclusion, these differences better explain the distinctive mechanisms of CsPUPs in caffeine transport in tea plants.

In summary, in total, eight PUP members from the whole genome of tea plant were identified. Then, bioinformatics and expression patterns of these genes were analyzed to determine their potential roles in caffeine metabolism. Among of them, CsPUP1, CsPUP3.1, and CsPUP10.1 showed higher expression level in ML and were significantly negatively correlated with caffeine content. Complementation assays in yeast revealed that CsPUP1, CsPUP3.1, and CsPUP10.1 functioned as caffeine transporters. Further heterologous expression analysis in Arabidopsis re-confirmed the function of CsPUP10.1. Combined with the results of subcellular localization and expression mode, it suggested that CsPUPs play an important role in the intracellular transport of caffeine in tea plants. Deposited caffeine in tea leaves might be transported by CsPUPs among different subcellular organelles for facilitating its metabolism. This study provides a new insight into the molecular mechanisms of caffeine regulatory network in tea plants.

Data availability statement

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

Author contributions

Conceived and designed the experiment: KW, HC, and LW. Performed the experiments: YZ, LG, and YL. Analyzed the data: YZ, KW, LG, and LW. Contributed reagents, materials, and analysis tools: HC, LW, and CC. Contributed to the writing of the manuscript: YZ, KW, and LW. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by China Agriculture Research System of MOF and MARA (CARS-19) and Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding Tea Plant (2021C02067-7-1).

Acknowledgments

We thank Hua Zhao (College of Horticulture and Forestry Sciences of Huazhong Agricultural University) for contributing the *fcy2* mutant.

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.2022.1033316/full#supplementary-material

ADDITIONAL FILE 2
The overall expression levels of CsPUPs in different tissues.

ADDITIONAL FILE 3
Growth conditions of Arabidopsis on normal 1/2 MS medium.

References

Alkhatib, R., Alkhatib, N., Al-Quraan, N., Al-Eitan, L., Abdo, N., and Muhaidat, R. (2016). Impact of exogenous caffeine on morphological, biochemical, and ultrastructural characteristics of *Nicotiana tabacum*. *Biol. Plantarum.* 60 (4), 706–714. doi: 10.1007/s10535-016-0600-z

Ashihara, H., Gillies, F. M., and Crozier, A. (1997). Metabolism of caffeine and related purine alkaloids in leaves of tea (Camellia sinensis L.). *Plant Cell Physiology* 38 (4), 413–419.

Ashihara, H., and Suzuki, T. (2004). Distribution and biosynthesis of caffeine in plants. *Front. Bioscience.* 9, 1864–1876. doi: 10.2741/1367

Ashihara, H., Yokota, T., and Crozier, A. (2013). Biosynthesis and catabolism of purine alkaloids. *Adv. Botanical Res.* 68, 111–138. doi: 10.1016/B978-0-12-408061-4.00004-3

Bard, M., Neuhauser, J., and Leeds, N. (1980). Caffeine resistance of *Saccharomyces cerevisiae*. *J. Bacteriology.* 141, 999–1002. doi: 10.1128/jb.141.2.999-1002.1980

Beevers, H. (1979). Microbodies in higher plants. *Annu. Rev. Plant Physiol.* 30, 159–193. doi: 10.1146/annurev.pp.30.060179.001111

Bonifacino, J. ,. S., and Glick, B. ,. S. (2004). The mechanisms of vesicle budding and fusion. *Cell* 116, 153–166. doi: 10.1016/S0092-8674(03)01079-1

Broughton, S., Castello, M., Liu, L., Killen, J., Hepworth, A., and O'Leary, R. (2020). The effect of caffeine and trifluralin on chromosome doubling in wheat anther culture. *Plants.* 9, 105. doi: 10.3390/plants9010105

Bürkle, L., Cedzich, A., Dopke, C., Stransky, H., Okumoto, S., Gillissen, B., et al. (2003). Transport of cytokinins mediated by purine transporters of the PUP family expressed in phloem, hydathodes, and pollen of arabidopsis. *Plant J.* 34, 13–26. doi: 10.1046/j.1365-313X.2003.01700.x

Celik, T., Iyisoy, A., and Amasyali, B. (2010). The effects of coffee intake on coronary heart disease: Ongoing controversy. *Int. J. Cardiol.* 144 (1), 118. doi: 10.1016/j.ijcard.2008.12.112

Chuang, Y. H., Liu, C. H., Sallach, J. ,. B., Hammerschmidt, R., Zhang, W., Boyd, S. A., et al. (2019). Mechanistic study on uptake and transport of pharmaceuticals in lettuce from water. *Environ. Int.* 131, 104976. doi: 10.1016/j.envint.2019.104976

Curlango-Rivera, G., Duclos, V. ,. D., Ebolo, J. ,. J., and Hawes, C. ,. M. (2010). Transient exposure of root tips to primary and secondary metabolites: Impact on root growth and production of border cells. *Plant Soil.* 332, 267–275. doi: 10.1007/s11104-010-0291-8

Dash, S. ,. S., and Gummadi, S. ,. N. (2006). Catabolic pathways and biotechnological applications of microbial caffeine degradation. *Biotechnol. Letters.* 28 (24), 1993–2002. doi: 10.1007/s10529-006-9196-2

Dastmalchi, M., Chang, L., Chen, R., Yu, L., Chen, X., Hagel, J. ,. M., et al. (2019). Purine permease-type benzylisoquinoline alkaloid transporters in opium poppy. *Plant Physiol.* 181 (3), 916–933. doi: 10.1104/pp.19.00565

Dong, C. ,. X., Li, F., Yang, T. ,. Y., Feng, L., Zhang, S. ,. P., Li, F. ,. D., et al. (2020). Theanine transporters identified in tea plants (*Camellia sinensis* l.). *Plant J.* 101, 57–70. doi: 10.1111/tpj.14517

Facchini, P. ,. J. (2001). Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 29–66. doi: 10.1146/annurev.arplant.52.1.29

Ferreira, T., Brèthes, D., Pinson, B., Napias, C., and Chevallier, J. (1997). Functional analysis of mutated purine-cytosine permease from *Saccharomyces cerevisiae*. *J. Biol. Chem.* 272 (150), 9697–9702. doi: 10.1074/jbc.272.15.9697

Gani, U., Vishwakarma, R. ,. A., and Misra, P. (2021). Membrane transporters: the key drivers of transport of secondary metabolites in plants. *Plant Cell Rep.* 40, 1–18. doi: 10.1007/s00299-020-02599-9

Gillissen, B., Bürkle, L., Andre, B., Kuhn, C., Rentsch, D., Brandl, B., et al. (2000). A new family of high-affinity transporters for adenine, cytosine, and purine derivatives in arabidopsis. *Plant Cell.* 12, 291–300. doi: 10.1105/tpc.12.2.291

Gramza-Michalowska, A. (2014). Caffeine in tea *camellia sinensis*- content, absorption, benefits and risks of consumption. *J. Nutrition Health& Aging.* 18 (2), 143–149. doi: 10.1007/s12603-013-0404-1

Hildreth, S. ,. B., Gehman, E. ,. A., Yang, H., Lu, R. ,. H., Ritesh, K. ,. C., Harich, K. ,. C., et al. (2011). Tobacco nicotine uptake permease (NUP1) affects alkaloid metabolism. *Proc. Natl. Acad. Sci.* 108 (44), 18179–18184. doi: 10.1073/pnas.1108620108

Jelesko, ,. J. ,. G. (2012). An expanding role for purine uptake permease-like transporters in plant secondary metabolism. *Front. Plant Science.* 3 (78), 1–5. doi: 10.3389/fpls.2012.00078

Jin, J. ,. Q., Yao, M. ,. Z., Ma, C. ,. L., Ma, ,. J. ,. Q., and Chen, L. (2016). Natural allelic variations of TCS1 play a crucial role in caffeine biosynthesis of tea plant and its related species. *Plant Physiol. Biochem.* 100, 18–26. doi: 10.1016/j.plaphy.2015.12.020

Kakegawa, H., Shitan, N., Kusano, H., Ogita, S., Yazaki, K., and Sugiyama, A. (2019). Uptake of adenine by purine permeases of *Coffea canephora*. *Bioscience Biotechnology Biochem.* 83 (7), 1300–1305. doi: 10.1080/09168451.2019.1606698

Kang, J., Lee, Y., Sakakibara, H., and Marinoia, E. (2017). Cytokinin transporters: go and stop in signaling. *Trends Plant Science.* 22 (6), 455–461. doi: 10.1016/j.tplants.2017.03.003

Kato, K., Shitan, N., Shoji, T., and Hashimoto, T. (2015). Tobacco NUP1 transports both tobacco alkaloids and vitamin B6. *Phytochemistry.* 113, 33–40. doi: 10.1016/j.phytochem.2014.05.011

Khursheed, T., Ansari, M., and Shahab, D. (2009). Studies on the effect of caffeine on growth and yield parameters in *Helianthus annuus* l. variety. *Modern. Biol. Med.* 2, 56–60. doi: 10.4172/0974-8369.1000014

Kumar, V., Satyanarayana, K. V., Ramakrishna, A., Chandrashekar, A., and Ravishankar, G. A. (2007). Evidence for localization of n-methyltransferase (MMT) of caffeine biosynthetic pathway in vacuolar surface of *Coffea canephora* endosperm elucidated through localization of GUS reporter gene driven by NMT promoter. *Curr. Science.* 93, 383–389. Available at: https://www.currentscience.ac.in/Volumes/93/03/0383.pdf

Livak, K. J., and 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 (4), 402–408. doi: 10.1006/meth.2001.1262

Lloyd, J. (1996). Metabolite efflux and influx across the lysosome membrane. *Subcellular Biochem.* 27, 361–386. doi: 10.1007/978-1-4615-5833-0_11

Mazzafera, P. (2002). Degradation of caffeine by microorganisms and potential use of decaffeinated coffee husk and pulp in animal feeding. *Scientia Agricola.* 59 (4), 815–821. doi: 10.1590/S0103-90162002000400030

Mazzafera, P. (2004). Catabolism of caffeine in plants and microorganisms. *Front. Bioscience.* 9, 1348–1359. doi: 10.2741/1339

Mohanpuria, P., Kumar, V., Joshi, R., Gulati, A., Ahuja, P. ,. S., and Yadav, S. ,. K. (2009a). Caffeine biosynthesis and degradation in tea [*Camellia sinensis*(L.)O. kuntze] is under developmental and seasonal regulation. *Mol. Biotechnol.* 43, 104–111. doi: 10.1007/s12033-009-9188-2

Mohanpuria, P., and Yadav, S. ,. K. (2009b). Retardation in seedling growth and induction of early senescence in plants upon caffeine exposure is related to its negative effect on rubisco. *Photosynthetica.* 47 (2), 293–297. doi: 10.1007/s11099-009-0045-0

Morita, M., Shitan, N., Sawada, K., Montagu, M. ,. V., Inze´, ,. D., Rischer, H., et al. (2009). Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in *Nicotiana tabacum*. *Proc. Natl. Acad. Sci.* 106 (7), 2447–2452. doi: 10.1073/pnas.0812512106

Ogawa, M., Herai, Y., Koizumi, N., Kusano, T., and Sano, H. (2001). 7-methylxanthine methyltransferase of coffee plants. gene isolation and enzymatic properties. *J. Biol. Chem.* 276, 8213–8218. doi: 10.1074/jbc.M009480200

Oikawa, K., Hayashi, M., Hayashi, Y., and Nishimura, M. (2019). Re-evaluation of physical interaction between plant peroxisomes and other organelles using live-cell imaging techniques. *J. Integr. Plant Biol.* 61 (7), 836–852. doi: 10.1111/jipb.12805

Pierattini, E., Francini, A., Raffaelli, A., and Sebastiani, L. (2016). Degradation of exogenous caffeine by *Populus alba* and its effects on endogenous caffeine metabolism. *Environ. Sci. pollut. Res.* 23, 7298–7307. doi: 10.1007/s11356-015-5935-z

Pucker, B., and Selmar, D. (2022). Biochemistry and molecular basis of intracellular flavonoid transport in plants. *Plants.* 11, 963. doi: 10.3390/plants11070963

Qi, Z., and Xiong, L. (2013). Characterization of a purine permease family gene OsPUP7 involved in growth and development control in rice. *J. Integr. Plant Biol.* 55 (11), 1119–1135. doi: 10.1111/jipb.12101

Shilo, L., Sabbah, H., Hadari, R., Kovatz, S., and Shenkman, L. (2002). The effects of coffee consumption on sleep and melatonin secretion. *Sleep Med.* 3 (3), 271–273. doi: 10.1016/S1389-9457(02)00015-1

Shitan, N. (2016). Secondary metabolites in plants: transport and self-tolerance mechanisms. *Bioscience Biotechnology Biochem.* 80 (7), 1283–1293. doi: 10.1080/09168451.2016.1151344

Shitan, N., Kato, K., and Shoji, T. (2014). Alkaloid transporters in plants. *Plant Biotechnol.* 31, 453–463. doi: 10.5511/plantbiotechnology.14.1002a

Shitan, N., and Yazaki, K. (2007). Accumulation and membrane transport of plant alkaloids. *Curr. Pharm. Biotechnol.* 8, 244–252. doi: 10.2174/138920107781387429

Shitan, N., and Yazaki, K. (2020). Dynamism of vacuoles toward survival strategy in plants. *Biochim. Biophys. Acta-Biomembranes.* 1862, 183127. doi: 10.1016/j.bbamem.2019.183127

Shoji, T., Inai, K., Yazaki, Y., Sato, Y., Takase, H., Shitan, N., et al. (2009). Multidrug and toxic compound extrusion-type transporters implicated in vacuolar sequestration of nicotine in tobacco roots. *Plant Physiol.* 149, 708–718. doi: 10.1104/pp.108.132811

Shoji, J. ,. Y., Kikuma, T., and Kitamoto, K. (2014). Vesicle trafficking, organelle functions, and unconventional secretion in fungal physiology and pathogenicity. *Curr. Opin. Microbiol.* 20, 1–9. doi: 10.1016/j.mib.2014.03.002

Suzuki, T., Ashihara, H., and Waller, G. ,. R. (1992). Purine and purine alkaloid metabolism in camellia and *Coffea* plants. *Phytochemistry* 31 (8), 2575–2584. doi: 10.1016/0031-9422(92)83590-U

Szydlowski, N., Bürkle, L., Pourcel, L., Moulin, M., Stolz, J., and Fitzpatrick, T. ,. B. (2013). Recycling of pyridoxine (vitamin B6) by PUP1 in arabidopsis. *Plant J.* 75 (1), 40–52. doi: 10.1111/tpj.12195

Tan, X., Li, K., Wang, Z., Zhu, K., Tan, X., and Cao, J. (2019). A review of plant vacuoles: formation, located proteins, and functions. *Plants.* 8, 327. doi: 10.3390/plants8090327

Tong, J., Yan, X., and Yu, L. (2010). The late stage of autophagy: cellular events and molecular regulation. *Protein Cell.* 1 (10), 907–915. doi: 10.1007/s13238-010-0121-z

Uemura, T., Ueda, T., Ohniwa, R. ,. L., Nakano, A., Takeyasu, K., and Sato, M. ,. H. (2004). Systematic analysis of SNARE molecules in arabidopsis: dissection of the post-golgi network in plant cells. *Cell Structure Funct.* 29, 49–65. doi: 10.1247/csf.29.49

Vannucchi, F., Traversari, S., Raffaelli, A., Francini, A., and Sebastiani, L. (2020). *Populus alba* tolerates and efficiently removes caffeine and zinc excesses using an organ allocation strategy. *Plant Growth Regulation.* 92, 597–606. doi: 10.1007/s10725-020-00664-7

Wang, Y. ,. X., Wei, K., Ruan, L., Bai, P. ,. X., Wu, L. ,. Y., Wang, L. ,. Y., et al. (2022). Systematic investigation and expression profiles of the nitrate transporter 1/peptide transporter family (NPF) in tea plant (*Camellia sinensis*). *Int. J. Mol. Sci.* 23 (12), 6663. doi: 10.3390/ijms23126663

Wei, K., Wang, L. ,. Y., Zhang, Y. ,. Z., Ruan, L., Li, H. ,. L., Wu, L. ,. Y., et al. (2019). A coupled role for *CsMYB75* and *CsGSTF1* in anthocyanin hyperaccumulation in purple tea. *Plant J.* 97 (5), 825–840. doi: 10.1111/tpj.14161

Wink, M. (2010). Introduction: biochemistry, physiology and ecological functions of secondary metabolites. *Annu. Plant Rev.* 40, 1–19. doi: 10.1002/9781444320503.ch1

Winters, C. ,. M., Hong-Brown, L. ,. Q., and Chiang H, L. (2020). Intracellular vesicle clusters are organelles that synthesize extracellular vesicle-associated cargo proteins in yeast. *J. Biol. Chem.* 295 (9), 2650–2663. doi: 10.1074/jbc.RA119.008612

Xiao, Y., Liu, D., Zhang, G., Gao, S., Liu, L., Xu, F., et al. (2018). Big grain3, encoding a purine permease, regulates grain size *via* modulating cytokinin transport in rice. *J. Integr. Plant Biol.* 61, 581–597. doi: 10.1111/jipb.12727

Xiao, Y., Zhang, J., Yu, G., Lu, X., Mei, W., and Deng, H. (2020). Endoplasmic reticulum-localized PURINE PERMEASE1 regulates plant height and grain weight by modulating cytokinin distribution in rice. *Front. Plant Science.* 11, 61856. doi: 10.3389/fpls.2020.618560

Yahyazadeh, M., Nowak, M., Kima, H., and Selmar, D. (2017). Horizontal natural product transfer: A potential source of alkaloidal contaminants in phytopharmaceuticals. *Phytomedicine.* 34, 21–25. doi: 10.1016/j.phymed.2017.07.007

Zhang, F., He, W., Yuan, Q. ,. Y., Wei, K., Ruan, L., Wang, L. ,. Y., et al. (2021). Transcriptome analysis identifies CsNRT genes involved in nitrogen uptake in tea plants, with a major role of CsNRT2.4. *Plant Physiol. Biochem.* 167, 970–979. doi: 10.1016/j.plaphy.2021.09.024

Zhang, Y. ,. Z., Wang, L. ,. Y., Wei, K., Ruan, L., Wu, L. ,. Y., He, M. ,. D., et al. (2020). Differential regulatory mechanisms of secondary metabolites revealed at different leaf positions in two related tea cultivars. *Scientia Horticulturae*. 272, 109579. doi: 10.1016/j.scienta.2020.109579

Zhao, J. (2015). Flavonoid transport mechanisms: how to go, and with whom. *Trends Plant Science*. 20 (9), 576–585. doi: 10.1016/j.tplants.2015.06.007

Zhao, J., and Dixon, R. (2009). The 'ins' and 'outs' of flavonoid transport. *Trends Plant Science*. 15 (2), 72–80. doi: 10.1016/j.tplants.2009.11.006

Zhao, J., Li, P. ,. H., Xia, T., and Wan, X. ,. C. (2020). Exploring plant metabolic genomics: chemical diversity, metabolic complexity in the biosynthesis and transport of specialized metabolites with the tea plant as a model. *Crit. Rev. Biotechnol* 40 (5), 667–688. doi: 10.1080/07388551.2020.1752617

Zhu, B. ,. Y., Chen, L. ,. B., Lu, M. ,. Q., Zhang, J., Han, J., Deng, W. ,. W., et al. (2019). Caffeine content and related gene expression: novel insight into caffeine metabolism in *Camellia* plants containing low, normal and high caffeine concentrations. *J. Agric. Food Chem*. 67 (12), 3400–3411. doi: 10.1021/acs.jafc.9b00240

Zürcher, E., Liu, J., Donato, M., Geisler,, and Müller, B. (2016). Plant development regulated by cytokinin sinks. *Science*. 353 (6303):1027–30. doi: 10.1126/science.aaf7254

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 22 December 2022
DOI 10.3389/fpls.2022.1022075

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Hai Lu,
Beijing Forestry University, China
Manvi Sharma,
Swedish University of Agricultural Sciences, Sweden

*CORRESPONDENCE
Tao Zhou
taozhou88@163.com

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 18 August 2022
ACCEPTED 01 December 2022
PUBLISHED 22 December 2022

CITATION
Xu J, Hu Z, He H, Ou X, Yang Y, Xiao C, Yang C, Li L, Jiang W and Zhou T (2022) Transcriptome analysis reveals that jasmonic acid biosynthesis and signaling is associated with the biosynthesis of asperosaponin VI in *Dipsacus asperoides*. *Front. Plant Sci.* 13:1022075. doi: 10.3389/fpls.2022.1022075

Transcriptome analysis reveals that jasmonic acid biosynthesis and signaling is associated with the biosynthesis of asperosaponin VI in *Dipsacus asperoides*

Jiao Xu, Zhengping Hu, Hua He, Xiaohong Ou, Yang Yang, Chenghong Xiao, Changgui Yang, Liangyuan Li, Weike Jiang and Tao Zhou*

Resource Institute for Chinese Medicine and Ethnic Materia Medica, Guizhou University of Traditional Chinese Medicine, Guiyang, China

Dipsacus asperoides is a perennial herb, the roots of which are abundant in asperosaponin VI, which has important medicinal value. However, the molecular mechanism underlying the biosynthesis of asperosaponin VI in *D. asperoides* remains unclear. In present study, a comprehensive investigation of asperosaponin VI biosynthesis was conducted at the levels of metabolite and transcript during root development. The content of asperosaponin VI was significantly accumulated in two-leaf stage roots, and the spatial distribution of asperosaponin VI was localized in the xylem. The concentration of asperosaponin VI gradually increased in the root with the development process. Transcriptome analysis revealed 3916 unique differentially expressed genes (DEGs) including 146 transcription factors (TFs) during root development in *D. asperoides*. In addition, α-linolenic acid metabolism, jasmonic acid (JA) biosynthesis, JA signal transduction, sesquiterpenoid and triterpenoid biosynthesis, and terpenoid backbone biosynthesis were prominently enriched. Furthermore, the concentration of JA gradually increased, and genes involved in α-linolenic acid metabolism, JA biosynthesis, and triterpenoid biosynthesis were up-regulated during root development. Moreover, the concentration of asperosaponin VI was increased following methyl jasmonate (MeJA) treatment by activating the expression of genes in the triterpenoid biosynthesis pathway, including acetyl-CoA acetyltransferase (*DaAACT*), 3-hydroxy-3-methylglutaryl coenzyme A synthase (*DaHMGCS*), 3-hydroxy-3-methylglutaryl coenzyme-A reductase (*DaHMGCR*). We speculate that JA biosynthesis and signaling regulates the expression of triterpenoid biosynthetic genes and facilitate the biosynthesis of asperosaponin VI. The results suggest a regulatory network wherein triterpenoids, JA, and TFs co-modulate the biosynthesis of asperosaponin VI in *D. asperoides*.

KEYWORDS

***Dipsacus asperoides*, asperosaponin VI, transcriptome, jasmonic acid, triterpenoid biosynthesis**

Introduction

Dipsacus asperoides is a perennial herb belonging to the Dipsacaceae family (Figure 1A). It is a Chinese herbal medicinal that is used for the treatment of bone fractures and joint diseases such as osteoporosis, lassitude in the loin and legs, and fractures by preventing bone loss, increasing osteoblastic activity, and decreasing osteoclastic activity (Niu et al., 2012; Niu et al., 2015). Notably, asperosaponin VI is the main active component of *D. asperoides*, which is an oleanane triterpenoid saponin. In addition, asperosaponin VI may increase bone formation by increasing BMP-2 synthesis and activating both p38 and ERK1/2 (Niu et al., 2011). Asperosaponin VI may be a potential agent for suppressing neuroinflammation related to Alzheimer's disease (Yu et al., 2012). Moreover, because it activates the PI3K/Akt and CREB pathways, asperosaponin VI has protective effects against hypoxia-induced cardiomyocytes apoptosis (Li et al., 2010). Therefore, *D. asperoides* has very important medicinal value because of its content of asperosaponin VI.

Asperosaponin VI is a triterpenoid saponin, and thus it is biosynthesized *via* the mevalonic acid (MVA) pathway which has been found in previous studies (Yang et al., 2017; Chen et al., 2018; Yang et al., 2018). Typically, the pathway involves the conversion of isoprenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) into farnesyl pyrophosphate (FPP), which then generates squalene and 2,3-oxidosqualene, and is followed by a series of reactions involving cyclization, oxidation, hydroxylation and glycosylation (Kuzuyama, 2002; Augustin et al., 2011; Niu et al., 2014), which is accompanied by a variety of enzymatic reactions involving various enzymes such as farnesyl pyrophosphate synthase (FPPS), β-amyrin synthase (β-AS), and dammarenediol synthase (DS). There are great differences in the quality of *D. asperoides* from different production areas, and the content of asperosaponin VI also varies greatly due to great differences in the environment, such as altitude, temperature, and light conditions (Jiang et al., 2013). Moreover, germplasm resources differ greatly in different producing areas, resulting in different qualities of *D. asperoides* (Xiao et al., 2018). Previous studies have illustrated that the genes encoding asperosaponin VI synthetase are involved in the triterpenoid biosynthesis pathway (Wang et al., 2016). However, the genes belong to multigene families that are complex and diverse. To date, little is known about the genes functioning in triterpenoid biosynthesis in *D. asperoides*. Additionally, the mechanisms of its biosynthesis and regulation are unclear. Hence, analyzing the biosynthesis pathway of triterpenoid saponins and verifying the gene function are important.

To explore the biosynthesis pathway of triterpenoid saponins, omics methods have been conducted in recent years. The genomes of *P. notoginseng* and *P. ginseng* have been sequenced. UDP-glycosyltransferases, which belong to one of the largest gene families, catalyze the formation of different ginsenosides. Five glycosyltransferase genes were identified to catalyze the formation of different ginsenosides in *P. notoginseng* (Xu et al., 2017; Jiang et al., 2020). Other than glycosyltransferase genes, oxidosqualene cyclases, and cytochrome P450 are also involved in the biosynthesis of triterpenoid saponins (Yang et al., 2021). Moreover, transcriptome analysis has already been applied to investigations on terpenoid metabolism (Garg et al., 2015; Fan et al., 2019). Plant growth conditions play important roles in the biosynthesis of triterpenoid saponins. Research shows that more gibberellins and steroids accumulate in the root with quick vegetative growth under favorable conditions, and abscisic acid (ABA) and ginsenosides also accumulate in the root to enhance resistance under abiotic stress. In a functional study using transgenic *Arabidopsis*, terpene synthase genes were expressed and terpenoid contents increased compared with the wild-type control. Transcriptome and metabolome data showed the complexity of metabolic genes in terpenoid-rich *Salvia guaranitica* (Ali et al., 2018). With respect to asperosaponin VI, transcriptome analysis is urgently required to reveal the mechanism of regulation and biosynthesis of asperosaponin VI

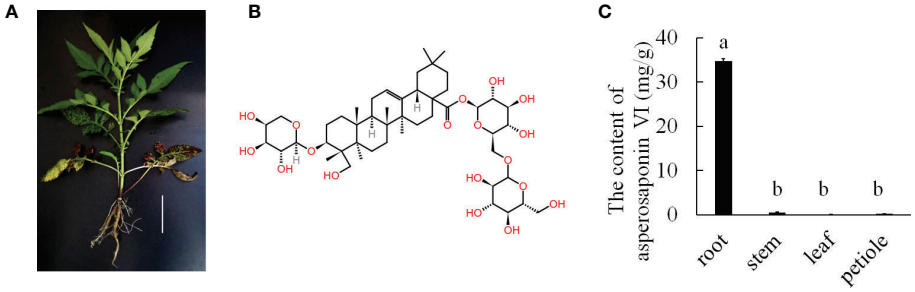


FIGURE 1
The major medicinally active triterpenoid in *D. asperoides*. **(A)**, the plant of *D. asperoides*. Scale bar = 10 cm. **(B)**, the typical triterpenoid of asperosaponin VI. **(C)**, the content of asperosaponin VI in different tissues of *D. asperoides*. The error bars represent the standard deviation with at least three biological replicates. Different lowercase letters denote significant differences by multiple comparisons using Statistix 8.0 software.

in *D. asperoides*. In addition, it is urgent to focus on molecular breeding to address the quality issue in *D. asperoides*.

In this study, matrix-assisted laser desorption ionization-mass spectrometry imaging (MALDI-MSI) was used to observe the spatial distribution pattern of asperosaponin VI. The roots at different stages of development were selected for detecting the content of asperosaponin VI. Three periods L1, L2 and M7 with large differences in asperosaponin VI content were selected for transcriptome sequencing by Illumina sequencing technology. Gene Ontology (GO) analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, and methyl jasmonate (MeJA) treatment were used to identify the gene expression and regulation signaling, thereby revealing new mechanistic insights into asperosaponin VI biosynthesis in *D. asperoides*.

Materials and methods

Plant materials and treatments

The *D. asperoides* plants were cultivated in March at Guizhou University of Traditional Chinese Medicine, Guizhou Province, China. The roots of one-leaf stage (L1), two-leaf stage (L2), and four-leaf stage (L4) plants, 7 months (M7) and 12 months (M12), as well as roots, leaves, stems, and petioles from plants that had been growing for approximate 2 years were frozen immediately in liquid nitrogen and stored in a −80 °C refrigerator for RNA isolation, JA and asperosaponin VI detection. The seedlings were cultured in ½ Murashige & Skoog (MS) solution under 16 h light/8 h dark conditions at 25 °C for MeJA treatment.

MALDI MSI and data analysis

Whole plants of *D. asperoides* were dug out from the field. The experiment was performed as previously described (Sun et al., 2021). The roots of *D. asperoides* were flash-frozen in liquid nitrogen for 15 s and then was transferred to a −80°C refrigerator. The root tissues were cryo-sectioned into 20-μm cross sections at −20°C on a cryostat microtome (Thermo CryoStar NX50 NOVPD) and mounted onto a conductive side of indium tin oxide (ITO)-coated glass slide. Root sections from different periods were prepared for MALDI matrix optimization. All these root sections were then transferred to a closed container and vacuum dried for 10 min. Then 10 mg/mL of DHB (2,5-Dihydroxybenzoic acid) in ACN/H_2O/TFA (Acrylonitrile/H_2O/trifluoroacetic acid, 70:30:0.1, v/v/v) was selected as the matrix. Matrix coating was performed on a HTX TM-Sprayer^{TM} (HTX Technologies). The flow rate of the sprayer was set to 0.075 mL/min at 55°C. The track speed was set to 800 mm/min and the track spacing was 3 mm. MALDI-MSI imaging was conducted on a Rapiflex MALDI Tissue Typer TM TOF/TOF MS (Bruker Daltonics). The laser was fired at a repetition rate of 5000 Hz, and the spatial resolution was set to 100 μm. Raw mass spectra data were acquired over the *m/z* range of 80–1200. Raw mass spectra were imported into Data Analysis 4.0 software (Bruker Daltonics) to perform internal mass calibration. The MS images were viewed and processed using SCiLS Lab 2018b (GmbH) software.

Paraffin section analysis

The roots of seedlings at the L2 stages were fixed in 50% FAA (Formalin-Aceto-Alcohol) as described previously (Hou and Huang, 2005). And the sections stained with hematoxylin solution and eosin dye.

Extraction and detection of asperosaponin VI

Asperosaponin VI was extracted from the roots of the *D. asperoides* samples with methanol according to our previous method with at least three independent biological replicates (Jiang et al., 2013). Approximately 0.1 g dried root powder was mixed with 5 mL methanol and extracted with ultrasonic treatment at 59 kHz for 30 min. Then the supernatant was collected after centrifugation at 4000 rpm for 10 min and filtered through a 0.45 μm filter. The standard of asperosaponin VI was dissolved in methanol and then mixed and diluted with methanol to obtain a series of mixture standard solutions of different concentrations. The solutions were filtered through a 0.45 μm syringe filter before high-performance liquid chromatography (HPLC) analysis.

Samples were separated on a Wondasil C18 column (4.6 mm × 250 mm, 5 μm) and detected at 212 nm. The column temperature was maintained at 30°C, and the flow rate was 1 mL/min. The mobile phase consisted of 30% acetonitrile. Asperosaponin VI was used as a standard to determine the asperosaponin VI content of the *D. asperoides* tissue.

RNA isolation and RNA sequencing

Nine samples at different stages of development were used for RNA sequencing, including L1, L2, M7 with three biological replicates. Total RNA was extracted following the instructions of Eastep® Super Total RNA Extraction Kit (Promega). The integrity and quality of the RNA were assessed using RNA agarose gel electrophoresis and NanoDrop2000 (Thermo Fisher Scientific). The RIN value was determined by the RNA 6000 Nano Kit of an Agilent Bioanalyzer 2100 (Agilent Technologies).

The cDNA libraries and the RNA-Seq were performed at Shanghai Majorbio Bio-pharm Technology Co., Ltd. (Shanghai,

China). The mRNA was purified using oligo (dT) magnetic beads. Taking these short fragments as templates, double-stranded cDNA was synthesized by a SuperScript double-stranded cDNA synthesis kit (Invitrogen). Libraries were size selected for cDNA target fragments of 200~300 bp. After quantification by TBS380, nine RNA-Seq libraries were sequenced by an Illumina NovaSeq 6000 sequencer (Illumina) for 2×150 bp paired-end reads. About 6 Gb reads were obtained from each sample for *de novo* assembly.

De novo assembly and annotation

To obtain high-quality clean data to ensure the success of subsequent analyses, the data quality control was used by FastQC, the raw reads with N ratios exceeding 10%, low-end sequence bases (mass values less than 20), or sequences with adaptors were all filtered using fastx_toolkit Version 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/). Then all clean reads were used for *de novo* assembly with Trinity (http://trinityrnaseq.sourceforge.net/). Afterward, the TPM was calculated by RSEM (http://deweylab.github.io/RSEM/software). Furthermore, all transcripts were compared with the Nr (NCBI Non-redundant Protein Library), String (Public Member Function), Swiss-Prot (Swiss-Prot protein database), Pfam (Protein Family Database), and KEGG databases using Blastx to obtain corresponding annotation information. To identify putative transcription factors (TFs), the Blastx was performed against PlantTFDB (http://planttfdb.cbi.pku.edu.cn/) (Jin et al., 2017).

Analysis of differentially expressed genes

The differentially expressed genes (DEGs) were analyzed by comparing the Fragments Per Kilobase of transcript per Million mapped reads (FPKM) of the unigenes in the L1, L2, and M7 samples. Based on padj <0.05, absolute value $|\log_2$ fold change$|$ > 2, and filtering according to (FPKM) max > 30, the unigenes were considered to be significantly differentially expressed. All DEGs were subjected to GO enrichment using Blast2GO with a corrected *P-value* < 0.05 (http://www.blast2go.com/b2ghome). GO functional analysis was conducted according previously study (Mishra et al., 2009). The KEGG pathway was used to obtain metabolic pathway analysis with a *P-value* < 0.05.

Expression correlation analysis

After the genes of interest are obtained, the correlation coefficients between genes are obtained by Pearson correlation algorithm based on the correlation of gene expression, and the visual network diagram was drawn. The data were analyzed on the online tool of Majorbio Cloud Platform (Ren et al., 2022).

Quantitative real-time PCR analysis

Candidate unigenes from the RNA-Seq were selected and validated by qRT-PCR. The first-strand cDNA was synthesized from 1 μg of total RNA using a cDNA synthesis kit (Promega). qRT-PCR was performed in 20 μL reactions using an Applied Biosystem 7500 real-time PCR system (Applied Biosystems) with at least three independent biological replicates. The gene *DaActin103* was used as the internal control to normalize and calculate the gene expression (Liang et al., 2020). The relative gene expression was calculated using the $2^{-\Delta Ct}$ method. The primers used for qRT-PCR are listed in Supplementary Table S1.

Endogenous JA extraction and quantification

To measure the concentration of JA, approximately 0.1g root samples were extracted in 80% cold methanol overnight. Further extraction and quantification analyses of JA were performed as described previously (Liu et al., 2012). The standard JA was purchased from Sigma-Aldrich.

MeJA treatment

In order to detect the regulation of JA on the biosynthesis of asperosaponin VI in *D. asperoides*, seedlings at the L2 stages were treated with 150 μM MeJA (Hu et al., 2021). The roots were collected following MeJA treatment for 6 h and 5 d, rapidly placed in liquid nitrogen, and frozen at –80 °C for RNA isolation and asperosaponin VI detection.

Results

Plant phenotype and asperosaponin VI accumulation in *D. asperoides*

D. asperoides is a medicinal herb. Our results showed that the roots of *D. asperoides* were rich in asperosaponin VI (Figure 1B). The other tissues in the aboveground part of the plant such as the petiole, stem and leaf had very low levels of asperosaponin VI (Figure 1C). To monitor the dynamics of asperosaponin VI metabolism, the roots were sampled at different developmental stages (L1, L2, L4, M7, and M12) (Figures 2A–E). The results showed that the accumulation of asperosaponin VI gradually increased during development. Only a small amount of asperosaponin VI was detected at the L1 stage, while it was significantly accumulated at the L2 stage. Additionally, it reached a high level at the M7 stage (Figure 2F, Figure S1). We suspect that asperosaponin VI

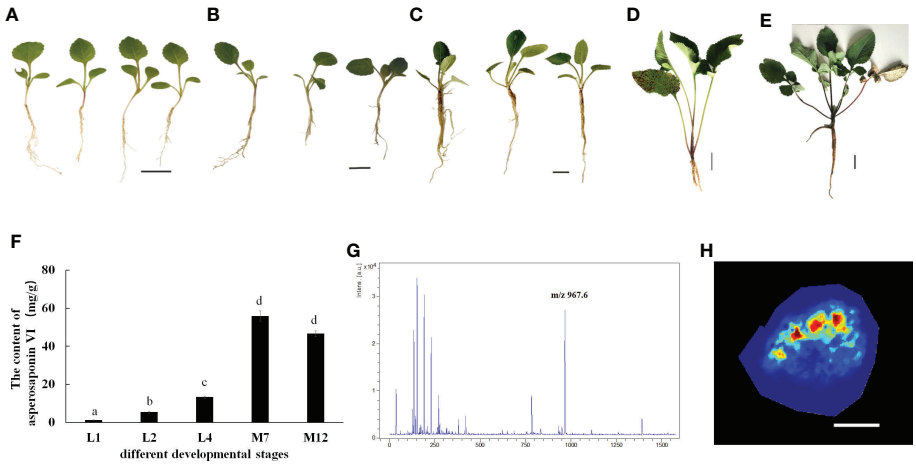


FIGURE 2
Accumulation Pattern of asperosaponin VI in root. **(A–E)**, the plants of D. asperoides in one leaf stage (L1), two leaf stage (L2), four leaf stage (L4), seven months (M7), one year (12M). Scale bars = 2 cm in **(A–C)**, Scale bars = 5 cm in **(D, E)**. **(F)**, the content of asperosaponin VI in **(A–E)**. The error bars represent the standard deviation with at least three biological replicates. Different lowercase letters denote significant differences by multiple comparisons using Statistix 8.0 software. **(G)**, the asperosaponin VI peak is most abundant (m/z 967). **(H)** Accumulation Pattern of asperosaponin VI in L2 root as observed by MALDI MSI. Scale bars = 1 mm.

gradually accumulated along with root development in the early growth stage in *D. asperoides*.

Spatial distribution pattern of asperosaponin VI

Given the accumulation of asperosaponin VI at initial period in the roots of L2, the roots were sampled to visualize the spatial distribution of asperosaponin VI by MALDI-MSI with a lateral resolution of 100 μm. The molecular ions detected in our study were identified according to the fragmentation of standard substances. The maximum peak area (Figure 2G) was identified as asperosaponin VI, which accumulated mainly in the xylem (Figure 2H, Figure S2). However, less was detected in the other parts. This result further illustrated that asperosaponin VI accumulated in the roots at stage L2, and furthermore, the accumulation of asperosaponin VI was associated with the development of the xylem in the roots of *D. asperoides*.

RNA sequencing and *De novo* assembly

To explore the mechanism of asperosaponin VI biosynthesis during the developmental stages and study the transcriptional regulatory networks for the biosynthesis of asperosaponin VI in *D. asperoides*, RNA-Seq was performed on *D. asperoides* roots during developmental stages of great difference (L1, L2, M7): the accumulation of asperosaponin VI at the beginning in the roots from L1 to L2, and the content of asperosaponin VI was the highest in the root of M7. Each sample had three biological replicates. Based on Illumina NovaSeq 6000, approximately 43 million clean reads of each sample were obtained with Q20 percentages of 97% or greater (Table S2). All clean reads were assembled by Trinity, resulting in 177195 transcripts with more than 75% of reads mapped. More than 124677 unigenes were expressed in the roots of *D. asperoides* (Table S3), 74463 unigenes were functionally annotated. The gene expression level was calculated based on the total mapped reads using FPKM.

Global analysis of DEGs

In the transcriptome database, a total of 27534 unigenes were detected and 3916 DEGs were obtained (Figure 3A). Compared with L1 and L2, there were more DEGs in M7 (Figure 3B). Compared with L1, 444 unigenes were up-regulated and 283 unigenes were down-regulated in L2. Only 108 unigenes were specifically differential expressed. By contrast, 1376 unigenes were up-regulated and 1456 unigenes were down-regulated in M7. Compared with L2, about 1463 unigenes were up-regulated and 1300 unigenes were down-regulated in M7. These results indicated that the content of asperosaponin VI increased and there were more DEGs as the roots developed. Thus, we believe that the biosynthesis of asperosaponin VI may regulated by DEGs in *D. asperoides*.

The GO functional classification analysis provided useful and potential contributors to the regulatory networks in the transcriptome (Figure 4A). All 3916 DEGs were subjected to GO

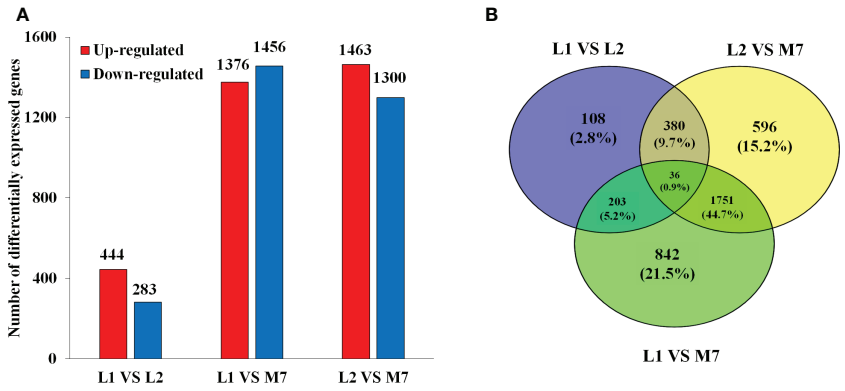


FIGURE 3
Differentially expressed genes (DEGs) in the root of *D. asperoides*. **(A)** The number of DEGs up- or down-regulated at L1, L2 and M7. **(B)** Venn diagram for the number of DEGs at different points in root of *D. asperoides*.

functional classification analysis to calculate the functional category based on the biological process. Most of the DEGs were assigned to 14 biological processes. Genes associated with isoprenoid biosynthetic process and isoprenoid metabolic process participated in terpenoid metabolism. Regulation of jasmonic acid mediated signaling pathway (1%), jasmonic acid metabolic process (1%), and jasmonic acid biosynthetic process (1%) were involved in the plant response to JA and abiotic stresses. The 36 DEGs that are common in all three categories enriched in 6 functional categories based on the biological process (Table S4). The most DEGs are response to light and may be the main factor for plant growth and development. And one of them had terpene synthase activity. 108 DEGs unique expressed in L2, including 10 genes were enriched in terpenoid biosynthetic and metabolic, 2 genes were enriched in jasmonic acid signaling. They may involve in regulation and biosynthesis of asperosaponin VI (Figure S3A). Therefore, the DEGs unique expressed in M7 compared with L1, approximately 103 genes enriched in cell wall metabolic, polysaccharide, xyloglucan metabolic, may relate to root development (Figure S3B). Moreover, the intersection between L1, M7 and L2, M7 was enriched most DEGs. Among the 1751 DEGs, 8 genes were enriched in terpene biosynthetic and metabolic, 12 genes were enriched in jasmonic acid signaling (Figure S3C).

The KEGG pathway analyses were performed for further functional enrichment analysis (Figure 4B). All of the DEGs from the three developmental stages were assigned to 18 KEGG pathways. Major pathways related to terpenoid metabolism included tropane, piperidine and pyridine alkaloid biosynthesis (15), sesquiterpenoid and triterpenoid biosynthesis (12), diterpenoid biosynthesis (6), terpenoid backbone biosynthesis (16). Several pathways were associated with phenylpropane metabolism, such as phenylpropanoid biosynthesis (47), flavonoid biosynthesis (13), and phenylalanine metabolism

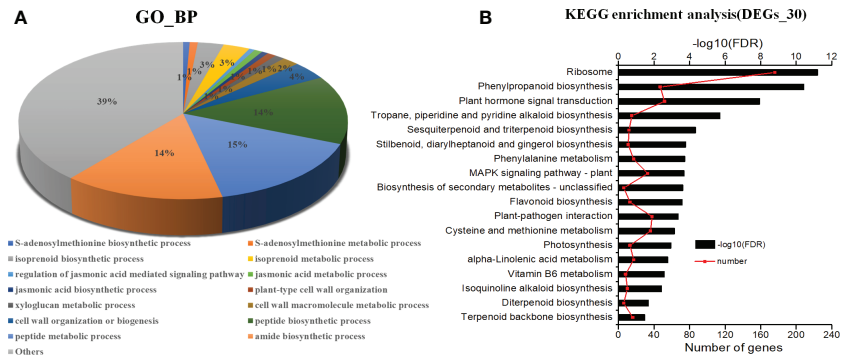


FIGURE 4
Analysis of KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways and GO (Gene Ontology) enrichments. **(A)** KEGG pathways (P <0.001). **(B)** GO enrichments (biological processes, FDR<0.001). The data are for the functional classification of all DEGs.

(17). There were several pathways related to JA metabolic pathway, such as plant hormone signal transduction (52), and α-linolenic acid metabolism (17). The results revealed that JA signal may play an important role in the biosynthesis of asperosaponin.

Genes potentially involved in asperosaponin VI biosynthesis

Combined with GO enrichment analyses and KEGG pathways, the genes related to asperosaponin VI biosynthesis, including transcription factors, plant hormone signal transduction pathways, α-linolenic acid metabolism, JA biosynthesis and signaling, and triterpenoid biosynthesis were identified.

Terpenoid biosynthesis

In our study, we found that most genes involved in the pathways of sesquiterpenoid and triterpenoid biosynthesis, and terpenoid backbone biosynthesis showed significant differences expression (Figure 5A). In our DEG data, 9 genes involved in terpenoid backbone biosynthesis, including acetyl-CoA acetyltransferase (*DaAACT*), 3-hydroxy-3- methylglutaryl coenzyme A reductase (*DaHMGCR*), 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (*DaMDS*), and mevalonate kinase (*DaMK*), were up-regulated during the developmental stages of the roots (Figure 5B). Moreover, two alpha-farnesene synthase (*DaAFS*) genes related to sesquiterpenoid biosynthesis were down-regulated. In the pathway of triterpenoid saponins biosynthesis, three *DaDS* genes involved in dammarane-type triterpene saponin biosynthase were down-regulated. Furthermore, the genes in the biosynthesis of asperosaponin VI were up-regulated, including squalene epoxidase (*SE*), (Figure 5B). To confirm the results of transcriptomic analysis, we identified the expression of several terpenoid biosynthesis genes by qRT-PCR. The expression of Da*AACT, DaGPPS, DaSE* and *DaSE1* in the roots of M7 was higher than in L1. The expression of *DaHMGCS*, *DaHMGCR-2* was up-regulated significantly in L2 and M7, compared with L1. Whereas *DaAFS*, *DaDS-1* and *DaDS-2* transcripts were inhibited in M7 (Figure 5C). These results suggested that asperosaponin VI

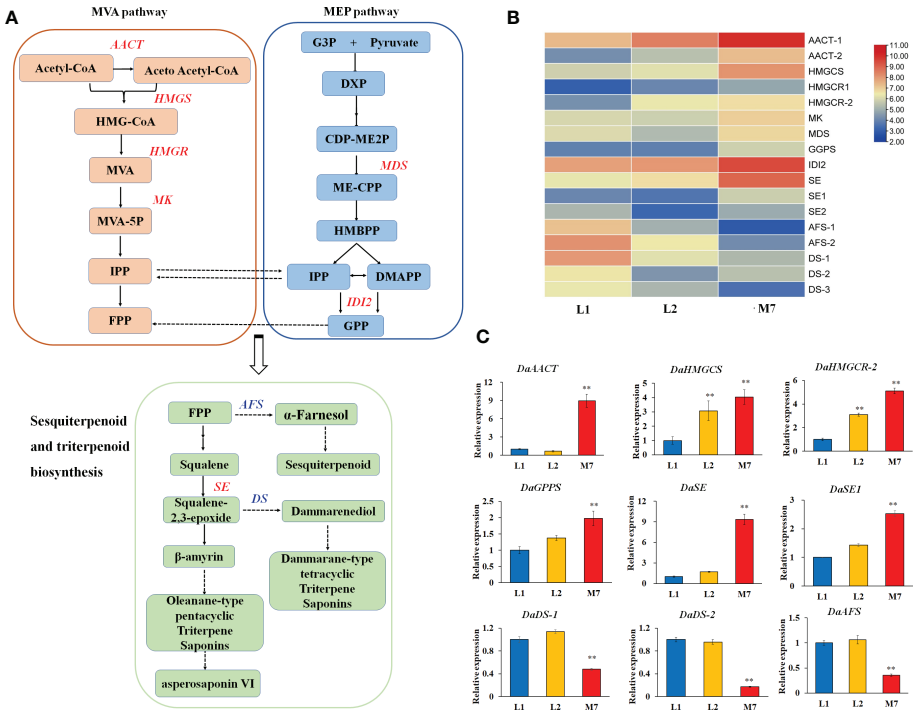


FIGURE 5
The biosynthesis of sesquiterpenoid and triterpenoid at different developmental stages. **(A)** The biosynthesis pathway of sesquiterpenoid and triterpenoid was produced according to the KEGG database. **(B)** The heat map of genes associated with terpenoid backbone, sesquiterpenoid and triterpenoid biosynthesis based on their expression level. **(C)** The expression of *DaAACT, DaHMGCS, DaHMGCR, DaGPPS, DaSE, DaSE1, DaDS-1,DaDS-2, DaAFS* in the pathway of A during root development. The values are the means ± SD; n = 3. Statistical analyses were performed using Student's t test compared with L1. **, P < 0.01. AACT, Acetyl-CoA acetyltransferase; SE, squalene epoxidase; DS, dammarenediol synthase; AFS, alpha-farnesene synthase.

biosynthesis pathways were activated during root development. In addition, the biosynthesis of sesquiterpenoid and dammarane-type triterpene saponins was inhibited in the roots of *D. asperoides*.

α-Linolenic acid metabolism, JA biosynthesis and signaling

The release of α-linolenic acid through the hydrolysis of the chloroplast membranes by phospholipases initiates JA biosynthesis (Scherer et al., 2010). In this study, several α-linolenic acid metabolism genes, JA biosynthesis and signaling related genes were differentially expressed during root development in *D. asperoides* (Figure 6A). A gene heatmap was illustrated based on the RNA-Seq analysis (Figure 6B). Genes related to α-linolenic acid metabolism and JA biosynthesis, such as two phospholipase genes (*DaPLA*), two allene oxide synthase (*DaAOS*), two allene oxide cyclase genes (*DaAOC*), and 12-oxophytodienoic acid reductase genes (*DaOPR3*, *DaOPR2*), were up-regulated during root development. JA negative regulator *JAZ* (*DaTIFY11A*, *DaTIFY11B*) and *DaMYC2* were down-regulated. Moreover, the concentration of JA was gradually increased during root development in *D. asperoides* (Figure 6C).

To further validate the expression data, we verified the expression of α-linolenic acid metabolism and JA biosynthesis genes by qRT-PCR. The expression of *DaPLA2A*, *DaAOS2* and *DaOPR3* was strongly up-regulated in the root of M7 compared with L1.And the gene expression of *DaAOS1*, *DaAOC1* in the root of L2, M7 was significantly higher than in the root of L1. In addition, JA signaling related genes *DaTIFY11A*, *DaTIFY11B*, and *DaMYC2* was significantly down-regulated expression in M7. (Figure 6D). The results illustrated that the α-linolenic acid metabolism and JA biosynthesis and signaling was activated during root development in *D. asperoides*. Additionally, we speculated that JA biosynthesis and signaling regulated the biosynthesis of asperosaponin VI in *D. asperoides*.

Plant hormone signal transduction

Hormones are important in plant growth, development, and the biosynthesis of secondary metabolites. Genes involved in plant hormone signaling pathways were identified in different developmental stage roots in *D. asperoides*, including auxin,

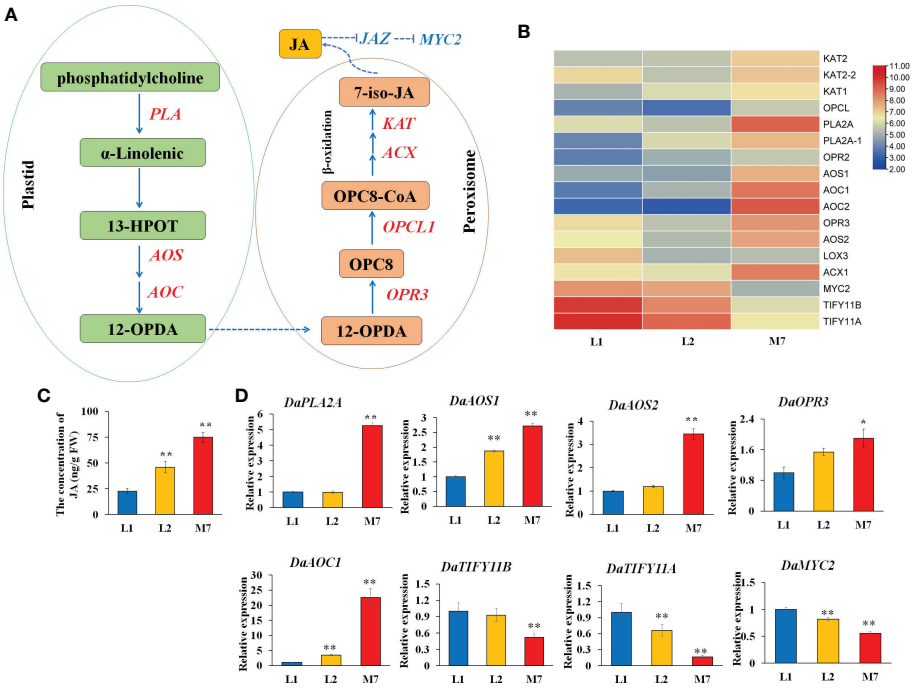


FIGURE 6
The JA biosynthesis at different developmental stages. **(A)** The pathway of α-linolenic acid metabolism and JA biosynthesis according to the KEGG database. **(B)** The heat map of genes associated with α-linolenic acid metabolism and JA biosynthesis based on their expression level. **(C)**.The concentration of JA during root development. **(D)** The expression of genes *DaPLA2A, DaAOS1, DaAOS2, DaOPR3, DaAOC1, DaTIFY11B, DaTIFY11A*, and *DaMYC2* during root development. AOS, allene oxide synthase. The values of **(C, D)** are the means ± SD; n = 3. Statistical analyses were performed using Student's t test compared with L1. **, P < 0.01, *, P < 0.05. OPR3, 12-oxophytodienoic acid reductase; PLA2A, Phospholipase A2A.

ethylene, JA, cytokinin, ABA, gibberellic acid (GA) and salicylic acid (SA) (Figure S4, Table S5). All hormone signaling pathways were predominantly induced. In IAA signaling pathways, the IAA receptor gene TIR1 and most small auxin-up RNA (SAUR) were up-regulated, while some auxin-responsive protein/indoleacetic acid-induced protein (Aux/IAA) were down-regulated during root development. Many DEGs in the JA signaling pathway, including a MYC-related gene and three TIFY genes were differentially expressed during root development. In the ABA signaling pathways, four ABA receptor genes (PYL) were up-regulated, while three serine/threonine-protein kinase SRK2 (snrk2) genes and two protein phosphatase 2C (PP2C) genes were down-regulated. Three DELLA proteins, three GID1 genes and one GID2 were up-regulated in the GA signaling pathway. The results suggested that plant hormones played complex roles in the biosynthesis of asperosaponin VI and roots development in *D. asperoides*.

Transcription factors

Transcription factors play an important role in regulating metabolite synthesis and metabolism, as well as regulating gene expression. This study focused on the TFs in the DEGs. A total of 146 DEGs encoding TFs were identified during different developmental stages in the roots, which belonged to 22 TF families (Table S6). The number of down-regulated genes was significantly greater than the up-regulated genes. The results showed that 18.6% (27 in 146) of total MYBs were down-regulated in the developmental stages of the roots, as well as 5.5% (8 in 146) WRKY, and 5.5% (9 in 146) NAC TFs. About 24.1% (35 in 146) of the ethylene-responsive TFs (ERF/AP2) were differentially expressed (Table S5). One MYB and 1 bHLH were unique expressed in L2. And 36 TFs were unique expressed in M7, including 5 NAC, 4 AP2/ERF, 6 MYB and 3 bHLH. Expression correlation analysis indicated that 34 TFs involved in terpenoid biosynthesis, including 11 AP2/ERF and 4 B3_superfamily (Figure S5, Table S8). qRT-PCR revealed that the expression of *DaERF1* was up-regulated during root development, especially in stage M7 (Figure S6).

JA biosynthesis and signaling pathway enhanced the biosynthesis of asperosaponin VI in *D. asperoides*

To confirm whether the JA biosynthesis and signaling plays important role in the biosynthesis of asperosaponin VI in D. asperoides, which was subjected to MeJA treatment. The results showed that the content of asperosaponin VI was significantly accumulated following MeJA treatment in the roots of *D. asperoides* (Figure 7A). Moreover, the terpenoid biosynthesis gene expression of *DaAACT, DaHMGCS*, and *DaHMGCR-2* was also up-regulated after MeJA treatment for 6 h (Figures 7B–D). In addition, the expression of *DaOPR3*, and *DaAOS1* was up-regulated following MeJA treatment for 6 h (Figures 7E, F). Furthermore, the expression of *DaDS* was down-regulated. And expression level of *DaAFS* was increased following MeJA

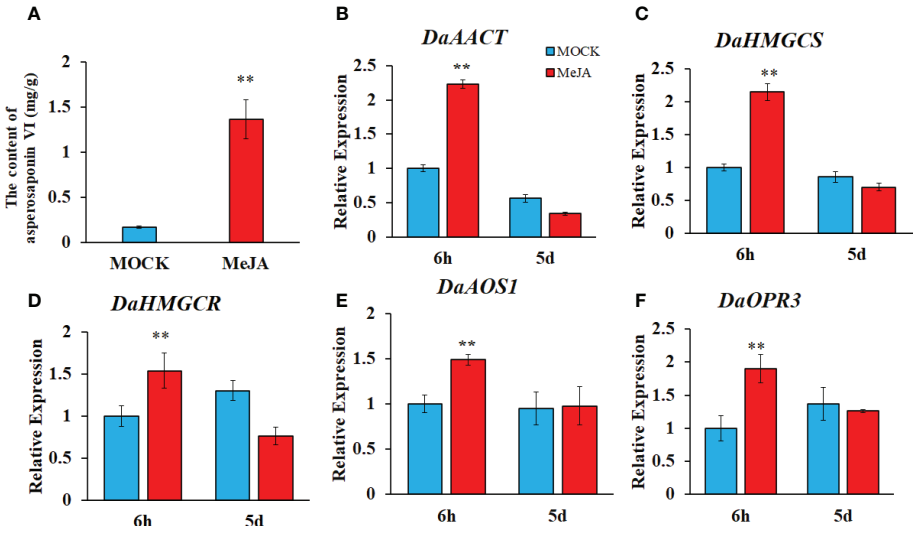


FIGURE 7
MeJA regulated the biosynthesis of asperosaponin VI in *D. asperoides*. **(A)** the content of asperosaponin VI in the roots following MeJA treatment for 5 d. **(B–F)** the expression of *DaAACT, DaOPR3, DaAOS1, DaHMGCS, DaHMGCR* in the roots following MeJA treatment for 6 h and 5d. All the values are the means ± SD; n = 3. Statistical analyses were performed using Student's t test compared MOCK. **, P < 0.01.

treatment (Figure S7). Hence, the JA biosynthesis pathway was activated. The results illustrated that JA biosynthesis and signaling regulated the expression of *DaAACT, DaHMGCS, DaHMGCR-2* and accelerated the biosynthesis of asperosaponin VI.

Discussion

Saponins are important secondary metabolites in Chinese herbal plants, especially triterpenoid saponins. Significant difference in content of asperosaponin VI is due to different growth periods (Jiang et al., 2013). Exploring the distribution of triterpene saponins and understanding the molecular and physiological mechanisms of the synthesis and metabolism of triterpene saponins are essential in *D. asperoides*. In this study, MALDI-MSI and transcriptomics were used to analyze the distribution and mechanisms of synthesis of triterpene saponins in *D. asperoides*. The results revealed a complex gene network involving JA biosynthesis pathways and other regulatory factors involved in the biosynthesis of triterpene saponins.

Root growth and development resulted in differences in the biosynthesis and distribution of asperosaponin VI

Secondary metabolites participate in plant growth, development, and adaptation, and may in addition play important roles in plant defense against pathogens and abiotic stress (Bednarek and Osbourn, 2009; Peng et al., 2021). Triterpenoid saponins are known as a vital class of secondary metabolites with important medicinal value (Ge et al., 2016; Yan et al., 2018). A few histochemical staining methods have been applied for the visualization of the distribution of various triterpenoid saponins, which play important roles in the efficacy of traditional Chinese medicine (Kubo et al., 1981). However, it is noteworthy that these staining methods are unable to distinguish different types of triterpenoid saponins, especially asperosaponin VI in *D. asperoides*. In this study, asperosaponin VI was visualized with MALDI MSI for the first time. Asperosaponin VI was distributed in the xylem and was enriched throughout the entire xylem from the two-leaf stage (Figure 2H). Thus, we speculate that the content of asperosaponin VI increased because of the increase of xylem proportion in roots during root development.

Complex gene networks involved in the biosynthesis of asperosaponin VI

Complex secondary metabolic regulation processes exist in Chinese herbal plants. Genomes have been used to elucidate the metabolic processes in some important plants, such as *Panax ginseng* Xu et al., 2017 *Panax notoginseng* (Wei et al., 2017; Chen et al., 2018; Jiang et al., 2020), and *Tripterygium wilfordii* (Tu et al., 2020). With the rapid development of sequencing technology, transcriptome sequencing has emerged as a rapid method for researching the regulation and molecular mechanisms of complex growth and development processes, especially for *D. asperoides*, which lacks a reference genome. In this study, we identified 3916 DEGs (Figure 3A), which were enriched in the pathways of sesquiterpenoid and triterpenoid biosynthesis, terpenoid backbone biosynthesis, α-linolenic acid metabolism and plant hormone signal transduction by KEGG pathway analysis (Figure 4B). Herein, the DEGs were abundant in jasmonic acid metabolic process and jasmonic acid biosynthetic process (Figure 4A). Moreover, the unique expressed genes in L2 stage enriched in terpenoid biosynthetic and metabolic, jasmonic acid signaling (Figure S3A). The results indicated that JA biosynthesis and signaling, sesquiterpenoid and triterpenoid biosynthesis genes, terpenoid backbone biosynthesis genes, play an important role in regulating the biosynthesis of asperosaponin VI.

TFs involved in JA biosynthesis and signaling and the biosynthesis of asperosaponin VI

Transcription factors play important roles in regulating plant growth and metabolite synthesis by regulating the expression of downstream genes (Hoang et al., 2017; Long et al., 2019). In this study, 146 TFs exhibited significant differences in expression. These TFs were classified into 20 families, including *ERF/AP2*, *MYB*, *NAC*, and *WRKY*, which indicated the regulation of metabolites synthesis in *D. asperoides* (Tables S6, S7). Expression correlation analysis that 11 AP2/ERF involved in terpenoid biosynthesis (Figure S5, Table S8). Previous studies have demonstrated that TFs function in regulating terpenoid synthesis. Tanshinone production was significantly increased in *SmERF1L1* overexpressed roots in *Salvia miltiorrhiza*. Additionally, *SmERF1L1*, a novel JA -responsive gene, positively regulated tanshinone biosynthesis by comprehensively upregulating tanshinone biosynthetic pathway genes (Huang et al., 2019). Moreover, the over-expression of *SmWRKY1* significantly increased tanshinone production in *S. miltiorrhiza*. *SmWRKY1* was responsive to MeJA and acted as a positive regulator to regulate tanshinone biosynthesis through activating *SmDXR* in the MEP pathway (Cao et al., 2018). In addition, a *R2R3-MYB* TF, *SmMYB98*, promoted the accumulation of tanshinone by activating the transcription of *SmGGPPS1*, *SmPAL1*, and *SmRAS1* (Hao et al., 2020). Therefore, many TFs may be regulated by JA biosynthesis and signaling. And studies on TFs will be

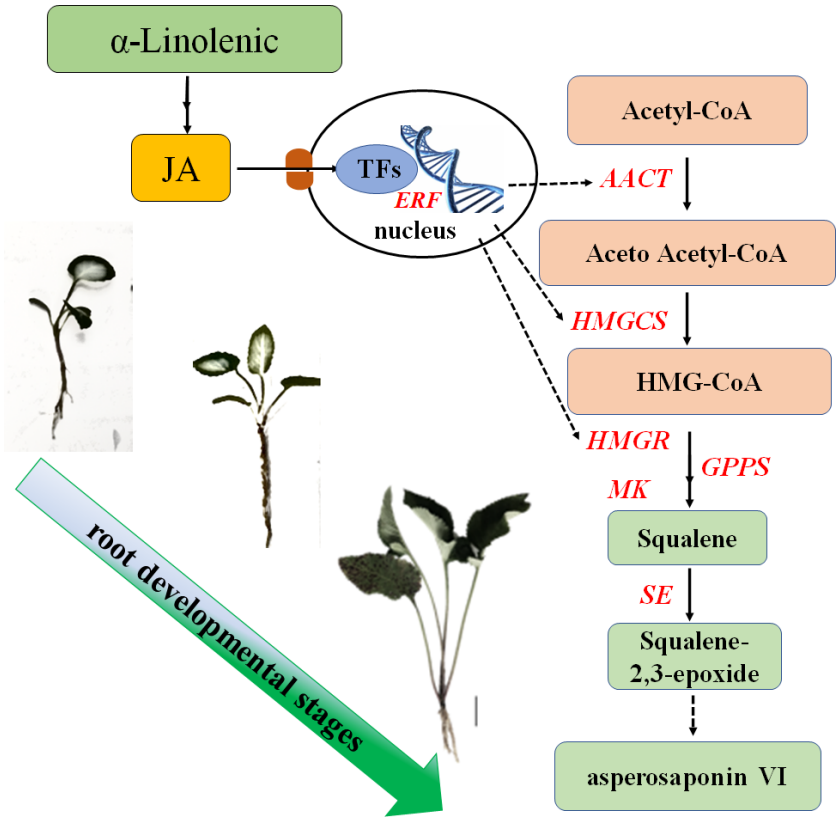


FIGURE 8
Schematic showing the role of α-linolenic acid and JA in the biosynthesis of asperosaponin VI during the root deveopmental in *D. asperoides*..

beneficial for illustrating the regulatory networks for the biosynthesis of triterpenoids, including asperosaponin VI.

JA biosynthesis and signaling promotes the biosynthesis of asperosaponin VI by activating the expression of triterpenoid biosynthetic genes

The phytohormone JA regulates plant development and defense processes, such as by synthesizing secondary metabolites to resist biotic or abiotic stresses (Browse, 2009; Zhou et al., 2016; Ho et al., 2020). There are some molecular links between JA biosynthesis, signaling and saponin biosynthesis. A JA biosynthetic 13-lipoxygenase gene, *PgLOX6*, promotes the production of ginsenoside, which is a triterpenoid saponin. Furthermore, *PgLOX6* up-regulates the expression of ginsenoside biosynthetic genes such as squalene synthase (*SS1*) and squalene epoxidase (*SE*) (Shadi et al., 2016). The JASMONATE ZIM DOMAINs (JAZs) are repressors in the JA signaling pathway (Bai et al., 2011), and *SmJAZ8* deregulates the yield of tanshinone in "*Danshen*" (Pei et al., 2018). JAs are formed from α -linolenic acid of chloroplast membranes by the lipoxygenase pathway (Wasternack and Song, 2016) (Figure 6A). In this study, the concentration of JA was increased (Figure 6C). Meanwhile, MeJA facilitated the accumulation of asperosaponin VI (Figure 7A). JA biosynthesis and signaling has a positive effect on the regulation of asperosaponin VI biosynthesis. In addition, the gene *DaAACT*, *DaHMGCS, DaHMGCR* in the pathway of saponin biosynthesis was rapidly up-regulated following MeJA treatment (Figures 7B–D), as well as in the developing roots. Hence, we suspect that JA biosynthesis and signaling may regulate the biosynthesis of asperosaponin VI by regulating the expression of triterpenoid biosynthetic genes including *DaAACT*, *DaHMGCS, DaHMGCR*. *DaAACT* is the first enzyme in the terpenoid synthesis pathway and catalyzes two units of acetyl-CoA into acetoacetyl-CoA. The application of AACT could increase the fraction of 3-hydroxyvalerate in *Escherichia coli* (Jeon et al., 2017). Therefore, we speculate that JA biosynthesis and signaling regulates the expression of TFs which active the expression of triterpenoid biosynthetic genes, thereby facilitating the biosynthesis of asperosaponin VI in *D. asperoides*.

Conclusion

In this study, we found that the content of asperosaponin VI and JA in the root increased gradually along with the development of the roots. Additionally, the spatial distribution of asperosaponin VI was localized in the xylem. A total of 3916 DEGs and 146 TFs were identified. The genes associated with JA biosynthesis, signaling and asperosaponin VI biosynthesis were up-regulated following MeJA treatment, as well as in the developing roots of *D. asperoides*. We speculate that JA biosynthesis and signaling regulates the expression of TFs to promote the expression of triterpenoid biosynthetic genes and facilitate the biosynthesis of asperosaponin VI (Figure 8). These results present a regulatory network by which asperosaponin VI, JA, and TFs co-modulate the biosynthesis of asperosaponin VI in *D. asperoides*.

Data availability statement

The original contributions presented in the study are publicly available. The data presented in the study are deposited in the the National Center for Biotechnology Information (NCBI) database, accession number GSE208123.

Author contributions

JX, TZ conceived and designed the study. JX wrote the manuscript. XO, YY, CX, WJ revised the manuscript. JX, ZH and CY performed the experiments. JX, HH and LL contributed to data analysis. All authors discussed the results and agreed to the published version of the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (81860675, 82160725), Key project at central government level: The ability to establish sustainable use of valuable Chinese medicine resources (2060302), Guizhou Provincial Science and Technology Projects (Qian Ke He Ji Chu [2020]1Y370), [2019]1026, the Research Platform Team Project of the Education Department of Guizhou Province [Qian Jiao Ji [2022]021], and the High-level Innovative Talents of Guizhou Province of China (Qian Ke He Ping Tai Ren Cai [2018] 5638-2).

Acknowledgment

We are grateful to Prof. Xiao Wang from the Key Laboratory for Applied Technology of Sophisticated Analytical Instruments of Shandong Province, Shandong Analysis and Test Center, Qilu University of Technology (Shandong Academy of Sciences), for assistance with MALDI-MSI.

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.2022.1022075/full#supplementary-material

References

Ali, M., Hussain, R. M., Rehman, N. U., She, G., Li, P., Wan, X., et al. (2018). *De novo* transcriptome sequencing and metabolite profiling analyses reveal the complex metabolic genes involved in the terpenoid biosynthesis in blue anise sage (*Salvia guaranitica* l.). *DNA Res.* 25, 597–617. doi: 10.1093/dnares/dsy028

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

Bai, Y., Meng, Y., Huang, D., Qi, Y., and Chen, M. (2011). Origin and evolutionary analysis of the plant-specific TIFY transcription factor family. *Genomics* 98, 128–136. doi: 10.1016/j.ygeno.2011.05.002

Bednarek, P., and Osbourn, A. (2009). Plant-microbe interactions: Chemical diversity in plant defense. *Science.* 324, 746–748. doi: 10.1126/science.1171661

Browse, J. (2009). Jasmonate passes muster: a receptor and targets for the defense hormone. *Annu. Rev. Plant Biol.* 60, 183–205. doi: 10.1146/annurev.arplant.043008.092007

Cao, W., Wang, Y., Shi, M., Hao, X., Zhao, W., Wang, Y., et al. (2018). Transcription factor *SmWRKY1* positively promotes the biosynthesis of tanshinones in *Salvia miltiorrhiza*. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00554

Chen, K., Liu, J., Ji, R., Chen, T., Zhou, X., Yang, J., et al. (2018). Biogenic synthesis and spatial distribution of endogenous phytohormones and ginsenosides

provide insights on their intrinsic relevance in *Panax ginseng*. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01951

Fan, H., Li, K., Yao, F., Sun, L., and Liu, Y. (2019). Comparative transcriptome analyses on terpenoids metabolism in field- and mountain-cultivated ginseng roots. *BMC Plant Biol.* 19, 82. doi: 10.1186/s12870-019-1682-5

Garg, A., Agrawal, L., Misra, R. C., Sharma, S., and Ghosh, S. (2015). *Andrographis paniculata* transcriptome provides molecular insights into tissue-specific accumulation of medicinal diterpenes. *BMC Genomics* 16, 659. doi: 10.1186/s12864-015-1864-y

Ge, Z. R., Xu, M. C., Huang, Y. U., Zhang, C. J., Lin, J. E., and Ruan, C. W. (2016). Cardioprotective effect of notoginsenoside R1 in a rabbit lung remote ischemic postconditioning model *via* activation of the TGF-beta1/TAK1 signaling pathway. *Exp. Ther. Med.* 11, 2341–2348. doi: 10.3892/etm.2016.3222

Hao, X., Pu, Z., Cao, G., You, D., Zhou, Y., Deng, C., et al. (2020). Tanshinone and salvianolic acid biosynthesis are regulated by *SmMYB98* in *Salvia miltiorrhiza* hairy roots. *J. Adv. Res.* 23, 1–12. doi: 10.1016/j.jare.2020.01.012

Hoang, X. L. T., Nhi, D. N. H., Thu, N. B. A., Thao, N. P., and Tran, L. P. (2017). Transcription factors and their roles in signal transduction in plants under abiotic stresses. *Curr. Genomics* 18, 483–497. doi: 10.2174/1389202918666170227150057

Ho, T. T., Murthy, H. N., and Park, S. Y. (2020). Methyl jasmonate induced oxidative stress and accumulation of secondary metabolites in plant cell and organ cultures. *Int. J. Mol. Sci.* 21, 716. doi: 10.3390/ijms21030716

Hou, Z. X., and Huang, W. D. (2005). Immunohistochemical localization of IAA and ABP1 in strawberry shoot apexes during floral induction. *Planta* 222, 678–687. doi: 10.1007/s00425-005-0014-1

Hu, Z. P., Xu, J., Zhou, T., Jiang, W. K., Guo, J, Xiao, C. H., et al. (2021). Methyl jasmonate stimulates the synthesis and accumulation of asperosaponin VI in the roots of *Dipsacus asper* [J]. *Acta Pharm. Sin.* 56, 2302–2307. doi: 10.16438/j.0513-4870.2021-0375

Huang, Q., Sun, M., Yuan, T., Wang, Y., Shi, M., Lu, S., et al. (2019). The AP2/ERF transcription factor *SmERF1L1* regulates the biosynthesis of tanshinones and phenolic acids in *Salvia miltiorrhiza*. *Food Chem.* 274, 368–375. doi: 10.1016/j.foodchem.2018.08.119

Jeon, J. M., Kim, H. J., Bhatia, S. K., Sung, C., Seo, H. M., Kim, J. H., et al. (2017). Application of acetyl-CoA acetyltransferase (AtoAD) in *Escherichia coli* to increase 3-hydroxyvalerate fraction in poly(3-hydroxybutyrate-co-3-hydroxyvalerate). *Bioprocess Biosyst. Eng.* 40, 781–789. doi: 10.1007/s00449-017-1743-9

Jiang, W., Ai, Q., Zhou, T., and Jin, Y. (2013). Analysis on geographical distribution trend of asperosaponin vi content in *Dipsacus asper* in guizhou province. *Guizhou Agric. Sci.* 41, 19–22. doi: 1001-3601(2013)08-0450-0019-04

Jiang, Z., Tu, L., Yang, W., Zhang, Y., Hu, T., Ma, B., et al. (2020). The chromosome-level reference genome assembly for *Panax notoginseng* and insights into ginsenoside biosynthesis. *Plant Commun.* 2, 100113. doi: 10.1016/j.xplc.2020.100113

Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982

Kubo, M., Tani, T., Katsuki, T., Ishizaki, K., and Arichi, S. (1981). Histochemistry. I. ginsenosides in ginseng (*Panax ginseng* C. A. Meyer, root). *J. Nat. Prod.* 44, 278–284. doi: 10.1021/np50016a002

Kuzuyama, T. (2002). Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units. *Biosci. Biotechnol. Biochem.* 66, 1619–1627. doi: 10.1271/bbb.66.1619

Liang, Q., Jiao, X., Xiao, C. H., Zhou, T., Jiang, W. K., Wang, H., et al. (2020). Cloning and expression stability analysis of actin, tubulin and GAPDH genes in *Dipsacus asper*. *Chin. Traditional Herbal Drugs* 51, 5571–5578. doi: 10.7501/j.issn.0253-2670.2020.21.022

Li, C., Tian, J., Li, G., Jiang, W., Xing, Y., Hou, J., et al. (2010). Asperosaponin VI protects cardiac myocytes from hypoxia-induced apoptosis *via* activation of the PI3K/Akt and CREB pathways. *Eur. J. Pharmacol.* 649, 100–107. doi: 10.1016/j.ejphar.2010.08.060

Liu, H., Li, X., Xiao, J., and Wang, S. (2012). A convenient method for simultaneous quantification of multiple phytohormones and metabolites: Application in study of rice-bacterium interaction. *Plant Methods* 8, 2. doi: 10.1186/1746-4811-8-2

Long, L., Yang, W., Liao, P., Guo, Y., Kumar, A., and Gao, W. (2019). Transcriptome analysis reveals differentially expressed ERF transcription factors associated with salt response in cotton. *Plant Sci.* 281, 72–81. doi: 10.1016/j.plantsci.2019.01.012

Mishra, B. S., Singh, M., Aggrawal, P., and Laxmi, A. (2009). Glucose and auxin signaling interaction in controlling *Arabidopsis thaliana* seedlings root growth and development. *PLoS One* 4 (2), e4502. doi: 10.1371/journal.pone.0004502

Niu, Y., Li, Y., Huang, H., Kong, X., Zhang, R., Liu, L., et al. (2011). Asperosaponin VI, a saponin component from *Dipsacus asper* wall, induces osteoblast differentiation through bone morphogenetic protein-2/p38 and extracellular signal-regulated kinase 1/2 pathway. *Phytother. Res.* 25, 1700–1706. doi: 10.1002/ptr.3414

Niu, Y. B., Li, Y. H., Kong, X. H., Zhang, R., Sun, Y., Li, Q., et al. (2012). The beneficial effect of radix dipsaci total saponins on bone metabolism *in vitro* and *in vivo* and the possible mechanisms of action. *Osteoporos Int.* 23, 2649–2660. doi: 10.1007/s00198-012-1932-y

Niu, Y., Li, C., Pan, Y., Li, Y., Kong, X., Wang, S., et al. (2015). Treatment of Radix Dipsaci extract prevents long bone loss induced by modeled microgravity in hindlimb unloading rats. *Pharm. Biol.* 53, 110–116. doi: 10.3109/13880209.2014.911920

Niu, Y., Luo, H., Sun, C., Yang, T. J., Dong, L., Huang, L., et al. (2014). Expression profiling of the triterpene saponin biosynthesis genes FPS, SS, SE, and DS in the medicinal plant *Panax notoginseng*. *Gene.* 533, 295–303. doi: 10.1016/j.gene.2013.09.045

Pei, T. L., Ma, P. D., Ding, K., Liu, S. J., Jia, Y. Y., Mei, Ru, et al. (2018). SmJAZ8 acts as a core repressor regulating JA-induced biosynthesis of salvianolic acids and tanshinones in *Salvia miltiorrhiza* hairy roots. *J. Exp. Bot.* 69, 1663–1678. doi: 10.1093/jxb/erx484

Peng, Z., Wang, Y., Zuo, W. T., Gao, Y. R., Li, R. Z., Yu, C. X., et al. (2021). Integration of metabolome and transcriptome studies reveals flavonoids, abscisic acid, and nitric oxide comodulating the freezing tolerance in *Liriope spicata*. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.764625

Ren, Y, Yu, G, Shi, C, Liu, L, Guo, Q, Han, C, et al (2022). Majorbio Cloud: A one-stop, comprehensive bioinformatic platform for multiomics analyses. *iMeta* 1, e12. doi: 10.1002/imt2.12

Scherer, G. F., Ryu, S. B., Wang, X., Matos, A. R., and Heitz, T. (2010). Patatin-related phospholipase a: Nomenclature, subfamilies and functions in plants. *Trends Plant Sci.* 15, 693–700. doi: 10.1016/j.tplants.2010.09.005

Shadi, R., Yu-Jin, K., Johan, S., Dabing, Z., and Deok-Chun, Y. (2016). *PgLOX6* encoding a lipoxygenase contributes to jasmonic acid biosynthesis and ginsenoside production in *Panax ginseng*. *J. Exp. Bot.* 67, 6007–6019. doi: 10.1093/jxb/erw358

Sun, C., Ma, S., Li, L., Wang, D., Liu, W., Liu, F., et al. (2021). Visualizing the distributions and spatiotemporal changes of metabolites in *Panax notoginseng* by MALDI mass spectrometry imaging. *J. Ginseng Res.* 45, 726–733. doi: 10.1016/j.jgr.2021.04.001

Tu, L., Su, P., Zhang, Z., Gao, L., Wang, J., Hu, T., et al. (2020). Genome of *Tripterygium wilfordii* and identification of cytochrome P450 involved in triptolide biosynthesis. *Nat. Commun.* 11, 971. doi: 10.1038/s41467-020-14776-1

Wang, J. Y., Liang, Y. L., Hai, M. R., Chen, J. W., Gao, Z. J., Hu, Q. Q., et al. (2016). Genome-wide transcriptional excavation of *Dipsacus asperoides* unmasked both cryptic asperosaponin biosynthetic genes and SSR markers. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00339

Wasternack, C., and Song, S. (2016). Jasmonates: biosynthesis, metabolism, and signaling by proteins activating and repressing transcription. *J. Exp. Bot.* 68, 1303–1321. doi: 10.1093/jxb/erw443

Wei, C., Kui, L., Zhang, G., Zhu, S., Zhang, J., Wang, X., et al. (2017). Whole-genome sequencing and analysis of the chinese herbal plant *Panax notoginseng*. *Mol. Plant* 10, 899–902. doi: 10.1016/j.molp.2017.02.010

Xiao, C., Zhou, T., and Jiang, W. (2018). Phenotypic diversity and characteristic analysis of cultivated *Dipsacus asper* from guizhou province. *Mol. Plant Breeding.* 16, 6509–6516. doi: 10.13271/j.mpb.016.006509

Xu, J., Chu, Y., Liao, B., Xiao, S., Yin, Q., Bai, R., et al. (2017). *Panax ginseng* genome examination for ginsenoside biosynthesis. *Gigascience*. 6, 1–15. doi: 10.1093/gigascience/gix093

Yang, Y., Ge, F., Sun, Y., Liu, D., and Chen, C. (2017). Strengthening triterpene saponins biosynthesis by over-expression of farnesyl pyrophosphate synthase gene and RNA interference of cycloartenol synthase gene in *Panax notoginseng* cells. *Molecules.* 22, 581. doi: 10.3390/molecules22040581

Yang, J. L., Hu, Z. F., Zhang, T. T., Gu, A. D., Gong, T., and Zhu, P. (2018). Progress on the studies of the key enzymes of ginsenoside biosynthesis. *Molecules.* 23, 589. doi: 10.3390/molecules23030589

Yang, Z., Liu, G., Zhang, G., Yan, J., Dong, Y., Lu, Y., et al. (2021). The chromosome-scale high-quality genome assembly of *Panax notoginseng* provides insight into dencichine biosynthesis. *Plant Biotechnol. J.* 19, 869–871. doi: 10.1111/pbi.13558

Yan, Y. T., Li, S. D., Li, C., Xiong, Y. X., Lu, X. H., Zhou, X. F., et al. (2018). *Panax notoginsenoside* saponins Rb1 regulates the expressions of akt/ mTOR/PTEN signals in the hippocampus after focal cerebral ischemia in rats. *Behav. Brain Res.* 345, 83–92. doi: 10.1016/j.bbr.2018.02.037

Yu, X., Wang, L. N., Ma, L., You, R., Cui, R., Ji, D., et al. (2012). Akebia saponin d attenuates ibotenic acid-induced cognitive deficits and pro-apoptotic response in rats: Involvement of MAPK signal pathway. *Pharmacol. Biochem. Be.* 101, 479–486. doi: 10.1016/j.bbr.2012.07.045

Zhou, Y., Sun, W., Chen, J., Tan, H., Xiao, Y., Li, Q., et al. (2016). *SmMYC2a* and *SmMYC2b* played similar but irreplaceable roles in regulating the biosynthesis of tanshinones and phenolic acids in *Salvia miltiorrhiza*. *Sci. Rep.* 6, 22852. doi: 10.1038/srep22852

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 03 February 2023
DOI 10.3389/fpls.2023.1125158

OPEN ACCESS

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

REVIEWED BY
Shihong Luo,
Shenyang Agricultural University, China
Lucas Busta,
University of Minnesota Duluth,
United States

*CORRESPONDENCE
Thu-Thuy T. Dang
thuy.dang@ubc.ca

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 15 December 2022
ACCEPTED 12 January 2023
PUBLISHED 03 February 2023

CITATION
Nguyen T-AM, Grzech D, Chung K, Xia Z, Nguyen T-D and Dang T-TT (2023) Discovery of a cytochrome P450 enzyme catalyzing the formation of spirooxindole alkaloid scaffold.
Front. Plant Sci. 14:1125158.
doi: 10.3389/fpls.2023.1125158

Discovery of a cytochrome P450 enzyme catalyzing the formation of spirooxindole alkaloid scaffold

Tuan-Anh M. Nguyen[1], Dagny Grzech[2], Khoa Chung[3], Zhicheng Xia[4], Trinh-Don Nguyen[1] and Thu-Thuy T. Dang[1]*

[1]Department of Chemistry, Irving K. Barber Faculty of Science, University of British Columbia, Kelowna, BC, Canada, [2]Department of Natural Product Biosynthesis, Max Planck Institute for Chemical Ecology, Jena, Germany, [3]Chemistry Research Laboratory, University of Oxford, Oxford, United Kingdom, [4]Department of Chemistry, Faculty of Science, University of British Columbia, Vancouver, BC, Canada

Spirooxindole alkaloids feature a unique scaffold of an oxindole ring sharing an atom with a heterocyclic moiety. These compounds display an extensive range of biological activities such as anticancer, antibiotics, and anti-hypertension. Despite their structural and functional significance, the establishment and rationale of the spirooxindole scaffold biosynthesis are yet to be elucidated. Herein, we report the discovery and characterization of a cytochrome P450 enzyme from kratom (*Mitragyna speciosa*) responsible for the formation of the spirooxindole alkaloids 3-*epi*-corynoxeine (3*R*, 7*R*) and isocorynoxeine (3*S*, 7*S*) from the corynanthe-type (3*R*)-secoyohimbane precursors. Expression of the newly discovered enzyme in *Saccharomyces cerevisiae* yeast allows for the efficient *in vivo* and *in vitro* production of spirooxindoles. This discovery highlights the versatility of plant cytochrome P450 enzymes in building unusual alkaloid scaffolds and opens a gateway to access the prestigious spirooxindole pharmacophore and its derivatives.

KEYWORDS

kratom, cytochrome P450, spirooxindole, CYP71, secoyohimbane

1 Introduction

Spirooxindole alkaloids constitute a subclass of monoterpenoid indole alkaloids (MIAs) with a substituted carbonyl group at the C-2 position in the indole ring (Figure 1A). Since the first isolation of a spirooxindole alkaloid from the root of yellow jessamine (*Gelsemium sempervirens*) in 1870, many spirooxindole alkaloids have been reported from various plant genera, including *Mitragyna*, *Rauwolfia*, and *Vinca* (Bindra, 1973) (Supplementary Figure 1). The majority of spirooxindole alkaloids feature the unique spirooxindole scaffold in which the oxindole ring shares a single atom at the C-3 position with a cycloalkyl or a heterocyclic moiety (Zhou et al., 2020) derived from monoterpenoid indole alkaloids biosynthesis. As a valuable pharmacophore, spirooxindoles have recently attracted significant attention from chemists and biochemists for their diverse range of bioactivities. Examples include the tetracyclic corynoxeine and isocorynoxeine used in treating hypertension and stroke (Zhao

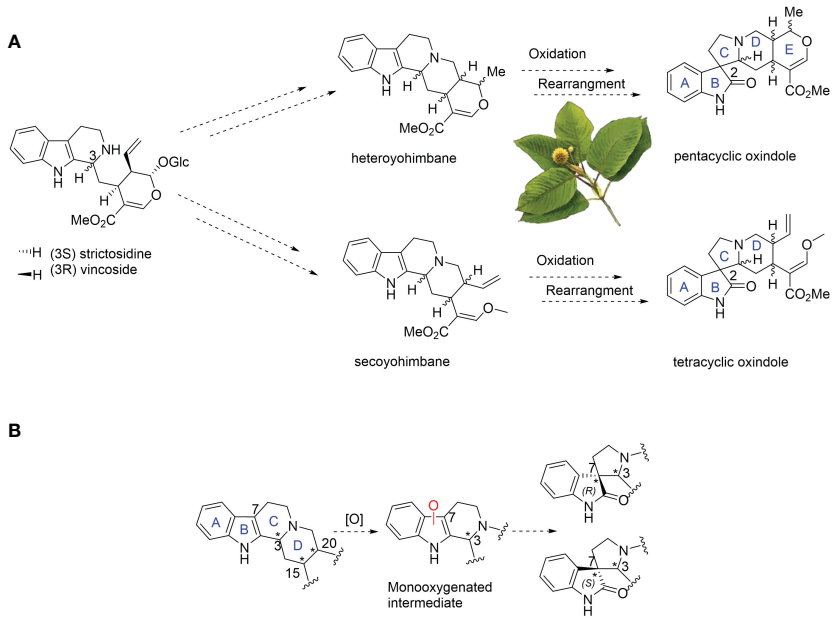
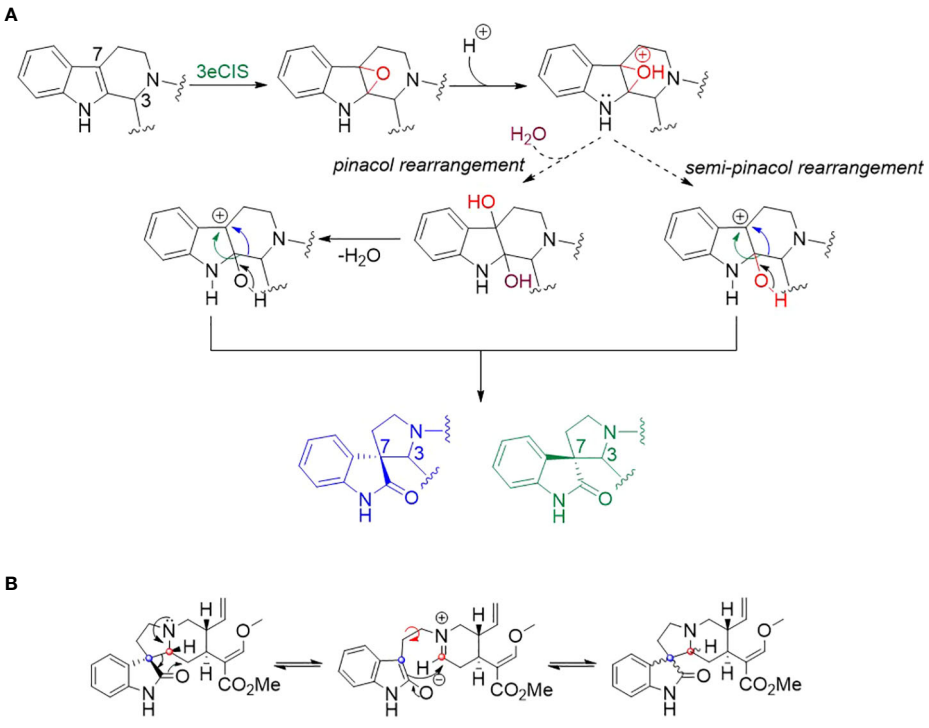


FIGURE 1
(A) Proposed biosynthesis of spirooxindole alkaloids from monoterpenoid indole alkaloid precursors in kratom (*Mitragyna speciosa*). The dotted arrows represent the unknown enzymatic steps. **(B)** Proposed oxidative rearrangement in the formation of spirooxindole alkaloids.

et al., 2016), corynoxine and its isomer corynoxine B as potential agents to treat Parkinson's disease (Chen et al., 2021), and mitraphylline with promising anti-tumour activity (Supplementary Figure 1) (García Giménez et al., 2010). Kratom (*Mitragyna speciosa*) and cat's claw (*Uncaria rhynchophylla*) from the plant family Rubiaceae are well known for their spirooxindole alkaloid contents and thus have been the focus in studying spirooxindole biosynthesis. However, the low abundance of spirooxindole alkaloids in these plants (Manwill et al., 2022) makes biosynthetic elucidation a formidable task. Typically, spirooxindole alkaloids occur in pairs of



SCHEME 1
(A) Proposed mechanism for the formation of spirooxindole *via* (semi-)pinacol rearrangement. **(B)** The isomerization of spirooxindoles *via* intramolecular Mannich reactions (Laus et al., 1996; Flores-Bocanegra et al., 2020). The red and blue dots correspond to C-3 and C-7, respectively.

interconvertible stereoisomers as the C-3–C-7 bond of the *p*-aminolactam group is prone to cleavage and reformation (Figure 1B) (Ahmad and Salim, 2015). The distinct stereogenic centers of the polycyclic scaffolds are the key feature contributing to the diversity of spirooxindole structures. Generally, the known spirooxindoles from *M. speciosa* are classified into two main structural types: secoyohimbane-type/tetracyclic and heteroyohimbane-type/pentacyclic structures (Figure 1A) (Bindra, 1973). While the underlying biochemistry is unknown, the spirooxindole group formation is speculated to be the result of an oxidative rearrangement at C-7 and C-3 positions in the tetrahydro-β-carboline moiety of the indole precursor in *M.* species (Figure 1B) (Shavel and Zinnes, 1962). Recent isotopic labelling studies supported that spirooxindole alkaloids could be generated in a one-step oxidative rearrangement from the tetrahydro-β-carboline moiety of seco- and hetero-yohimbine-type alkaloids such as corynantheine methyl ether and ajmalicine (Lopes et al., 2019). Among oxidative enzymes, cytochrome P450 enzymes (CYPs) are ubiquitous in plant specialized metabolism (Nguyen and Dang, 2021), especially members of the CYP71 family are key drivers of MIA diversification from simple seco- and hetero-yohimbine to various scaffolds, including sarpagan (Dang et al., 2018), strychnos (Tatsis et al., 2017; Hong et al., 2022; Wang et al., 2022), akuammilan (Wang et al., 2022), iboga (Farrow et al., 2019), and aspidosperma (Caputi et al., 2018; Qu et al., 2018). Therefore, we hypothesized that a CYP71 catalyzes the oxidative rearrangement of tetrahydro-β-carbolines to spirooxindoles. Using available *M. speciosa* transcriptome and genome (Brose et al., 2021) and OrthoFinder (Emms and Kelly, 2015), we identified and characterized a CYP71 enzyme that converts a secoyohimbine scaffold to a spirooxindole scaffold. This discovery opens a window into the largely unknown biosynthesis of spirooxindole alkaloids and offers a pioneering biocatalyst for sustainable synthetic routes of spirooxindoles from the tetrahydro-β-carboline scaffold.

2 Materials and methods

2.1 Identification, cloning of candidates and protein expression

Transcriptomes of spirooxindole alkaloids-producing plants (*M. speciosa*, *Rauwolfia serpentina*, and *G. sempervirens*), MIAs-producing but spirooxindole alkaloids-free plants (*Camptotheca acuminata*, *Amsonia hubrichtii*, *Cinchona ledgeriana*, *Nothapodytes nimmoniana*, *Ophiorrhiza pumila*, and *Catharanthus roseus*) and a MIAs-free plant (*Arabidopsis thaliana*) are publicly available from the Medicinal Plant Genomic Resources (http://mpgr.uga.edu/), the PhytoMetaSyn database (https://bioinformatics.tugraz.at/phytometasyn/), TAIR10 (https://www.arabidopsis.org/), and previous studies (Rather et al., 2018; Rai et al., 2021). TransDecoder was used to generate the proteomes, which were subsequently subjected to OrthoFinder for orthogroups analysis. Candidates belonging to the CYP71 family that act on the B ring of the indole moiety were selected to test for activities in the spirooxindole scaffold formation. We focused on the orthogroups containing CYP orthologues unique to *M. speciosa* (MsCYP). The phylogenetic tree of the MsCYP candidates with other reported CYPs involved in the biosynthesis of MIAs from different species was constructed by the Geneious Tree Builder program in the Geneious Prime software package (Biomatters). The amino acid sequence alignment of the MsCYP candidates was performed by the Geneious Prime software package (Biomatters). The open reading frames of MsCYP candidates were obtained from the available transcriptome (Brose et al., 2021). The sequences combining overhangs of *Spe*I and *Not*I restriction sites at the multiple cloning site 1 of the pESC-Leu2d plasmid were synthesized by TwistBioscience (CA, USA) (Ro et al., 2008; Nguyen et al., 2021). The constructs were inserted into pESC already containing the required redox partner cytochrome P450 reductase (CPR) (Ro et al., 2008) by using 5X In-Fusion cloning system (Takara Bio USA Inc.). The yeast strain YPL 150 C:PEP4KO was used for heterologous expression of the CYP candidates following the procedure described before (Nguyen et al., 2022).

2.2 Enzyme assays

The *in vivo* CYP activity screening assays were conducted using the established protocol for yeast whole-cell assays in the 96-well plate (Nguyen et al., 2022). Various MIA alkaloid substrates were fed at a final concentration of 10μM into the yeast cultures for 48 h (Supplementary Figure 13). The *in vitro* assays were conducted in different buffers: 1 M citrate pH 4, 1 M citrate pH 5, 1 M HEPES pH 6, 1 M HEPES pH 7, 1 M HEPES pH 8, 1 M Tris pH 9, and 1 M Tris pH 10. The *in vitro* reaction condition was performed with 100 μL of 100 μM buffer, 250 μM NADPH, 10 mg total microsomal protein, and 10 μM hirsuteine at 37 °C for 1 h. The yeast cells containing plasmid without CYP construct were used as empty vector controls. The reaction supernatants were collected by centrifugation and filtration with a 0.2 μm syringe filter (Sartorius). The supernatants were injected to ultra-performance liquid chromatography (UPLC) coupled with a Xevo TQ-S Cronos Triple Quadrupole Mass Spectrometer (MS). All UPLC-MS analyses were conducted on an XBridge BEH XP (50 x 2.1 mm, 1.7 μm) column at a flow rate of 0.6 $mL.min^{-1}$. The column was pre-equilibrated in 90% solvent A (water + 0.1% formic acid), and 10% solvent B (acetonitrile + 0.01% formic acid). The eluting conditions were: 0–8 min, 10–50% B; 8.0–8.5 min, 50–100% B; 8.5–9.5 min, 100% B; and 9.5–11 min, 100–10% B to re-equilibrate the column. Immunoblotting experiment of recombinant MsCYP72056 enzyme was conducted as described before (Nguyen et al., 2021). Steady-state enzyme kinetics was conducted by varying the concentration of hirsuteine substrate from 0 to 300 μM in HEPES pH 7.5, at a fixed concentration of NADPH at 250 μM and analyzed using GraphPad Prism 9.4.1 (GraphPad software).

2.3 Enzymatic product purification and structural elucidation

To obtain enzymatic products at sufficient yields for structural elucidation, we performed multiple *in vitro* assays of MsCYP72056 containing 10 mL 100 μM HEPES pH 7.5, 250 μM NADPH, 10 mg microsomal protein, and 50 μM hirsuteine at 37 °C for 1 h. Reactions were stopped by adding 1 mL of methanol. The *in vitro* assay

supernatants were combined after centrifugation. The crude enzymatic products mixture was extracted from the supernatant by liquid-liquid extraction with chloroform, which was removed *in vacuo* by GeneVac. Concentrated samples were subjected to a Varian semi-preparative HPLC equipped with a Kinetex® 5 µm EVO C18 column (100 Å, 100 × 250 mm) at a flow rate of 1.5 mL.min^{-1}. The column was equilibrated in 90% solvent A (water, 0.1% formic acid) and 10% solvent B (acetonitrile, 0.1% formic acid). The eluting conditions were conducted: 0–5 min, 10–20% B; 5–25 min, 20–70% B; 25–27 min, 70–90% B; 27–30 min, 90% B; 30–31 min, 90–10% B; and 31–34 min, 10% B to re-equilibrate the column. Approximately 0.2 mg of each product was dissolved in 600 µL $CDCl_3$ and subjected to 1D NMR (^{1}H, ^{13}C) and 2D NMR (HSQC, HMBC, COSY NOESY) analyses on a Bruker Avance 600 MHz NMR spectrometer. CD analyses were performed with 0.2 mg/mL samples in CH_3OH on the Jasco J-815 CD spectrophotometer from 200–400 nm.

3 Results

3.1 Discovery of the first plant spirooxindole synthase

Using OrthoFinder, we generated the orthogroups from the publicly available transcriptomes of ten species, including non-MIA-producing plants, MIA-producing and spirooxindoles-free plants, and spirooxindole-producing plants (see Materials and method). We focus on orthogroups containing CYP71 orthologues specific for spiroxindole alkaloid-producing plants such as *M. speciosa* (Supplementary Figure 2A) (Bindra, 1973; Ahmad and Salim, 2015; Manwill et al., 2022) as we speculated that member(s) of the CYP71 subfamily could oxidize the seco-/hetero-yohimbine alkaloids to oxygenated intermediates, which would be rearranged to spirooxindole pairs (Figure 1). Our analysis identified six candidates, namely MsCYP53813, MsCYP72054, MsCYP72056, MsCYP9580, MsCYP9583, and MsCYP9585 from the orthogroup OG0016157.

From the spirooxindole structures reported *in M. speciosa*, we traced back to the plausible corynanthe- and ajmalicine-type precursors. Among these, ajmalicine, tetrahydroalstonine, mitragynine, 9-hydroxycorynantheidine, yohimbine, corynanthine, and hirsuteine were available and used for the functional validation of CYP candidates. To test the enzyme activities, 10 µM of the putative substrates were fed to 100-µL YPL154C:PEP4KO yeast cultures for 48 hr. Only yeast cultures harbouring the construct pESC-Leu2d::*CPR/MsCYP72056* showed the consumption of hirsuteine ($[M+H]^+$ *m/z* 367.5) and the formation of two new products ($[M+H]^+$ *m/z* 383.5), **1** at 2.4 min and **2** at 2.6 min, as analyzed by LC-MS/MS (Supplementary Figure 5A). A 16-amu difference between the products and the substrate indicated that MsCYP72056 catalyzed an oxygenation/oxidation reaction. No enzymatic product was observed when hirsuteine was incubated with yeast transformed with an empty vector or constructs containing other CYP candidates. *In vitro* assays with microsomal protein of yeast expressing pESC-leu2d::*CPR/MsCYP72056* also showed that in the presence of NADPH, hirsuteine was consumed, resulting in the formation of products **1** and **2** (Figure 2A). Michaelis–Menten kinetics characterization of MsCYP72056 with hirsuteine revealed a K_M value of 68.33 µM. We also investigated the *in vitro* activity of MsCYP72056 in the pH range of 4–10 and found that the enzymatic reaction was more favourable at pH

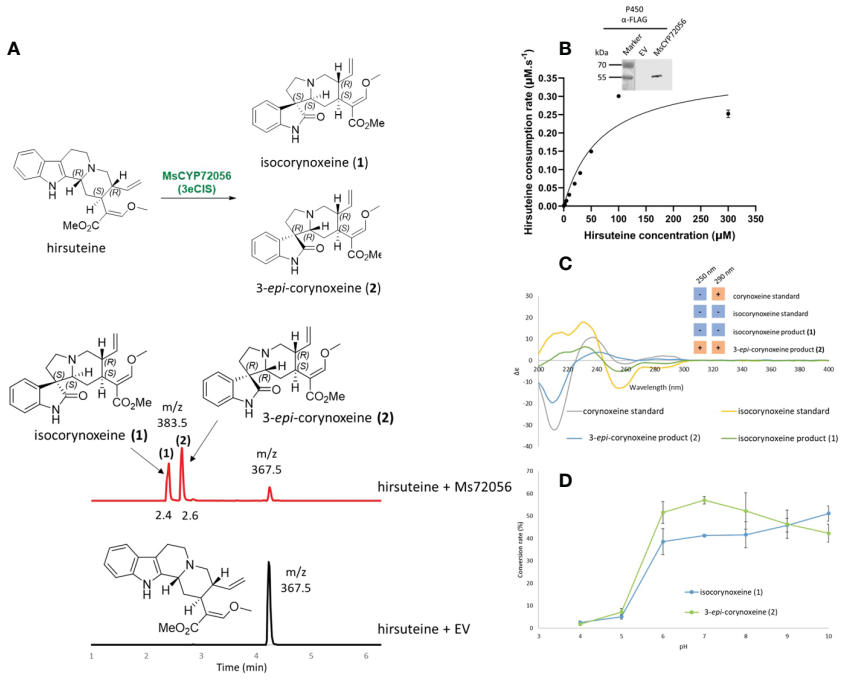


FIGURE 2
The activity of 3eCIS (MsCYP72056) with hirsuteine as substrate. **(A)** Enzymatic reaction and extracted ion chromatograms from LC-MS analysis showing the *in vitro* activity of MsCYP72056 with hirsuteine in HEPES pH 7.5 buffer. **(B)** Immunoblotting and enzyme kinetics of MsCYP72056 using total microsomal protein extraction of *S. cerevisiae* expressing MsCYP72056/CPR. **(C)** Circular dichroism spectra and Cotton effects at 250 nm and 290 nm of corynoxeine, isocorynoxeine, and enzymatic products **1** and **2**. **(D)** Product profile of *in vitro* assays of recombinant MsCYP72056 at different pHs.

6–10 (Figure 2B). There was a 20-fold increase of isocorynoxeine levels from the same amount of substrate as the pH increased from 4 to 10, of which the most dramatic increase (5% to 50%) was observed as the pH increased from 5 to 6 (Figure 2D).

In addition to hirsuteine, 16 other structurally various indole alkaloids were used to assess the substrate scope of MsCYP72056 (Supplementary Figure 13). Both *in vivo* and *in vitro* assays using cultures of yeast containing pESC-leu2d::*CPR/MsCYP72056* and its microsomal fractions, respectively, showed that *MsCYP72056* accepted hirsutine beside hirsuteine but not the other alkaloids (Supplementary Figures 5, 13). Similar to hirsuteine, hirsutine is also a (3*R*) secoyohimbine alkaloid with an ethyl group at C-20 in place of a vinyl group in hirsuteine (Supplementary Figure 5). LC-MS analysis showed the consumption of hirsutine by MsCYP72056, yielding two products **3** and **4** with 16 amu (*m/z* 385.5) greater than the substrate (*m/z* 369.5). Multiple reaction monitoring (MRM) and daughter scan analyses were performed to detect the oxindole scaffold of the enzymatic products. The specific daughter ion of oxindole ($[M+H]^+$ *m/z* 160.0) was observed, confirming that hirsutine was also converted into two spirooxindole products (Avula et al., 2015). Based on MsCYP72056 activity with hirsuteine, we speculated that the products from hirsutine were (3-*epi*-) rhynchophylline-type spirooxindoles, some of which have recently been isolated from *M. speciosa* (Flores-Bocanegra et al., 2020).

3.2 Structural elucidation of spirooxindole enzymatic products

To elucidate the stereochemistry of the enzymatic products **1** and **2**, large-scale *in vitro* reactions were conducted. Approximately 0.2 mg of the two products were purified and subjected to 1D NMR (^{1}H, ^{13}C), 2D NMR (HSQC, HMBC, COSY, NOESY), and circular dichroism (CD) analyses (Supplementary Figures 7-9). The 1D NMR data confirmed the spirooxindole skeleton of two enzymatic products compared to previously reported compounds corynoxeine and isocorynoxeine (Supplementary Tables 1, 2; Supplementary Figures 7, 10) (Kitajiina et al., 2001; Flores-Bocanegra et al., 2020). In particular, the spectra of compound **2** resemble those of corynoxeine, and **1** had identical spectra with isocorynoxeine. In 2D NMR analysis, key NOE signals of isocorynoxeine, such as H-3/H-6, H-5/H-21, and H-20/H-21 were observed in NOESY spectrum of **1** (Supplementary Table 2; Supplementary Figures 11B). These signals initially indicated the 3*S* and 20*R* configurations of C-3 and C-20 in **1** (Seki et al., 1993; Qi et al., 2015). Based on the previous studies in conformational analysis of spirooxindole alkaloids, **1** possessed the *normal*-type conformation as that of mitraphylline and formosanine (Shamma et al., 1967; Seki et al., 1993). As the deshielding effect at H-9 has been reported to differentiate the 7*S* and 7*R* isomers in the *normal*-type spirooxindole alkaloids (Seki et al., 1993), the downfield shift of H-9 in **1** (δ 7.45) depicted the 7*S* configuration (Supplementary Table 2). Although **2** had similar 1D NMR data with corynoxeine, we could not observe the key NOE signals of the *normal*-type conformation (3*S*, 20*R*) in the spectra of **2** (Supplementary Table 1; Supplementary Figure 8B). However, we could detect the NOE signals of the *pseudo*-type alkaloids (3*R*, 20*R*) as recently described in two compounds 3-*epi*-rhynchophylline and 3-*epi*-corynoxine B (Flores-Bocanegra et al., 2020). The NOE correlations were among H-3/H-9, H-9/H-6, H-5/H-3, H-18/H-20 (Supplementary Figure 8B). Therefore, we proposed that **2** was an epimer of corynoxeine, which was 3-*epi*-corynoxeine (3*R*, 7*R*, 20*R*).

To further confirm our proposed structures, we examined the C-3 and C-7 stereocenters of the enzymatic products of MsCYP72056 and the authentic standards of corynoxeine and isocorynoxeine using circular dichroism (CD) spectroscopic analysis (Figure 2C). Intriguingly, the positive Cotton effect at 250 nm confirmed the 3*R* configuration of product **2**. On the contrary, a negative Cotton effect at 250 nm suggested the 3*S* configuration of product **1**. By comparing the CD and NMR spectra of product **1** with isocorynoxeine standard data (Figure 2C, Supplementary Table 2), we concluded that **1** was isocorynoxeine (3*S*, 7*S*) (Figure 2A). Meanwhile, **2** was a 3*R* spirooxindole since it has the opposite Cotton effect at 250 nm, as compared to the corynoxeine standard (Figure 2C). Therefore, product **2** (3*R*, 7*R*) was 3-*epi*-corynoxeine (Figure 2A). Based on the product profile of the enzymatic reaction, we named this enzyme 3-*epi*-corynoxeine/isocorynoxeine synthase (3eCIS).

4 Discussion

Spirooxindole alkaloids have become highly sought-after scaffolds thanks to their stereochemical diversity and substantial range of bioactivities (Nasri et al., 2021). The total synthesis of spiroxindole involves lengthy procedures (up to 17 steps) (Wanner et al., 2013; Zhang et al., 2019), some of which require costly as well as toxic catalysts such as palladium and result in minute yields of products in complex mixtures (Nasri et al., 2021). Although hundreds of tetra- and pentacyclic spirooxindole alkaloids have been discovered (Nasri et al., 2021) and MIA biosynthesis has been studied extensively (Pan et al., 2016; Dang et al., 2017; Dang et al., 2018; Caputi et al., 2018; Farrow et al., 2019; Lopes et al., 2019; Hong et al., 2022; Wang et al., 2022), no enzymes catalyzing the oxidative rearrangement of corynanthe alkaloids to spirooxindole alkaloids have been reported. Our discovery of 3eCIS in *M. speciosa* has provided an answer to the historical question of spirooxindole biosynthesis in plants and highlights the versatility of CYP71 enzymes in MIA scaffolding. From secoyohimbine precursors, 3eCIS catalyzes the formation of polycyclic spirooxindole alkaloids with high yield (10 μg/mL purified product from *in vitro* reaction). In particular, our LC-MS/MS analysis showed a total conversion of hirsuteine substrate to the spirooxindole products at physiological pH and 37°C within 1 hour (Supplementary Figure 6). CYP71 enzymes are well known for their roles in MIA metabolism, such as sarpagan bridge enzyme from serpentine wood (*R. serpentina*) (Dang et al., 2018), and geissochizine oxidases from Madagascar periwinkle (*C. roseus*) (Tatsis et al., 2017; Qu et al., 2018) and blackboard tree (*Alstonia scholaris*) (Wang et al., 2022). These notable examples and 3eCIS reported here provide the entry points to various MIA subgroups, including spirooxindole (this work), sarpagan (Dang et al., 2018), strychno (Tatsis et al., 2017; Hong et al., 2022) and akuammilan (Wang et al., 2022) alkaloids. While a flavoprotein monooxygenase was recently reported to catalyze the formation of a spirooxindole alkaloid in fungus, 3eCIS is the first plant enzyme reported to bear the spirooxindole formation catalytic activity (Tsunematsu et al., 2013; Liu et al., 2021). Intriguingly, oxidative rearrangements of the biosynthetic motif of seco-/

heteroyohimbine alkaloids have been used to chemically synthesize the spirooxindole core (Xu et al., 2019), and extensive studies in spirooxindole synthesis have been focused on more sustainable synthetic strategies, including less toxic catalysts and less chemical waste (Yan and Wang, 2016; Zhang et al., 2017; Xu et al., 2019). Towards this end, our finding paves the way to access spirooxindole alkaloids and their derivatives through biocatalysis and enzyme engineering.

Spirooxindoles from kratom are highly diverse in terms of scaffold and stereochemistry (Flores-Bocanegra et al., 2020). The stereogenic centers (C-3, C-7, C-15 and C-20) of spirooxindoles are considered key in deciphering the upstream biosynthetic pathways of spirooxindoles (Figure 1A). Previous studies hypothesized that these configurations remain unchanged between an oxindole product and its tetrahydro-β-carboline precursor; therefore, the products **1** and **2** of 3eCIS were initially expected to share the stereo-configurations in rings C and D with those of hirsuteine (3*R*, 20*R*) (Figure 1B) (Lopes et al., 2019). Since the *pseudo*-conformation of 3*R* spirooxindoles are not stable due to the interaction of the spirooxindole ring with ring D, the 3*R* spirooxindoles could spontaneously isomerize to the 3*S* spirooxindoles (*normal*-conformation) *via* intramolecular Mannich reaction to reduce the steric hindrance (Scheme 1B) (Seki et al., 1993). The C–C single bonds in C-5–C-6–C-7 could freely rotate, resulting in interconvertible 3*R* and 3*S* spirooxindoles through a zwitterion intermediate. In contrast to the configuration retention hypothesis, our newly found enzyme could transform the 3*R* secoyohimbine alkaloid to a mixture of 3*R* and 3*S* spirooxindoles, which can be partially controlled by different pH (Figure 2D). In the enzymatic reaction with hirsuteine catalyzed by 3eCIS, in addition to the two characterized products 3-*epi*-corynoxeine and isocorynoxeine, it is possible that other stereoisomers, corynoxeine or 3-*epi*-isocorynoxeine, were generated as we noticed a small peak of *m/z* 383 ($[M+H]^+$) in the extracted ion and MRM chromatograms (Figure 2A).

Computational and synthetic chemistry studies suggested that a seco-/hetero-yohimbine alkaloid such as hirsuteine can be converted to a pair of spirooxindole epimers by epoxidation on the indole ring followed by (semi-)pinacol rearrangement (Scheme 1A) (Xu et al., 2019; Liu et al., 2021). The 3eCIS enzyme is proposed to catalyze the initial oxygenation step (Xu et al., 2019). Then, the formation of carbocation at C-7 could occur without water addition via the (semi-) pinacol mechanism. Subsequently, the ring opening likely allows the alkyl chain at C-3 to rearrange on both sides of the indole ring to yield the spirooxindoles (Scheme 1A).

Most spirooxindole alkaloids occur at trace abundance levels *in planta* (Nasri et al., 2021). The discovery and characterization of 3eCIS as reported here provide insights into the long puzzling biosynthesis of plant spirooxindoles and open the gateway to obtaining the elusive spirooxindole alkaloids. The high selectivity towards the (3*R*) tetracyclic corynanthe-type alkaloids of 3eCIS could serve as a starting point for gene discovery and enzyme engineering toward accessing and diversifying spirooxindole core-containing molecules.

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

T-TTD conceived and designed the project. T-AMN, T-DN, and T-TTD designed the experiments and wrote the manuscript. T-AMN and DG characterized the CYPs *in vitro* and *in vivo*. T-AMN, KC, and ZX characterized the products. All authors contributed to the article and approved the submitted version.

Funding

T-AMN received the University of British Columbia's Graduate Entrance Scholarship and the University of British Columbia Okanagan Graduate Research Scholarships. T-TTD received funding from Canada's Natural Science and Engineering Research Council (NSERC, RGPIN-2019-05473), the Canada Foundation for Innovation (38167), the University of British Columbia's Eminence Fund, and the Michael Smith Foundation for Health Research Scholar (SCH-2020-0401).

Acknowledgments

We sincerely thank Dr. Jakob Franke (Institute of Botany, Leibniz University of Hannover, Germany) for his helpful discussion on this manuscript. We thank Dr. Susan Murch and Dr. Kirsten Wolthers (University of British Columbia) for their helps with plant material and equipment. We thank Mr. Benjamin Herring, Mr. Oleg Sannikov and Ms. Yuki Liang (University of British Columbia) for their assistance in circular dichroism analysis. We also thank Ms. Brooke Kwan (University of British Columbia) for her helpful support in OrthoFinder analysis.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1125158/full#supplementary-material

References

Ahmad, R., and Salim, F. (2015). Oxindole alkaloids of uncaria (Rubiaceae, subfamily cinchonoideae). *Stud. Natural Products Chem.* (Amsterdam: Elsevier), 485–525. doi: 10.1016/B978-0-444-63473-3.00012-5

Avula, B., Sagi, S., Wang, Y. H., Wang, M., Ali, Z., Smillie, T. J., et al. (2015). Identification and characterization of indole and oxindole alkaloids from leaves of mitragyna speciosa korth using liquid chromatography - accurate QToF mass spectrometry. *J. AOAC Int.* 98, 13–21. doi: 10.5740/jaoacint.14-110

Bindra, J. S. (1973). Chapter 2 oxindole alkaloids. *Alkaloids: Chem. Physiol.* (New York), 83–121. doi: 10.1016/S1876-0813(08)60219-5

Brose, J., Lau, K. H., Dang, T. T. T., Hamilton, J. P., Martins, L. D. V., Hamberger, B., et al. (2021). The mitragyna speciosa (Kratom) genome: A resource for data-mining potent pharmaceuticals that impact human health. *G3: Genes Genomes Genet.* 11. doi: 10.1093/g3journal/jkab058

Caputi, L., Franke, J., Farrow, S. C., Chung, K., Payne, R. M. E., Nguyen, T. D., et al. (2018). Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle. *Sci. (1979)* 360, 1235–1239. doi: 10.1126/science.aat4100

Chen, L., Huang, Y., Yu, X., Lu, J., Jia, W., Song, J., et al. (2021). Corynoxine protects dopaminergic neurons through inducing autophagy and diminishing neuroinflammation in rotenone-induced animal models of parkinson's disease. *Front. Pharmacol.* 12. doi: 10.3389/fphar.2021.642900

Dang, T. T. T., Franke, J., Carqueijeiro, I. S. T., Langley, C., and Courdavault, V. (2018). And O'Connor, sSarpagan bridge enzyme has substrate-controlled cyclization and aromatization modes. *E.Nat. Chem. Biol.* 14, 760–763. doi: 10.1038/s41589-018-0078-4

Dang, T. T. T., Franke, J., and Tatsis, E. (2017). And O'Connor, sDual catalytic activity of a cytochrome P450 controls bifurcation at a metabolic branch point of alkaloid biosynthesis in rauwolfia serpentina. *E.Angewandte Chemie - Int. Edition* 56, 9440–9444. doi: 10.1002/anie.201705010

Emms, D. M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16, 157. doi: 10.1186/s13059-015-0721-2

Farrow, S. C., Kamileen, M. O., Caputi, L., Bussey, K., Mundy, J. E. A., McAtee, R. C., et al. (2019). Biosynthesis of an anti-addiction agent from the iboga plant. *J. Am. Chem. Soc.* 141, 12979–12983. doi: 10.1021/jacs.9b05999

Flores-Bocanegra, L., Raja, H. A., Graf, T. N., Augustinović, M., Wallace, E. D., Hematian, S., et al. (2020). The chemistry of kratom [mitragyna speciosa]: Updated characterization data and methods to elucidate indole and oxindole alkaloids. *J. Nat. Prod* 83, 2165–2177. doi: 10.1021/acs.jnatprod.0c00257

García Giménez, D., García Prado, E., Sáenz Rodríguez, T., Fernández Arche, A., and de la Puerta, R. (2010). Cytotoxic effect of the pentacyclic oxindole alkaloid mitraphylline isolated from uncaria tomentosa bark on human ewing's sarcoma and breast cancer cell lines. *Planta Med.* 76, 133–136. doi: 10.1055/s-0029-1186048

Hong, B., Grzech, D., Caputi, L., Sonawane, P., López, C. E. R., Kamileen, M. O., et al. (2022). Biosynthesis of strychnine. *Nature* 607, 617–622. doi: 10.1038/s41586-022-04950-4

Kitajiina, M., Yokoya, M., Takayama, H., and Aimi, N. (2001). Co-Occurrence of harman and b-carboline-type monoterpenoid glucoindole alkaloids in ufia de gato (Uncaria tomentosa). *Natural Medicines* 55, 308–310.

Laus, G., Brössner, D., Senn, G., and Wurst, K. (1996). Analysis of the kinetics of isomerization of spiro oxindole alkaloids. *J. Chem. Soc. Perkin Trans. 2* 1931–1936. doi: 10.1039/P29960001931

Liu, Z., Zhao, F., Zhao, B., Yang, J., Ferrara, J., Sankaran, B., et al. (2021). Structural basis of the stereoselective formation of the spirooxindole ring in the biosynthesis of citrinadins. *Nat. Commun.* 12. doi: 10.1038/s41467-021-24421-0

Lopes, A. A., Chioca, B., Musquiari, B., Crevelin, E. J., França, S., de, C., et al. (2019). Unnatural spirocyclic oxindole alkaloids biosynthesis in uncaria guianensis. *Sci. Rep.* 9. doi: 10.1038/s41598-019-47706-3

Manwill, P. K., Flores-Bocanegra, L., Khin, M., Raja, H. A., Cech, N. B., Oberlies, N. H., et al. (2022). Kratom (Mitragyna speciosa) validation: Quantitative analysis of indole and oxindole alkaloids reveals chemotypes of plants and products. *Planta Med.* doi: 10.1055/a-1795-5876

Nasri, S., Bayat, M., and Mirzaei, F. (2021). Recent strategies in the synthesis of spiroindole and spirooxindole scaffolds. *Top. Curr. Chem.* 379, 25. doi: 10.1007/s41061-021-00337-7

Nguyen, T.-D., and Dang, T.-T. T. (2021). Cytochrome P450 enzymes as key drivers of alkaloid chemical diversification in plants. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.682181

Nguyen, T. A. M., McConnachie, M., Nguyen, T. D., and Dang, T. T. T. (2022). Discovery and characterization of oxidative enzymes involved in monoterpenoid indole alkaloid biosynthesis. *Methods Mol. Biol.* (New York), 141–164. doi: 10.1007/978-1-0716-2349-7_11

Nguyen, T. A. M., Nguyen, T. D., Leung, Y. Y., McConnachie, M., Sannikov, O., Xia, Z., et al. (2021). Discovering and harnessing oxidative enzymes for chemoenzymatic synthesis and diversification of anticancer camptothecin analogues. *Commun. Chem.* 4. doi: 10.1038/s42004-021-00602-2

Pan, Q., Mustafa, N. R., Tang, K., Choi, Y. H., and Verpoorte, R. (2016). Monoterpenoid indole alkaloids biosynthesis and its regulation in catharanthus roseus: a literature review from genes to metabolites. *Phytochem. Rev.* 15, 221–250. doi: 10.1007/s11101-015-9406-4

Qi, W., Chen, F., Sun, J., Simpkins, J. W., and Yuan, D. (2015). Isolation and identification of twelve metabolites of isocorynoxeine in rat urine and their neuroprotective activities in HT22 cell assay. *Planta Med.* 81, 46–55. doi: 10.1055/s-0034-1383357

Qu, Y., Easson, M. E. A. M., Simionescu, R., Hajicek, J., Thamm, A. M. K., Salim, V., et al. (2018). Solution of the multistep pathway for assembly of corynanthean, strychnos, iboga, and aspidosperma monoterpenoid indole alkaloids from 19E-geissoschizine. *Proc. Natl. Acad. Sci. U.S.A.* 115, 3180–3185. doi: 10.1073/pnas.1719979115

Rai, A., Hirakawa, H., Nakabayashi, R., Kikuchi, S., Hayashi, K., Rai, M., et al. (2021). Chromosome-level genome assembly of ophiorrhiza pumila reveals the evolution of camptothecin biosynthesis. *Nat. Commun.* 12. doi: 10.1038/s41467-020-20508-2

Rather, G. A., Sharma, A., Pandith, S. A., Kaul, V., Nandi, U., Misra, P., et al. (2018). *De novo* transcriptome analyses reveals putative pathway genes involved in biosynthesis and regulation of camptothecin in nothapodytes nimmoniana (Graham) mabb. *Plant Mol. Biol.* 96, 197–215. doi: 10.1007/s11103-017-0690-9

Ro, D. K., Ouellet, M., Paradise, E. M., Burd, H., Eng, D., Paddon, C. J., et al. (2008). Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnol.* 8, 1–14. doi: 10.1186/1472-6750-8-83

Seki, H., Takayama, H., Aimi, N., Sakai, S., and Ponglux, D. (1993). A nuclear magnetic resonance study on the eleven stereoisomers of heteroyohimbine-type oxindole alkaloids. *Chem. Pharm. Bull. (Tokyo)* 41, 2077–2086. doi: 10.1248/cpb.41.2077

Shamma, M., Shine, R. J., Kompis, I., Sticzay, T., Morsingh, F., Poisson, J., et al. (1967). The stereochemistry of the pentacyclic oxindole alkaloids. *J. Am. Chem. Soc.* 89, 1739–1740. doi: 10.1021/ja00983a041

Shavel, J., and Zinnes, H. (1962). Oxindole alkaloids. i. oxidative-rearrangement of indole alkaloids to their oxindole analogs. *J. Am. Chem. Soc.* 84, 1320–1321. doi: 10.1021/ja00866a063

Tatsis, E. C., Carqueijeiro, I., Dugé de Bernonville, T., Franke, J., Dang, T.-T. T., Oudin, A., et al. (2017). A three enzyme system to generate the strychnos alkaloid scaffold from a central biosynthetic intermediate. *Nat. Commun.* 8, 316. doi: 10.1038/s41467-017-00154-x

Tsunematsu, Y., Ishikawa, N., Wakana, D., Goda, Y., Noguchi, H., Moriya, H., et al. (2013). Distinct mechanisms for spiro-carbon formation reveal biosynthetic pathway crosstalk. *Nat. Chem. Biol.* 9, 818–825. doi: 10.1038/nchembio.1366

Wang, Z., Xiao, Y., Wu, S., Chen, J., Li, A., and Tatsis, E. C. (2022). Deciphering and reprogramming the cyclization regioselectivity in bifurcation of indole alkaloid biosynthesis. *Chem. Sci.* doi: 10.1039/D2SC03612F

Wanner, M. J., Ingemann, S., van Maarseveen, J. H., and Hiemstra, H. (2013). Total synthesis of the spirocyclic oxindole alkaloids corynoxine, corynoxine b, corynoxeine, and rhynchophylline. *Eur. J. Org Chem.* 2013, 1100–1106. doi: 10.1002/ejoc.201201505

Xu, J., Liang, L., Zheng, H., Chi, Y. R., and Tong, R. (2019). Green oxidation of indoles using halide catalysis. *Nat. Commun.* 10. doi: 10.1038/s41467-019-12768-4

Yan, L. J., and Wang, Y. C. (2016). Recent advances in green synthesis of 3,3′-spirooxindoles *via* isatin–based one–pot multicomponent cascade reactions in aqueous medium. *ChemistrySelect* 1, 6948–6960. doi: 10.1002/slct.201601534

Zhang, J. H., Wang, R. B., Li, D. F., and Zhao, L. M. (2017). Green method to preparing oxindole-fused spirotetrahydrofuran scaffolds through methanesulfonic acid-catalyzed cyclization reactions of 3-allyl-3-hydroxy-2-oxindole in water. *ACS Omega* 2, 7022–7028. doi: 10.1021/acsomega.7b01279

Zhang, Z., Zhang, W., Kang, F., Ip, F. C. F., Ip, N. Y., and Tong, R. (2019). Asymmetric total syntheses of rhynchophylline and isorhynchophylline. *J. Org Chem.* 84, 11359–11365. doi: 10.1021/acs.joc.9b01977

Zhao, L., Zang, B., Qi, W., Chen, F., Wang, H., Kano, Y., et al. (2016). Pharmacokinetic study of isocorynoxeine metabolites mediated by cytochrome P450 enzymes in rat and human liver microsomes. *Fitoterapia* 111, 49–57. doi: 10.1016/j.fitote.2016.04.008

Zhou, L. M., Qu, R. Y., and Yang, G. F. (2020). An overview of spirooxindole as a promising scaffold for novel drug discovery. *F.Expert Opin. Drug Discovery* 15, 603–625. doi: 10.1080/17460441.2020.1733526

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 13 February 2023
DOI 10.3389/fpls.2023.1140509

Identification of drought-responsive phenolic compounds and their biosynthetic regulation under drought stress in *Ligularia fischeri*

Yun Ji Park[1], Do Yeon Kwon[2], Song Yi Koo[3], To Quyen Truong[1,4], Sung-Chul Hong[1], Jaeyoung Choi[1], Jinyoung Moon[1] and Sang Min Kim[1,4]*

[1]Smart Farm Research Center, KIST Gangneung Institute of Natural Products, Gangneung, Republic of Korea, [2]Euseed Inc, Daejeon, Republic of Korea, [3]Natural Product Informatics Center, KIST Gangneung Institute of Natural Products, Gangneung, Republic of Korea, [4]Department of Bio-medical Science & Technology, Korea Institute of Science and Technology (KIST) School, University of Science and Technology, Seoul, Republic of Korea

OPEN ACCESS

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

REVIEWED BY
Choonkyun Jung,
Seoul National University, Republic of Korea
Anh Tuan Pham,
University of Manitoba, Canada

*CORRESPONDENCE
Sang Min Kim
kimsm@kist.re.kr

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 09 January 2023
ACCEPTED 30 January 2023
PUBLISHED 13 February 2023

CITATION
Park YJ, Kwon DY, Koo SY, Truong TQ, Hong S-C, Choi J, Moon J and Kim SM (2023) Identification of drought-responsive phenolic compounds and their biosynthetic regulation under drought stress in *Ligularia fischeri*. *Front. Plant Sci.* 14:1140509. doi: 10.3389/fpls.2023.1140509

Ligularia fischeri, a leafy edible plant found in damp shady regions, has been used as an herbal medicine and is also consumed as a horticultural crop. In this study, we investigated the physiological and transcriptomic changes, especially those involved in phenylpropanoid biosynthesis, induced by severe drought stress in *L. fischeri* plants. A distinguishing characteristic of *L. fischeri* is a color change from green to purple due to anthocyanin biosynthesis. We chromatographically isolated and identified two anthocyanins and two flavones upregulated by drought stress using liquid chromatography-mass spectrometry and nuclear magnetic resonance analyses in this plant for the first time. In contrast, all types of caffeoylquinic acids (CQAs) and flavonol contents were decreased under drought stress. Further, we performed RNA sequencing to examine the molecular changes in these phenolic compounds at the transcriptome level. In an overview of drought-inducible responses, we identified 2,105 hits for 516 distinct transcripts as drought-responsive genes. Moreover, differentially expressed genes (DEGs) associated with phenylpropanoid biosynthesis accounted for the greatest number of both up- and downregulated DEGs by Kyoto Encyclopedia of Genes and Genomes enrichment analysis. We identified 24 meaningful DEGs based on the regulation of phenylpropanoid biosynthetic genes. Potential drought-responsive genes included upregulated flavone synthase (*LfFNS*, TRINITY DN31661 c0 g1 i1) and anthocyanin 5-*O*-glucosyltransferase (*LfA5GT1*, TRINITY DN782 c0 g1 i1), which could contribute to the high levels of flavones and anthocyanins under drought stress in *L. fischeri*. In addition, the downregulated shikimate *O*-hydroxycinnamolytransferase (*LfHCT*, TRINITY DN31661 c0 g1 i1) and hydroxycinnamoyl-CoA quinate/shikimate transferase (*LfHQT4*, TRINITY DN15180 c0 g1 i1) genes led to a reduction in CQAs. Only one or two BLASTP hits for *LfHCT* were obtained for six different Asteraceae species. It is possible that

the HCT gene plays a crucial role in CQAs biosynthesis in these species. These findings expand our knowledge of the response mechanisms to drought stress, particularly regarding the regulation of key phenylpropanoid biosynthetic genes in *L. fischeri*.

KEYWORDS

***Ligularia fischeri*, drought stress, transcriptome, drought-responsive gene, phenylpropanoid biosynthesis, phenolic compounds, anthocyanin**

1 Introduction

Ligulari fischeri, also referred to as *Gomchi* in Korea, is a medicinal plant species in the family Asteraceae mn. This edible leafy plant is found in wet shady regions and is mainly distributed in China, Korea, Japan, Europe, and the far-east region of Russia (Xie et al., 2010). Notably, this plant is used as a cooking ingredient as well as a traditional remedy to treat various diseases in China and its leaves have been historically consumed as either fresh or salted fermented vegetables in Korea. In addition, *L. fischeri* leaf tea, which is brewed by blanching fresh leaves in hot water, has recently been considered a high-value nutritional food, and its primary ingredients in tea infusion have been investigated (Kim et al., 2010). Currently, *L. fischeri* can be found in the wild and is also massively cultivated in farms and greenhouses, with a continuous increase in its cultivation area over time to meet the commercial demand (Kim et al., 2012). Among herbal medicines, *L. fischeri* leaves have been used to treat hepatic failure, jaundice, rheumatoid arthritis, and scarlet fever (Park et al., 2016). In accordance with several *in vitro* and *in vivo* studies, *L. fischeri* has anti-hepatotoxic, anti-inflammatory, and anti-obesity properties (Kim et al., 2019). Moreover, the solvents and aqueous extracts of this plant show high radical-scavenging activities and are a key source of dietary antioxidants (Kim et al., 2010). Most biological activities can be attributed to the functional compounds found in plants. Species in the genus *Ligularia* have been reported to contain a wide variety of phytochemicals, including alkaloids, flavonoids, steroids, terpenoids, and other compounds associated with various biological activities (Yang et al., 2011). Previous studies have shown that the therapeutic effects of *L. fischeri* are determined by active compounds such as caffeoylquinic acid (CQA), sesquiterpenoids, phenolic compounds, terpenoids, monocyclosqualene, spiciformisins, and norsesquiterpene derivatives (Azam Ansari et al., 2019). Of these, CQA and most CQA derivatives have been identified as the predominant phenolic compounds in *L. fischeri* leaves (Shang et al., 2010).

Environmental constraints substantially influence plant growth and development, with effects ranging from mild to severe, based on the intensity and duration of stress and plant growth stage (Sourour et al., 2017). Several global regions are dealing with global warming-induced water shortages and droughts, which in turn result in reduced water uptake, restricting subsequent nutrient absorption and affecting crop growth, gene expression, distribution, productivity, and quality (Guo et al., 2018). Water scarcity impairs events involving cell mitosis, such as division, elongation, and turgor pressure, as well as critical enzymes responsible for nutrient assimilation, resulting in nutrient deficiency and reductions in height, leaf area, number, and plant biomass (Farooq et al., 2009). The most severe effect of drought is observed in photosynthesis, which is an essential process that provides energy for all metabolic activities in plants. Water shortages impair enzymes involved in carbon fixation, the electron transport chain, and photorespiration, and cause deterioration of the photosynthetic apparatus (Farooq et al., 2009). Moderate to severe water stress affects a variety of morpho-physiological properties, including chlorophyll fluorescence, water usage efficiency, dry matter output, water content, water potential, membrane ability, and pigment content stability (Sourour et al., 2017). Previous studies have revealed that the accumulation of organic and inorganic solutes is induced to maintain the actual intercellular water level and to promote carbon dioxide uptake *via* the stomata and water *via* roots. Moreover, the regulation of plant growth factors and phytohormones is also modified by drought stress. Plant cells have developed an antioxidant defense system because insufficient water supply is linearly associated with the increase in and accumulation of reactive oxygen species (Farooq et al., 2009). Therefore, plants need to develop a drought resistance system to overcome water constraints and survive under drought conditions.

Drought stress also causes a wide range of biochemical and physiological reactions that assist in the maintenance of water and ionic homeostasis and prevent plants from wilting and desiccation (Singh et al., 2017). For instance, glycolysis, hormone synthesis, photosynthesis, sugar synthesis, and the tricarboxylic acid cycle are involved in plant responses to drought stress (Guo et al., 2018). Drought-induced metabolites such as glycine-betaine, proline, and soluble sugars reduce the osmotic potential of cells and improve water infiltration in plants without disrupting normal metabolic processes. This reaction, referred to as osmotic adjustment, allows plants to sustain cell turgor for development and survival under stressful conditions (Takahashi et al., 2020). Plants activate an array of mechanisms for survival, especially the biosynthesis of secondary metabolites, including phenolics/flavonoids, when subjected to water deficits. Drought stress affects the accumulation of polyphenolic substances including anthocyanins and other flavonoids. For

Abbreviations: CQA, caffeoylquinic acids; DEGs, differentially expressed genes; FC, fold change; GO, gene ontology; HPLC, high-performance liquid chromatography; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC/MS, liquid chromatography-mass spectrometry; NMR, nuclear magnetic resonance; TFA, trifluoroacetic acid.

instance, flavonoids and anthocyanins in pea plants and graph berries are significantly increased by drought (Nogués et al., 1998; Castellarin et al., 2007). Additionally, stress-induced synthesis of antioxidants, such as flavonoids (e.g., flavonols and anthocyanins), enables plants to serve as free radical scavengers, mitigating oxidative and dehydration stress (Nakabayashi et al., 2014).

Several studies have investigated the effects of different environmental conditions, including weather, cultivation systems, and sunlight on the production of active compounds in *L. fischeri* (Kim et al., 2012; Rekha et al., 2015; Hao et al., 2018). We previously demonstrated that the major phenolic components, CQAs, were induced by sunlight exposure in *L. fischeri*. Because this plant is typically grown in moist and shady regions, several growth conditions can be altered to improve its functional constituents (Kim et al., 2012). In this study, we investigated the effect of severe drought stress in *L. fischeri* plants to evaluate its effect on polyphenolic composition and its induction of regulatory changes. We identified and characterized two flavonoids and two anthocyanins for the first time in this plant under drought stress. We compared changes in major phenolic compounds between the control and drought-treated plants. Furthermore, we performed RNA sequencing to determine differentially expressed genes and explore the drought-related network that is adapted to drought stress in *L. fischeri*.

2 Materials and methods

2.1 Reagents

Analytical grade and high-performance liquid chromatography (HPLC)-graded solvents were obtained from Daejung (Gyeonggi, Korea) and Fisher Scientific (Pittsburgh, PA, USA). Deuterated liquids for nuclear magnetic resonance (NMR) spectroscopy were obtained from the Cambridge Isotope Laboratory (Andover, MA, USA). Hyperoside, potassium persulfate, formic acid, and trifluoroacetic acid (TFA) were obtained from Sigma-Aldrich (St. Louis, MO, USA).

2.2 Plant materials

Three-year-old *L. fischeri* roots were grown under sunlight in Apr. 2019, in Gangneung, Korea. Plants were grown for 50 days under well-watered conditions. Subsequently, two groups of plants with two or three leaves each were established. The control group was maintained under the same conditions and the drought group was subjected to drought stress (with no additional water). Based on the data observed from Korea Meteorological Administration (http://data.kma.go.kr), the average temperature was 18.3°C (min 13.5°C – max 23.2°C) and the relative humidity was 54% during cultivation. After four weeks, the leaves were harvested from each group and dried under shade for one week for liquid chromatography-mass spectrometry (LC/MS) analysis and extraction for chromatography. Some leaves were stored at – 80°C for RNA extraction.

2.3 HPLC-MS and HPLC analysis

Dried leaves (200 mg) were extracted with 10 mL of methanol containing 0.1 N HCl for 12 h at room temperature. After filtration, the extract was analyzed using an Agilent LC-MS system (Agilent Technologies, Palo Alto, CA, USA), consisting of an analytical 1200 HPLC system with a Shiseido MG II C_{18} analytical column (250 × 4.6 mm i.d., 5 μm particle size) and a 6120-quadrupole mass spectrometer with electrospray ionization (ESI). The mobile phase was made up of acetonitrile (A) and water (B), with 0.3% formic acid, respectively. The gradient was run as follows: 15% solvent A for 10 min, 15–40% solvent A for 18 min, 40–90% solvent A for 7 min, held at 90% solvent A for 5 min, and then returned to 15% solvent A for column equilibration. The separated components were detected at 330 nm for CQAs and flavonoids, and 520 nm for anthocyanins. Mass data were collected in full scan mode in positive ion mode from *m/z* 50 to 1000 under a 5 mL min^{-1} drying gas flow, 150°C vaporizing temperature, 60 psi nebulizing gas (N_2) pressure, and 30°C drying gas temperature.

To determine and quantify phenolic compounds, we used the same HPLC conditions described above. An Agilent 1200 HPLC system equipped with a binary pump (G1312A), auto sampler (G1347B), PDA detector (G1315D), column oven (G1316A), and ChemStation, using an external standard method with a calibration curve, was used for HPLC analysis. In total, ten phenolic compounds, including four hydroxycinnamic acids, two flavonols, two flavones, and two anthocyanins, were used for calibration curve conduction in the range of 31.25 to 250 ppm. 2”-Acetylhyperoside was quantified as a relatively equivalent value to hyperoside. All experiments were performed in triplicate.

2.4 Isolation of flavonoids and anthocyanins

L. fischeri leaves (300 g) were extracted thrice with methanol for 12 h at room temperature and then partitioned with *n*-hexane, dichloromethane, ethyl acetate, *n*-butanol, and water to obtain two unknown flavonoids. As indicated in a previous study, the ethyl acetate fractions were analyzed by HPLC (Kim et al., 2012). The mobile phase used for separation consisted of water and methanol. The methanol concentration was increased from 20 to 50% over 30 min at a flow rate of 10 mL min^{-1}. Flavonoid compounds were detected at 330 nm. The two flavonoids were concentrated to yield 4.08 mg of Peak 3 and 5 mg of Peak 4.

For the two unknown anthocyanins, dried powder (150 g) of drought-treated leaves was extracted three times for 12 h each with methanol containing 0.1 N HCl. The extract was filtered and condensed using a vacuum evaporator. During purification, all solvents contained 0.1 N HCl to maintain the stability of anthocyanins. The crude extract (10 g) was dissolved in 100 mL of 0.1 N HCl and subjected to Amberlite XAD-7 HP (Supelco, Bellefonte, PA, USA) column chromatography (7 cm × 50 cm). The sample was eluted with aqueous methanol by gradually increasing the methanol concentration from 0 to 100%. Purple output fractions were collected and further purified by preparative HPLC as previously described (Kim et al., 2012). An acetonitrile gradient from 0 to 60%

over 60 min was used to elute anthocyanins, and chromatogram peaks were observed at 520 nm. Peaks 9 and 10 were concentrated in the two anthocyanins, yielding 4 mg of peak 9 and 3 mg of peak **10**. In total, four compounds were successfully isolated and analyzed for structure determination and quantification.

2.5 NMR spectroscopy analysis

^{1}H and ^{13}C NMR data of the separated compounds were collected using a Varian NMR system operating at 500 MHz (Varian, Palo Alto, CA, USA).

2.5.1 Luteolin-7-O-β-glucoside (peak 3)

This was a yellow powder; ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.45 (dd, 1H, J = 2.5, 8.3 Hz, H-6′), 7.41 (d, 1H, J = 2.5 Hz, H-2′), 6.89 (d, 1H, J = 8.3 Hz, H-5′), 6.78 (d, 1H, J = 2.5 Hz, H-8), 6.76 (s, 1H, H-3), 6.44 (d, 1H, J = 2.5 Hz, H-6) (Luteolin moiety); δ 5.08 (d, 1H, J = 7.3 Hz, H-1′′), 3.15–3.71 (m, 5H, H-2′′–H-6′′) (Glucose moiety). ^{13}C NMR (125 MHz, DMSO-d_6) δ 182.4 (C-4), 164.9 (C-2), 163.4 (C-7), 161.6 (C-5), 157.4 (C-9), 150.4 (C-4′), 146.2 (C-3′), 121.8 (C-1′), 119.6 (C-6′), 116.4 (C-5′), 114.0 (C-2′), 105.8 (C-10), 103.6 (C-3), 100.0 (C-6), 95.1 (C-8) (Luteolin moiety); δ 100.2 (C-1′′), 77.6 (C-5′′), 76.8 (C-3′′), 73.5 (C-2′′), 69.9 (C-4′′), 61.0 (C-6′′) (Glucose moiety).

2.5.2 Luteolin-7-O-β-glucuronide (peak 4)

This was a yellow powder; ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.42 (d, 1H, J = 2.2 Hz, H-2′), 7.37 (dd, 1H, J = 2.2, 8.4 Hz, H-6′), 6.84 (d, 1H, J = 8.4 Hz, H-5′), 6.77 (d, 1H, J = 1.8 Hz, H-8), 6.68 (s, 1H, H-3), 6.39 (d, 1H, J = 1.8 Hz, H-6) (Luteolin moiety); δ 13.01 (s, 1H, -COOH), 5.08 (d, 1H, J = 7.3 Hz, H-1′′), 3.60 (d, 1H, J = 9.9 Hz, H-5′′), 3.26–3.40 (m, 3H, H-2′′–H4′′) (Glucuronide moiety). ^{13}C NMR (125 MHz, DMSO-d_6) δ 182.4 (C-4), 164.9 (C-2), 162.9 (C-7), 161.6 (C-5), 157.4 (C-9), 150.4 (C-4′), 146.2 (C-3′), 121.8 (C-1′), 119.6 (C-6′), 116.4 (C-5′), 114.0 (C-2′), 105.8 (C-10), 103.6 (C-3), 99.5 (C-6), 94.9 (C-8) (Luteolin moiety); δ 170.6 (C-6′′), 99.8 (C-1′′), 76.0 (C-3′′), 75.8 (C-5′′), 73.2 (C-2′′), 71.7 (C-4′′) (Glucuronide moiety).

2.5.3 Cyanidin-3-O-β-glucoside (peak 9)

This was a purple powder; ^{1}H NMR (500 MHz, CD_3OD/TFA-d_1 (9: 1, v/v)) δ 9.06 (s, 1H, H-4), 8.29 (dd, 1H, J = 2.3, 8.7 Hz, H-6′), 8.09 (d, 1H, J = 2.3 Hz, H-2′), 7.06 (d, 1H, J = 8.7 Hz, H-5′), 6.94 (d, 1H, J = 1.2 Hz, H-8), 6.70 (d, 1H, J = 2.0 Hz, H-6) (Cyanidin moiety); δ 5.33 (d, 1H, J = 7.8 Hz, H-1′′), 3.40–4.00 (m, 5H, H-2′′–H-6′′) (Glucose moiety). ^{13}C NMR (125 MHz, CD_3OD/TFA-d_1 (9: 1, v/v)) δ 171.0 (C-7), 165.2 (C-2), 160.2(C-5), 158.7 (C-9), 156.7 (C-4′), 148.3 (C-3′), 146.6 (C-3), 137.9 (C-4), 129.1 (C-6′), 122.2 (C-1′), 119.3 (C-2′), 118.2 (C-5′), 114.2 (C-10), 104.6 (C-6), 96.0(C-8) (Cyanidin moiety); δ 104.2 (C-1′′), 79.7 (C-5′′), 79.0 (C-3′′), 75.6 (C-2′′), 71.9 (C-4′′), 63.2 (C-6′′) (Glucose moiety).

2.5.4 Cyanidin-3-O-β-(6′′-malonylglucoside) (peak 10)

This was a purple powder; ^{1}H NMR (500 MHz, CD_3OD/TFA-d_1 (9: 1, v/v)) δ 8.79 (s, 1H, H-4), 8.21 (dd, 1H, J = 2.4, 8.8 Hz, H-6′), 7.97 (d, 1H, J = 2.4 Hz, H-2′), 7.01 (d, 1H, J = 8.8 Hz, H-5′), 6.87 (d, 1H, J = 1.2 Hz, H-8), 6.69 (d, 1H, J = 2.0 Hz, H-6) (Cyanidin moiety); δ 5.37 (d, 1H, J = 7.8 Hz, H-1′′), 4.44 (m, 2H, H-6′′), 3.80 (m, 1H, H-5′′), 3.51 (t, 1H, H-2′′), 3.40 (t, 1H, H-3′′), 3.35 (d, 2H, J = 2.2 Hz, H-2′′′), 3.23 (t, 1H, H-4′′) (Malonylglucose moiety). ^{13}C NMR (125 MHz, CD_3OD/TFA-d_1 (9: 1, v/v)) δ 168.6 (C-7), 162.0 (C-2), 157.7 (C-5), 156.4 (C-9), 154.6 (C-4′), 146.3 (C-3′), 144.4 (C-3), 134.7 (C-4), 127.2 (C-6′), 119.8 (C-1′), 117.6 (C-2′), 117.0 (C-5′), 112.0 (C-10), 102.6 (C-6), 94.4 (C-8) (Cyanidin moiety); δ 168.2 (C-3′′′), 167.3(C-1′′′), 101.9 (C-1′′), 76.3 (C-3′′), 74.5 (C-5′′), 73.1 (C-2′′), 70.1 (C-4′′), 64.6 (C-6′′), 41.4 (C-2′′′) (Malonylglucose moiety).

2.6 RNA extraction and RNA sequencing

Frozen leaf material (100 mg) was ground in liquid nitrogen, and RNA was extracted using Trizol (Invitrogen, Carlsbad, CA, USA) and the Qiagen RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturers' instructions. RNase-free DNase I was used to remove genomic DNA. A NanoDrop ND-1000 spectrometer was used to measure the quality and quantity of RNA. RNA was run on an Agilent 2100 Bioanalyzer to determine RNA integrity. RNA integrity number (RIN) values were > 8.

Six RNA-Seq libraries were created from three biological replicates using a TruSeq RNA Library Prep Kit (Illumina Inc. San Diego, CA, USA) using SEEDERS (Daejeon, Republic of Korea) and processed on an Illumina HiSeq X system. In the pre-processing steps, adapter sequences were removed using Cutadapt and then trimmed using the DynamicTrim and LengthSort of the SolexaQA package (Cox et al., 2010; Martin, 2011). RNA-Seq data were deposited in the NCBI Sequencing Read Archive public database with Accession No number PRJNA626533, consisting of six libraries (Accession No. SRR11585851 to SRR11585856).

2.7 *De novo* transcriptome assembly and read mapping

The cleaned reads were assembled according to the manufacturer's protocol using the Trinity program (v2.8.6) (https://github.com/trinityrnaseq/trinityrnaseq/wiki) (Grabherr et al., 2011). In addition, read counts were calculated by mapping clean reads, which were obtained as the total mapping reads of each transcript, using Bowtie2 (v2.1.0) software (Langmead and Salzberg, 2012). Mismatches were ≤ 2 bp. After read mapping, the number of cleaned reads was normalized using the DESeq library in R to avoid bias due to differences in coverage depth between the genes (Anders and Huber, 2010).

2.8 Functional annotation and differentially expressed genes (DEGs) analysis

To characterize the functions of the genes, annotation was conducted using amino acid sequences of Viridiplantae DB from NCBI NR and BLASTX (e-value ≤ 1e-10). InterProscan fulfilled the

default criteria using the tools offered by EMBL. To determine the DEGs, fold change (FC) in gene expression and a bionomical test (adjusted p-value ≤ 0.01) were used. In this study, we selected DEGs with 2-fold or greater changes ($\log_2$FC ≥ 1 for upregulation and $\log_2$FC ≤ -1 for downregulation). All DEGs were subjected to gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses using in-house scripts.

2.9 Taxonomic analysis

The datasets of translated representative transcripts were processed by aligning the contigs against the NCBI NR protein database, which was provided on Oct. 18, 2022, with the default parameter (E-value < 1e-30). In addition, the metagenomic community was visualized using Krona (version 2.8.1) (Ondov et al., 2011). Further, the major transcripts related to phenylpropanoid biosynthesis were translated and compared with the proteome sequences of six species of Asteraceae. The same E-value threshold (1e-30) was applied to remove spurious hits.

2.10 Statistical analysis

Data are expressed as mean ± standard deviation of three biological replicates. The two groups were compared using unpaired Student's t-test. Differences were considered statistically significant at $p < 0.05$.

3 Results

3.1 Identification of phenolic compounds and drought stress-induced effects on metabolic events

We investigated the response of *L. fischeri* plants to drought by not watering them for four weeks. As shown in Figure 1A, the color change (from green to purple) was the most apparent feature of the plant. We determined the phytochemical changes in *L. fischeri* leaves under drought stress by analyzing leaf extracts using HPLC-MS analysis (Figure 1B). Preliminary identification of constituents in the leaf extract was conducted based on the UV-Vis absorption chromatogram (λ_{max} value) and fragmentation pattern using ESI-MS in the full scan mode (Table 1). Eight peaks (peaks **1**–**8**) and two peaks (peaks **9** and **10**) were observed at 330 and 520 nm, respectively. These compounds were identified by comparing their spectroscopic data with those of previous *L. fischeri* studies (Kim et al., 2012). According to the reported MS data, peaks **1**, **5**, **6**, and **7** were classified as derivatives of hydroxycinnamic acid (λ_{max} 330 nm) and identified as CQA, 3,4-di-*O*-caffeoylquinic acid, 3,5-di-*O*-caffeoylquinic acid, and 4,5-di-*O*-caffeoylquinic acid, respectively. Peaks **2** and **8** shared the same carbon skeletons of typical flavonoids, with molecular ions at *m/z* 464 and 507 in the positive mode, respectively. These two peaks were determined to be hyperoside (λ_{max} 256 and 357 nm) and 2″-acetylhyperoside (λ_{max} 256 and 353 nm), respectively (Piao et al., 2009; Park et al., 2016).

Peaks 3 and 4 were significantly induced in drought-treated samples, and their MS data have not been reported for this plant until now. Therefore, these compounds were isolated using chromatographic methods and subjected to NMR analysis to determine their chemical structures. Based on the ^{1}H and ^{13}C NMR spectral data, peaks 3 and 4 shared the same carbon skeleton of luteolin as the aglycon moiety of the flavonoids (Figure 1C) (Lee et al., 2011; Guvenalp et al., 2015). According to the ^{13}C NMR data, peak **3** contains six extra carbons, except for the luteolin moiety. Its chemical shifts of ^{1}H and ^{13}C NMR indicated a glucosyl moiety (δ_H 5.08 and 3.15–3.71/δ_C 100.2, 77.6, 76.8, 73.5, 69.9, and 61.0) attached to the C-7 position. The glycone part was confirmed by ESI-MS in the positive mode with molecular ions at *m/z* 449 $[M + H]^+$. Meanwhile, the ^{13}C NMR data of peak 4 were slightly different from those of peak **3** in that there were seven extra carbons. The chemical shifts of ^{1}H and ^{13}C NMR suggested that a glucuronosyl moiety (δ_H 5.08, 3.26–3.60, and 13.01/δ_C 170.6, 99.8, 76.0, 75.8, 73.2, and 71.7) was attached to the same position of peak 3 with molecular ions at m/z 463 $[M + H]^+$. Thus, peaks 3 and 4 were identified as luteolin-7-*O*-glucoside and luteolin-7-*O*-glucuronide, respectively (Fecka and Turek, 2008).

As shown in Figure 1C, peaks 9 and 10 corresponding to two anthocyanins were purified and their chemical structures were characterized using the same procedures described above. In the positive ion mode, the molecular formulae of these compounds were confirmed by ESI-MS with molecular ions at *m/z* 450 and 536 $[M + H]^+$, respectively. With the analysis of ^{1}H and ^{13}C NMR spectral data, the characteristic signals of cyanidin aglycon at the C-4 position were observed with chemical shifts of δ_C 137.9 and δ_H 9.06 for peak 9, and δ_C 134.7 and δ_H 8.79 for peak 10. In the glycone part, peak 9 had the same glucosyl group attached to a different position (C-3) than that of peak 3 (C-7). Signals of one anomeric proton and the other methines were confirmed as a sugar moiety with the range of δ_H 3.40–5.33 for peak 9. In contrast, the remaining carbons (δ_C 168.2, 167.3, and 41.4) of peak 10 indicated a substitution with 6′′-malonylglucoside at the C-3 position. Moreover, the molecular ions detected at *m/z* 450 and 536 $[M + H]^+$ suggest that peaks 9 and 10 were cyanidin-3-*O*-glucoside and cyanidin-3-*O*-(6′′-malonylglucoside), respectively (Narayan and Venkataraman, 2000).

The coupling constants of anomeric protons of each peak (δ_H 5.08, J = 7.3 Hz for peaks 3 and **4**/δ_H 5.33, J = 7.8 Hz for peak 9/δ_H 5.37, J = 7.8 Hz for peak 10) indicated that all of the glycone units were β-forms. Peaks 3, 4, 9, and 10 were identified as luteolin-7-*O*-β-glucoside, luteolin-7-*O*-β-glucuronide, cyanidin-3-*O*-β-glucoside, and cyanidin-3-*O*-β-(6′′-malonylglucoside), respectively, and were first isolated from *L. fischeri*.

According to our previous work, the major phenolic compounds in *L. fischeri* are hydroxycinnamic acid derivatives such as 5-*O*-caffeoylquinic acid and di-caffeoylquinic acids. These compounds accounted up to 9% of the phenolic compounds in the dried leaves of *L. fischeri* (Kim et al., 2012). In the present study, the major hydroxycinnamic acid derivatives in *L. fischeri* were downregulated after drought treatment (Table 2). Moreover, the content of two flavonols, hyperoside and 2″-acetylhyperoside, was decreased in

10.3389/fpls.2023.1140509

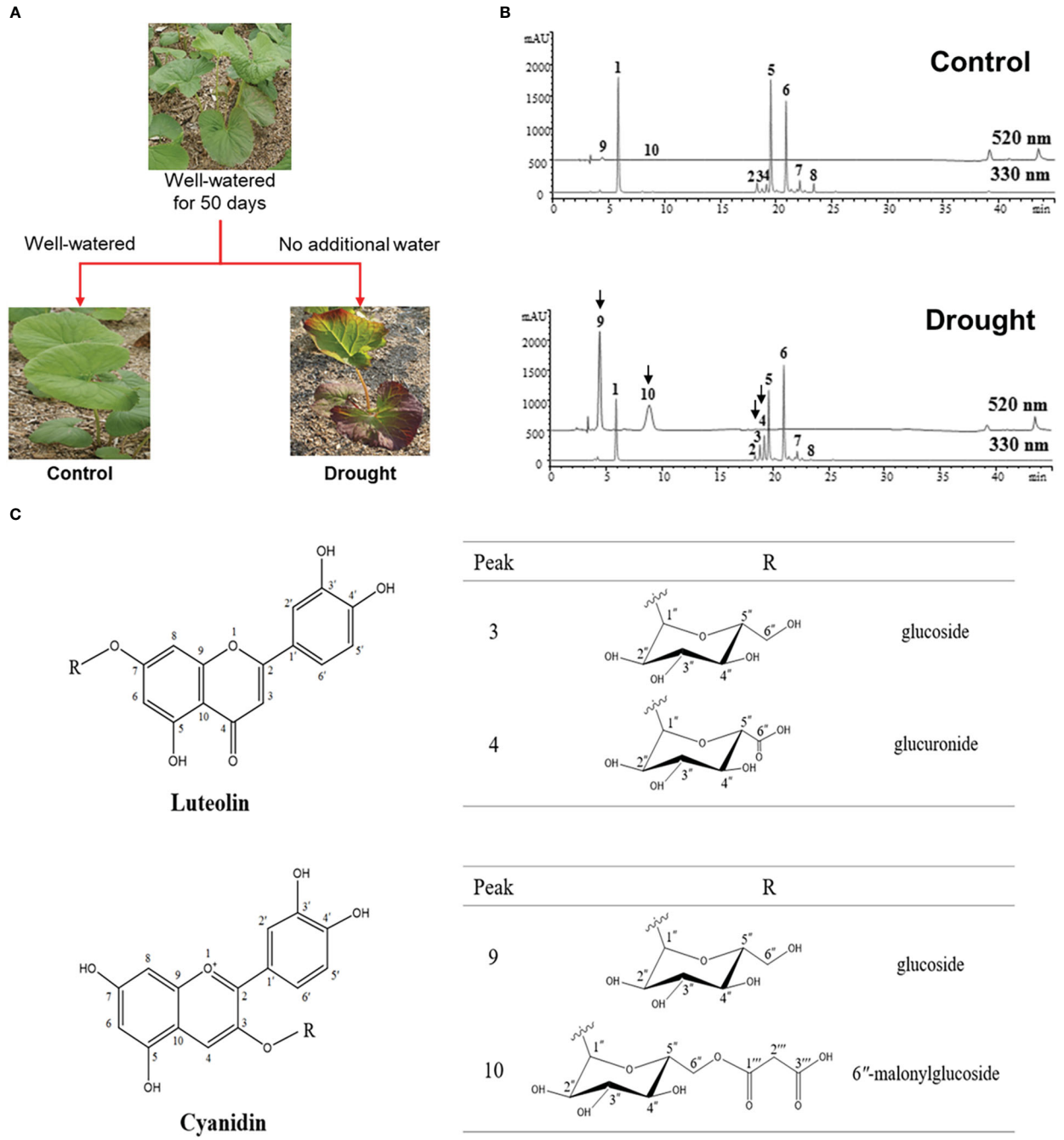


FIGURE 1
(A) Experimental scheme of phenotypes of *L. fischeri* under drought stress **(B)** Representative HPLC chromatograms of leaf extracts from control and drought treatments, monitored at 330 nm and 520 nm. Peak 1, 5-*O*-Caffeolyquinic acid; 2, Hyperoside; 3, Luteolin-7-*O*-β-glucoside; 4, Luteolin-7-*O*-*β*-glucuronide; 5, 3,4-di-*O*-Caffeoylquinic acid; 6, 3,5-di-*O*-Caffeoylquinic acid; 7, 4,5-di-*O*-Caffeoylquinic acid; 8, 2″-Acetylhyperoside; 9, Cyanidin-3-*O*-*β*-glucoside; and 10, Cyanidin-3-*O*-*β*-(6″-malonylglucoside). **(C)** Chemical structures of flavones and anthocyanins newly isolated from *L. fischeri* leaves.

drought-treated *L. fischeri* leaves. However, the newly identified flavones, luteolin-7-*O*-*β*-glucoside (peak 3) and luteolin-7-*O*-*β*-glucurunide (peak 4), were increased 2-fold and 3.34-fold, respectively, under drought stress. The content of two anthocyanins, including cyanidin-3-*O*-*β*-glucoside (peak 9) and cyanidin-3-*O*-*β*-(6″-malonylglucoside) (peak 10) was also significantly upregulated, corresponding to the color change. These appear to be the main changes brought on by drought stress in this plant.

3.2 Sequence generation and transcript assembly

For transcriptome analysis, we generated cDNA libraries of the control and drought groups using Illumina RNA Sequencing. *L. fischeri* leaves were sequenced in triplicate, and the raw reads are available online in the SRA database at NCBI with the accession number PRJNA626533. Clean reads, ranging from 13.55 to 17.33 million, were obtained by pre-processing with an average of 93.18 to

TABLE 1 UV/Vis and ESI/MS data of phenolic compounds in *L. fischeri* leaves.

Peak	Compound	UV/Vis (λmax, nm)	$[M+H]^+$ (*m/z*)
1	5-*O*-Caffeoylquinic acid	326, 298	355
2	Hyperoside	357, 256	465
3	Luteolin-7-*O*-*β*-glucoside	348, 260	449
4	Luteolin-7-*O*-*β*-glucuronide	348, 260	463
5	3,4-di-*O*-Caffeoylquinic acid	330, 232	517
6	3,5-di-*O*-Caffeoylquinic acid	330, 246	517
7	4,5-di-*O*-Caffeoylquinic acid	330, 246	517
8	2″-Acetylhyperoside	353, 256	507
9	Cyanidin-3-*O*-*β*-glucoside	517, 282	450
10	Cyanidin-3-*O*-*β*-(6″-malonylglucoside)	517, 282	536

94.24% read mapping efficiency. The number of reads generated for each sample is shown in Supplementary Table 1. Clean reads from the samples were used to generate the assembled transcriptome of *L. fischeri* using the Trinity program. The total assembled transcripts consisted of 74,873 transcripts, with an N50 value of 1,583 bp (Supplementary Table 2). Of these, 35,084 transcripts are indicated as representative, with an N50 value of 1,424 bp. The assembled transcripts ranged in length from 500 kb to > 14 kb. Prior to DEG analysis, a total of 4,122 DEGs were identified as being greater than the thresholds (p-value ≤ 0.001, $\log_2$FC ≥ 1 or ≤ -1) after read normalization. Overall, differential gene expression in the drought-treated and control groups revealed 1,896 upregulated and 2,792 downregulated transcripts. Of these, 1,514 up-regulated and 2,608 down-regulated genes were annotated using the NR public database (Supplementary Figure 1A). All DEGs were extracted from the expression matrix, and the results are shown in a volcano plot (Supplementary Figure 1B).

3.3 Functional enrichment analysis of DEGs

Based on recent advances in drought stress signaling in plants, we selected numerous putative drought-responsive genes and classified them into several regulatory processes, such as abscisic acid (ABA) signaling, phosphorylation, osmotic stress sensing, MAPK and calcium signaling, transcription factors regulating gene expression during drought stress, epigenetic regulation, and drought stress-inducible genes encoding functional proteins (Supplementary Table S3) (Takahashi et al., 2018). A total of 75 protein sequences encoded by the drought-responsive genes of Arabidopsis were searched against representative transcripts of *L. fischeri* using BLASTP (E-value < 1e-30). As a result, 2,105 hits for 516 distinct transcripts were identified, and their distribution of expression levels showed intricate patterns, ranging across up- and downregulation under drought stress in *L. fischeri* (Figure 2).

To understand the function of the genes affected by drought stress, DEGs were classified into functional groups using GO analysis based on a corrected p-value of 0.05. Among the various GO classifications, we isolated the specific subcategories "response to water deprivation" and "flavonoid metabolic process" from the main functional category, biological processes. The term "response to water deprivation" includes any process that induces a change in the state or activity of a cell or organism through enzyme production, gene expression, movement, and secretion, as a result of a water scarcity stimulus. The flavonoid biosynthetic pathway is enhanced or suppressed by drought stress in *L. fischeri*. As shown in Figure 3, the

TABLE 2 Contents of phenolic compounds in *L. fischeri* under drought stress (mg/g dry weight).

Subgroup	Compound	Control	Drought
Hydroxycinna-mic acid	5-*O*-Caffeoylquinic acid	16.46 ± 0.87	7.26 ± 0.25**
	3,4-di-*O*-Caffeoylquinic acid	13.83 ± 0.49	7.70 ± 0.21**
	3,5-di-*O*-Caffeoylquinic acid	18.31 ± 0.10	14.09 ± 0.32**
	4,5-di-*O*-Caffeoylquinic acid	7.16 ± 0.34	3.05 ± 0.10**
Flavonol	Hyperoside	4.38 ± 0.15	2.33 ± 0.08**
	2″-Acetylhyperoside	3.72 ± 0.16	0.32 ± 0.03**
Flavone	Luteolin-7-*O*-*β*-glucoside	1.03 ± 0.03	2.06 ± 0.04**
	Luteolin-7-*O*-*β*-glucuronide	2.86 ± 0.22	9.54 ± 0.34**
Anthocyanin	Cyanidin-3-*O*-*β*-glucoside	0.09 ± 0.01	2.79 ± 0.16*
	Cyanidin-3-*O*-*β*-(6″-malonylglucoside)	0.07 ± 0.02	1.82 ± 0.08**

Each value represents the mean ± SD of triplicate experiments; the control and drought groups were compared using an unpaired Student's t-test (* $p < 0.05$; ** $p < 0.01$).

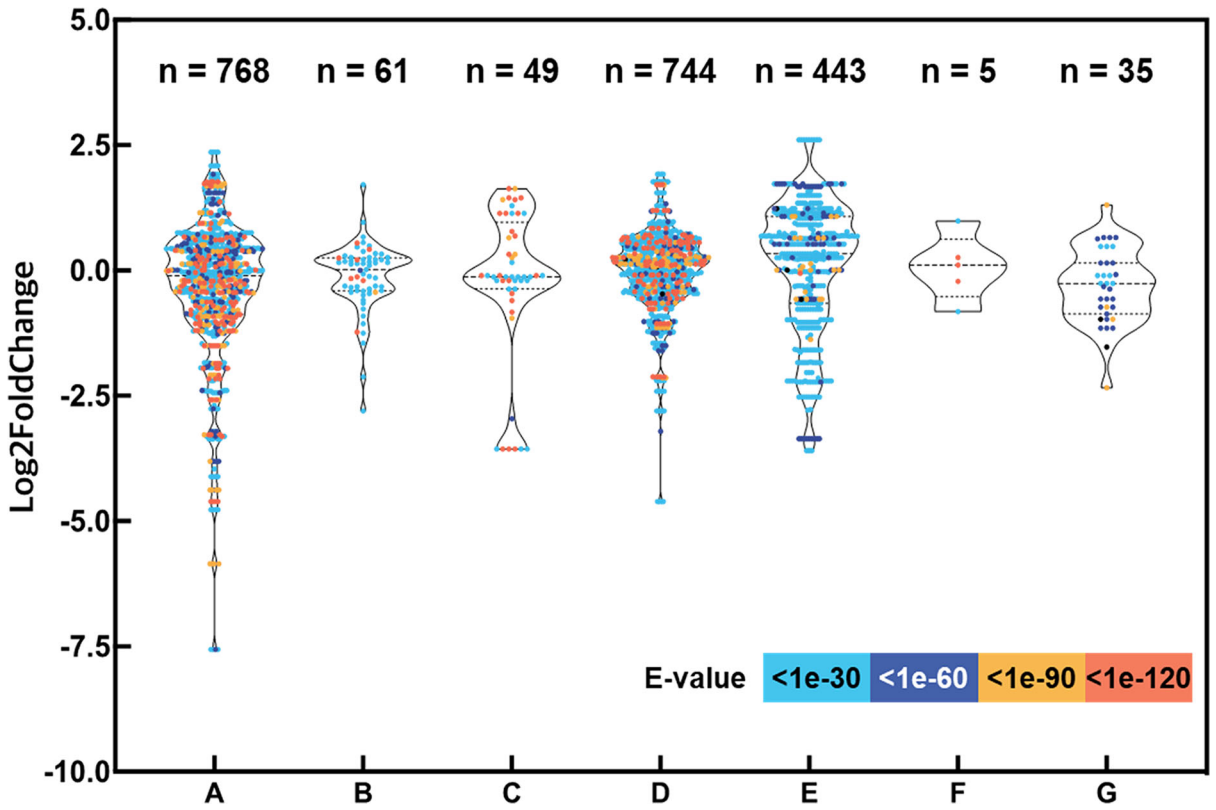


FIGURE 2
Expression levels of transcripts aligned with the drought-responsive genes by the seven categories and evaluation of the homology between the protein sequences encoding drought-responsive genes. The y-axis shows log2 fold change of each transcript and the x-axis indicates the gene categories shown in Supplementary Table S3: **(A)** ABA signaling, **(B)** Phosphorylation, **(C)** Osmotic stress sensing, **(D)** MAPK and Calcium signaling, **(E)** Transcription factors regulating gene expression during drought stress, **(F)** Epigenetic regulation, and **(G)** Drought stress-inducible genes encoding functional proteins. The number of hits (E-value < 1e-30) are shown at the top of the figure. The dots are displayed in different colors based on E-values from the homology searches.

number of downregulated genes outnumbered the number of upregulated genes in both the groups. Specifically, the "response to water deprivation" gene set included 38 upregulated and 76 downregulated genes. Moreover, the "flavonoid metabolic process" gene set contained 23 upregulated and 65 downregulated genes, further confirming the metabolic changes observed in *L. fischeri* under drought stress. The gene information for the DEGs associated with these subcategories is described in Supplementary Table S4.

To better explain the metabolic functions of the DEGs in *L. fischeri* leaves under drought stress, KEGG enrichment analysis was performed using the KEGG database and BLASTX, where the e-value of the filter standard was ≤ 1e-10. Based on the number of assigned DEGs, an enrichment analysis of KEGG terms related to DEGs was performed to further understand the potential metabolic roles in secondary metabolism in *L. fischeri* under drought stress. DEGs responsible for phenylpropanoid biosynthesis were most abundant in both up- (21) and downregulated (59) DEGs (Table 3). In addition, many DEGs were annotated for flavonoid biosynthesis, including 33 down-regulated and 12 up-regulated DEGs. These results are consistent with our findings that the contents of hydroxycinnamic acids and some flavonoids decrease under drought stress.

3.4 DEGs in phenylpropanoid biosynthesis

As described above, various functional metabolites were significantly altered due to drought stress in *L. fischeri*. We observed that the contents of hydroxycinnamic acids and flavonols were downregulated, whereas the newly identified flavones and anthocyanins were significantly increased under drought stress. The principal cause of drought-induced phenolic compound production is regulation of the phenylpropanoid biosynthetic pathway. Drought modulates numerous essential genes responsible for phenylpropanoid biosynthesis and drought-induced changes in gene expression. Therefore, we isolated DEGs-related to phenylpropanoid biosynthesis based on the BLAST tool (TBLASTN) with an e-value cutoff of $1 \times e^{-10}$ and identity ≥ 40%, and examined the distinct changes in *L. fischeri* subjected to drought stress. As shown in Table 4, we identified 24 DEGs encoding multiple enzymes associated with phenylpropanoid biosynthesis. This included nearly all of the enzymes required in the general pathway, including phenylalanine ammonia-lyase (*PAL*), cinnamate-4-hydroxylase (*C4H*), 4-coumarate:CoA ligase 1 (*4CL1*), chlacone synthase (*CHS*), chalcone isomerase (*CHI*), flavone synthase (*FNS*), flavanone 3-hydroxylase (*F3H*), flavonoid 3' -hydroxylase (*F3'H*), anthocyanin 5-*O*-

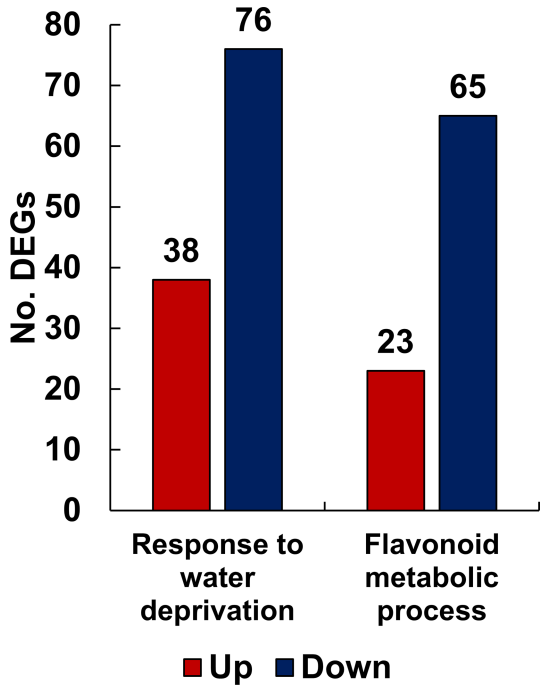


FIGURE 3
The number of up- and down-regulated DEGs related to response to water deprivation and flavonoid metabolic process functionally categorized by GO analysis from *L. fischeri* under drought stress. The x-axis represents the number of DEGs and the y-axis indicates each subcategory's information. In bar graphs, red and blue colors are shown as up- and down-regulated DEGs, respectively.

glucosyltransferase (*A5GT*), 4-coumarate 3-hydroxylase (*C3H*), shikimate *O*-hydroxycinnamoyltransferase (*HCT*), and hydroxycinnamoyl-CoA quinate/shikimate transferase (*HQT*).

A schematic overview of the phenylpropanoid biosynthetic pathways and transcript levels of the genes involved in *L. fischeri* is shown in Figure 4. Also, fragments per kilobase of transcript per million mapped reads (FPKM) values obtained by RNA-Seq analysis were shown in Supplementary Figure 2. Among the 24 DEGs, 8 DEGs were upregulated, while 16 were downregulated under drought stress. In the initial step of phenylpropanoid biosynthesis, two DEGs of *LfPAL* (TRINITY_DN299_c0_g1_i1; log_2FC=-0.89 and TRINITY_DN299_c0_g2_i1; log_2FC=-2.00) were downregulated in *L. fischeri*. Moreover, one gene encoding *LfC4H* (TRINITY_DN22101_c0_g1_i2, log_2FC=0.74) was specifically upregulated, whereas one gene encoding *Lf4CL* (TRINITY_DN399_c0_g2_i1, log_2FC=-1.85) was downregulated under drought stress. Subsequently, we found that DEGs encoding *LfCHS* (TRINITY_DN612_c0_g1_i1, log_2FC=0.44) and *LfCHI* (TRINITY_DN772_c0_g2_i3, log_2FC=-3.80) were upregulated and downregulated in *L. fischeri* under drought stress, respectively. We isolated one gene encoding *LfFNS* (TRINITY_DN59627_c0_g1_i1, log_2FC=1.43) that was significantly upregulated. However, three *LfF3H* (TRINITY_DN26907_c0_g1_i1; log_2FC=-1.24, TRNITY_DN505_c0_g1_i1, log_2FC=-1.40, and TRINITY_DN505_c0_g2_i1; log_2FC=-1.44) and four *LfF3'H* (TRINITY_DN3519_c0_g1_i1; log_2FC=-1.22, TRINITY_DN6034_c0_g2_i4; log_2FC=-1.29, TRINITY_DN1861_c0_g2_i2; log_2FC=-2.13, and TRINITY_DN81_c0_g1_i1; log_2FC=-2.16) genes were downregulated under drought stress. Interestingly, two DEGs encoding *LfA5GT* were found to be upregulated (TRINITY_DN782_c0_g1_i1, log_2FC=1.36) and downregulated (TRINITY_DN12729_c0_g1_i1, log_2FC=-4.03) in *L fischeri*. Among the mapped DEGs related to hydroxycinamic acid biosynthesis, we obtained three genes encoding *LfC3H* (TRINITY_DN2611_c1_g1_i5; log_2FC=1.30, TRINITY_DN6034_c0_g2_i4; log_2FC=-1.29, and TRINITY_DN11373_c0_g3_i1; log_2FC=-3.88), one gene encoding *LfHCT* (TRINITY_DN31661_c0_g1_i1, log_2FC=-3.99), and four genes encoding *LfHQT* (TRINITY_DN5048_c0_g1_i1; log_2FC=1.07, TRINITY_DN12670_c0_g1_i1, log_2FC=1.29, TRINITY_DN430_c0_g1_i6; log_2FC=1.07, and TRINITY_DN15180_c0_g1_i1; log_2FC=-2.35) in *L. fischeri*.

TABLE 3 KEGG pathway annotation and number of DEGs in biosynthesis of other secondary metabolites.

KEGG pathway item	Up	Down
Caffeine metabolism	1	0
Monobactam biosynthesis	2	4
Phenylpropanoid biosynthesis	21	59
Flavonoid biosynthesis	12	33
Anthocyanin biosynthesis	0	1
Isoflavonoid biosynthesis	2	4
Flavone and flavonol biosynthesis	0	5
Stilbenoid, diarylheptanoid and gingerol biosynthesis	9	18
Isoquinoline alkaloid biosynthesis	1	3
Tropane, piperidine and pyridine alkaloid biosynthesis	1	3
Glucosinolate biosynthesis	3	4
Biosynthesis of various secondary metabolites - part 2	0	2

3.5 Taxonomic analysis of *L. fischeri* transcriptome

Translated sequences of 35,084 representative transcripts of *L. fischeri* were aligned against the NCBI NR protein database (released on Oct. 18, 2022) using default parameters. The taxonomic profile of the BLASTP results was visualized using the Krona Tools (v2.8.1) (Ondov et al., 2011). The majority of the hits (72%) were found in plant species, especially the ones belonging to the family Asteraceae (69%) (Figure 5A). In addition, various Asteraceae plants, such as *Smallanthus sonchifolius* (14%), *Erigeron canadensis* (8%), *Arctium lappa* (8%), *Cynara cardunculus* var. *scolymus* (8%) and *Artemisia annua* (7%) showed the highest taxonomic levels based on the BLASTP results against the NCBI NR database (Figure 5B). Of these, six species, *A. annua*, *C. cardunculus*, *E. canadensis*, *Helianthus annuus*, *Lactuca sativa*, and *Mikania micrantha*, which are frequently found in the BLASTP results, were selected to further investigate the distribution of key transcripts across the family Asteraceae (Figure 4 and Supplementary Table S5) (Scaglione et al., 2016; Badouin et al., 2017; Reyes-Chin-Wo et al., 2017; Shen et al., 2018; Laforest et al., 2020; Liu et al., 2020).

10.3389/fpls.2023.1140509

TABLE 4 List of genes related to phenylpropanoid biosynthesis in *L. fischeri*.

Category	Gene ID	Symbol	Uniprot ID	Identity (%)	FC*
Upstream	TRINITY_DN299_c0_g1_i1	*LfPAL1*	AT2G37040	83.10	-1.89
	TRINITY_DN299_c0_g2_i1	*LfPAL2*		82.88	-2.00
	TRINITY_DN22101_c0_g1_i2	*LfC4H*	AT2G30490	40.00	0.74
	TRINITY_DN399_c0_g2_i1	*Lf4CL1*	AT1G51680	70.42	-1.85
Flavone	TRINITY_DN612_c0_g1_i1	*LfCHS*	AT5G13930	85.71	0.44
	TRINITY_DN772_c0_g2_i3	*LfCHI*	AT3G55120	59.36	-3.80
	TRINITY_DN59627_c0_g1_i1	*LfFNS*	AT5G07990	43.24	1.43
Flavonol	TRINITY_DN26907_c0_g1_i1	*LfF3H1*	AT3G51240	90.91	-1.24
	TRINITY_DN505_c0_g1_i1	*LfF3H2*		79.30	-1.40
	TRINITY_DN505_c0_g2_i1	*LfF3H3*		77.03	-1.44
	TRINITY_DN3519_c0_g1_i1	*LfF3'H1*	AT5G07990	70.11	-1.22
	TRINITY_DN6034_c0_g2_i4	*LfF3'H2*		46.93	-1.29
	TRINITY_DN1861_c0_g2_i2	*LfF3'H3*		42.29	-2.13
	TRINITY_DN81_c0_g1_i1	*LfF3'H4*		41.20	-2.16
Anthocyanin	TRINITY_DN782_c0_g1_i1	*LfA5GT1*	AT4G14090	42.64	1.36
	TRINITY_DN12729_c0_g1_i1	*LfA5GT2*		40.18	-4.03
Hydroxycinnamic acid	TRINITY_DN2611_c1_g1_i5	*LfC3H1*	AT2G40890	46.55	1.30
	TRINITY_DN6034_c0_g2_i4	*LfC3H2*		40.57	-1.29
	TRINITY_DN11373_c0_g3_i1	*LfC3H3*		43.56	-3.88
	TRINITY_DN31661_c0_g1_i1	*LfHCT*	AT5G48930	79.62	-3.99
	TRINITY_DN5048_c0_g1_i1	*LfHQT1*	AT5G48930	57.24	1.07
	TRINITY_DN12670_c0_g1_i1	*LfHQT2*		59.18	1.29
	TRINITY_DN430_c0_g1_i6	*LfHQT3*		56.16	1.07
	TRINITY_DN15180_c0_g1_i1	*LfHQT4*		67.19	-2.35

*Log_2FoldChange. PAL, phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylase; 4CL1, 4-coumarate:CoA ligase 1, CHS, chlacone synthase; CHI, chalcone isomerase; FNS, flavone synthase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; A5GT, anthocyanin 5-O-glucosyltransferase; C3H, 4-coumarate 3-hydroxylase; HCT, shikimate O-hydroxycinnamoyltransferase; HQT, hydroxycinnamoyl-CoA quinate/shikimate transferase.

The key transcripts shown in Figure 4 were translated and searched against the proteome sequences of the six selected species belonging to the family Asteraceae using BLASTP. The same E-value criterion (< 1e-30) was used to filter out spurious hits (Table 5 and Supplementary Tables S5). Interestingly, two transcripts were found to be *L. fischeri*-specific at the amino acid sequence level. For the transcript TRINITY_DN22101_c0_g1_i2 encoding *C4H*, there was only one hit from BLASTN searches, showing an alignment length of 297 bp found in *C. cardunculus* genome (E-value = 4.10e-74). For the transcript TRINITY_DN26907_c0_g1_i1 encoding *F3H*, BLASTN hits with short alignment lengths ranging from 113 bp to 126 bp were found in the genome sequences of *A. annua*, *E. canadensis*, *H. annuus*, and *M. micrantha*. This suggests that genes corresponding to the two transcripts are drought (or stress)-responsive and regulate the phenylpropanoid biosynthetic pathway. Notably, only one or two BLASTP hits were found in the proteomes of the six species for the translated transcript TRINITY_DN31661_c0_g1_i1. Thus, HCT plays a pivotal role in regulating the production of CQA and its derivatives in Asteraceae species.

4 Discussion

4.1 Metabolic response to drought stress in *L. fischeri*

In the present study, newly isolated flavones such as luteolin-7-*O*-*β*-glucoside (peak 3) and luteolin-7-*O*-*β*-glucuronide (peak 4), and anthocyanins, including cyanidin-3-*O*-*β*-glucoside (peak 9) and cyanidin-3-*O*-*β*-(6″-malonylglucoside) (peak 10), were considerably increased under drought stress in *L. fischeri*, while the most common phenolic compounds, called CQAs and flavonol content, were decreased. Compared to earlier findings, different metabolic responses have been revealed depending on plant, treatment. For instance, *Agave salmiana* under *in vitro* drought stress had the lowest

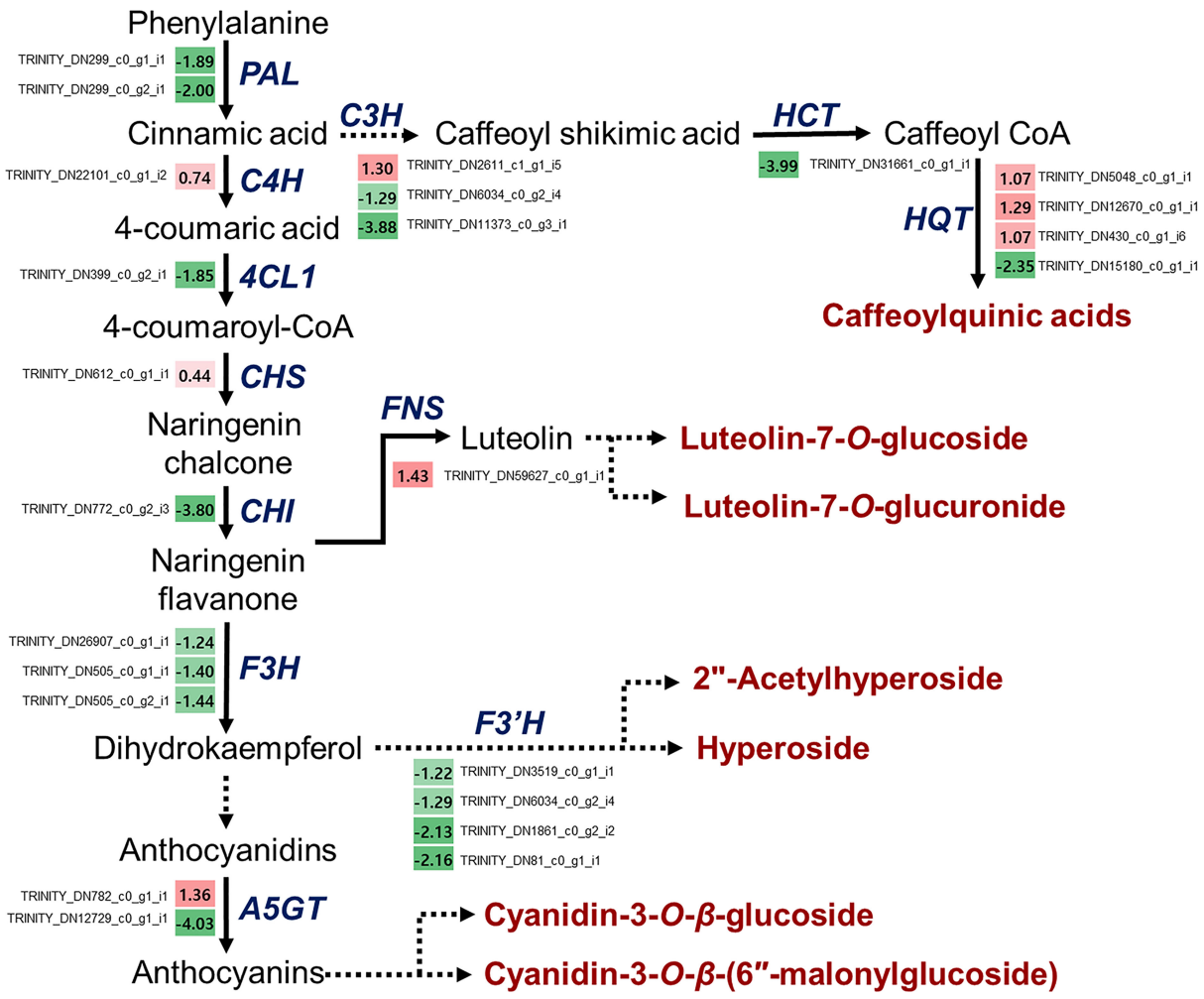


FIGURE 4
Drought-responsive genes in the phenylpropanoid biosynthesis from *L. fischeri*. A color gradient from low (green) to high (red) represents the relative levels of expression. The numbers in the squares indicate FC in gene expression in drought-induced *L. fischeri* plants, compared to control.

levels of flavonols such as kaempferol glycosides and quercetin glycosides (Puente-Garza et al., 2017). Most of the polyphenols including luteolin-7-O-glycoside, 1,3-dicaffeoylquinic acid were upregulated with increasing drought stress duration in *Achillea pachycephala* Rech. f (Gharibi et al., 2019). These were similar to our finding. In contrast, the higher levels of flavonols were indicated under extreme drought stress in Arabidopsis. It has been investigated that flavonoids response to drought stress are variable, and the severity and length of the drought stress may have a significant impact on the types, quantities and localization of flavonoids in response to various levels of water shortage (Shojaie et al., 2016). The accumulation of phenolic acids and flavonoids is crucial for mitigating the deleterious effects of drought stress in plants. Enhanced levels of these compounds act as antioxidants and protect plants from detrimental effects of water scarcity. For instance, flavonoids, including kaempferol and quercetin, increased in tomato plants, similar to drought resistance. Flavonoid production in the cytoplasm may detoxify the detrimental H_2O_2 molecules caused by drought stress (Sharma et al., 2019). Additionally, flavonoid levels increased and showed that a water deficit had an impact on flavonoid accumulation, possibly by controlling hormone metabolism (Yuan et al., 2012). In our previous study, we observed that sunlight increased the total phenolic content in four weeks, whereas the total flavonoid content showed no significant change (Kim et al., 2012). Here, *L. fischeri* plants were grown under identical sunlight conditions and the water supply was removed. Compared to our present findings, the biosynthesis of hydroxycinnamic acids, such as CQA and its derivatives, may be positively affected by sunlight but negatively by drought stress.

Plant cultivars with high anthocyanin levels tend to be more tolerant to drought conditions (Cirillo et al., 2021). Drought reduces grain weight and substantially decreases anthocyanin, protein, and carbohydrate content in *Triticum aestivum* L. cv. Guizi 1 is a drought-tolerant purple wheat cultivar (Li et al., 2020). Plant leaves synthesize anthocyanins for adaptive purposes, such as photoprotection, cold hardiness, and antioxidative capabilities (Archana et al., 2020). Drought tolerance is an adaptive response (Egilla et al., 2001). It has been proposed that anthocyanins operate as osmoregulators by controlling water homeostasis in stressed plants. These compounds build up in tissues under water shortage, and tissues with higher anthocyanin levels are more drought-resistant. It is well known that anthocyanins can remove

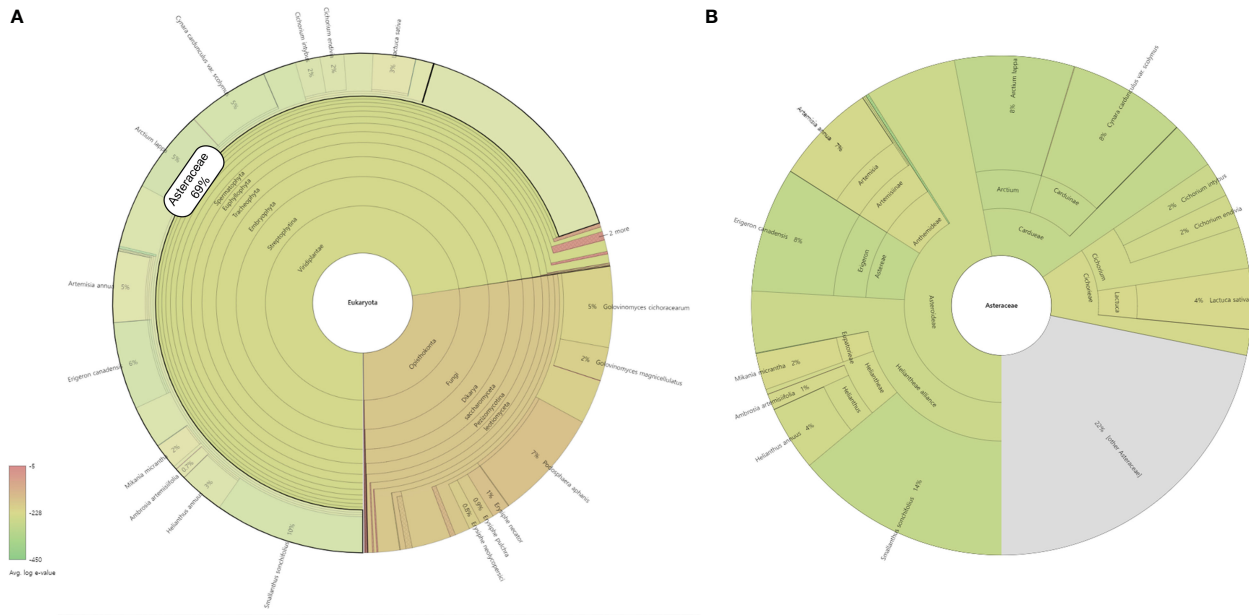


FIGURE 5
A taxonomic profile of BLASTP results against the NCBI NR database with the highest taxon as **(A)** Eukaryota and **(B)** Asteracea. The circles from inside to outside represent different taxonomic rankings, while the area of the sector indicates the fraction of various OUT annotation findings.

reactive oxygen species (ROS) and transport them to vacuoles (Zahedi et al., 2021). Similar to other Asteraceae species, safflower (*Carthamus tinctorius*) responded to drought stress by producing significantly greater levels of total flavonoid, anthocyanins (Toupchi Khosrowshahi et al., 2019). Also, water stress reduced chlorophyll content while increasing flavonoids, anthocyanin, phenolic compounds, and soluble sugars in *Helianthus annuus* and *Silybrum marianum* (Ebrahimian and Bybordi, 2012). Therefore, the pronounced induction of anthocyanins in *L. fischeri* leaves could be viewed as a physiological response to water deficits.

4.2 Drought-induced transcriptome responses classified by functional categories

Transcriptome sequencing is an effective approach to identify the molecular mechanisms underlying plant responses to several abiotic stressors. In a variety of crops, genome-scale transcriptome studies based on Illumina RNA Sequencing have been used to assess gene expression under cold, drought, and heat stresses (Qiu et al., 2017). In addition, identifying DEGs across treatments/conditions is an important phase, and frequently the primary objective, in statistical RNA-Seq data analysis. The discovery of DEGs contributes in the understanding of gene function when cells react to various chemo. Furthermore, finding DEGs can be used as a preliminary to grouping gene expression profiles or assessing gene set enrichments (Kvam et al., 2012). Several studies have been conducted on drought stress in various crops, such as wheat, sweet potato, maize, and rice, using transcriptome analysis (Zhao et al., 2021). To the best of our knowledge, this is the first RNA-Seq study of *L. fischeri* in drought response analysis. Therefore, this study provides important information on drought-responsive processes at the transcriptional level.

Various drought-responsive genes have been identified using RNA-Seq analysis in *L. fischeri*. Most of these genes were found to be involved in phytohormone signaling, MAPK and calcium signaling pathways, transcription factors regulating gene expression, and other drought-inducible regulatory processes. Phytohormones, including ABA, auxin, cytokinin, ethylene, gibberellic acid, and jasmonic acid mediate a variety of activities and enable plants to withstand drought (Wilkinson et al., 2012). ABA is an essential player that controls physiological and molecular reactions to water deficit, including gene expression, osmoprotectants, stomatal closure, and stress protein production (Takahashi et al., 2018). According to our results, eight transcripts (TRINITY_DN22327_c0_g1_i1, TRINITY_DN1659_c0_g1_i3, TRINITY_DN4970_c1_g1_i2, TRINITY_DN14995_c0_g1_i1, TRINITY_DN1678_c0_g1_i2, TRINITY_DN17252_c0_g1_i1, TRINITY_DN73142_c0_g1_i1, and TRINITY_DN15780_c0_g3_i1) matched with 13 query proteins that showed the highest upregulation rates under drought stress in *L. fischeri* (Supplementary Table 6). Interestingly, all of these transcripts were associated with ABA signaling. In addition, many DEGs were annotated and displayed various distributions in the MAPK signaling pathway and transcription factors under drought stress. Upon drought stimulation, phytohormone levels typically increase, activating morphophysiological and other biochemical pathways. These pathways may consist of the calcium signaling pathway, MAPK signaling pathway, regulation of transcription factors, and higher levels of antioxidant enzymes (Ilyas et al., 2021).

4.3 Phenylpropanoid biosynthesis under drought stress

Plants have developed complex molecular, physiological, and biochemical processes to deal with the consequences of drought

10.3389/fpls.2023.1140509

TABLE 5 The number of BLASTP hits (E-value < 1e-30) of the translated transcripts shown in Figure 4.

Transcript	Symbol	Aa*	Cc*	Ec*	Ha*	Ls*	Mm*	Sum
TRINITY_DN299_c0_g1_i1	PAL	6	6	6	9	5	11	43
TRINITY_DN299_c0_g2_i1	PAL	6	7	6	9	5	12	45
TRINITY_DN22101_c0_g1_i2	C4H	0	0	0	0	0	0	0
TRINITY_DN399_c0_g2_i1	4CL	54	28	34	49	33	30	228
TRINITY_DN612_c0_g1_i1	CHS	15	10	11	23	12	13	84
TRINITY_DN772_c0_g2_i3	CHI	7	2	1	3	2	6	21
TRINITY_DN26907_c0_g1_i1	F3H	0	0	0	0	0	0	0
TRINITY_DN505_c0_g1_i1	F3H	124	97	101	143	113	76	654
TRINITY_DN505_c0_g2_i1	F3H	116	94	98	136	110	71	625
TRINITY_DN782_c0_g1_i1	A5GT	182	102	129	234	154	123	924
TRINITY_DN12729_c0_g1_i1	A5GT	213	124	155	293	174	164	1123
TRINITY_DN2611_c1_g1_i5	C3H	63	65	52	99	73	67	419
TRINITY_DN6034_c0_g2_i4	C3H	190	166	204	280	224	198	1,262
TRINITY_DN11373_c0_g3_i1	C3H	183	176	211	303	243	216	1,332
TRINITY_DN59627_c0_g1_i1	FNS	167	164	200	272	209	177	1,189
TRINITY_DN3519_c0_g1_i1	F3'H	321	189	229	341	265	254	1,599
TRINITY_DN6034_c0_g2_i4	F3'H	190	166	204	280	224	198	1,262
TRINITY_DN1861_c0_g2_i2	F3'H	199	176	221	317	249	216	1,378
TRINITY_DN81_c0_g1_i1	F3'H	312	186	228	340	264	252	1,582
TRINITY_DN31661_c0_g1_i1	HCT	1	1	1	1	1	2	7
TRINITY_DN5048_c0_g1_i1	HQT	42	15	32	55	13	29	186
TRINITY_DN12670_c0_g1_i1	HQT	22	10	19	36	7	15	109
TRINITY_DN430_c0_g1_i6	HQT	55	22	42	73	26	38	256
TRINITY_DN15180_c0_g1_i1	HQT	27	10	19	38	8	17	119

Aa (*Artemisia annua*), Cc (*Cynara cardunculus*), Ec (*Erigeron canadensis*), Ha (*Helianthus annuus*), Ls (*Lactuca sativa*), and Mm (*Mikania micrantha*). See Table S1 for more information on the genomes used in this analysis.

(Zhao et al., 2021). Physiological reactions include stomatal closure, decreases in photosynthesis rate, effects on photosynthetic proton and electron transport, and modifications to photosynthetic carbon reduction and carbon oxidation cycles when there is a shortage of water in the soil (Cheng et al., 2018). Additionally, drought stress can affect secondary metabolic systems, and various findings on secondary metabolite formation in medicinal plants under drought stress have been investigated (Yadav et al., 2014; Zahir et al., 2014a; Jia et al., 2015).

Phenolic compounds are the most abundant secondary metabolites in plants, with simple to complex aromatic rings. They are classified into various classes including phenolic acids, flavonoids, stilbenes, and lignans, each of which has unique qualities. Because these compounds are derived from phenylalanine, they are also referred to as phenylpropanoids (Sharma et al., 2019). In the initial step of the phenylpropanoid biosynthetic pathway, *PAL* exists at the interface between the primary and secondary metabolism to convert L-phenylalanine to cinnamic acid (Pappi et al., 2021). The subsequent stages of the pathway, catalyzed by *C4H* and *4CL*, are required and serve as the basis for all further branches and resulting metabolites (Deng & Lu, 2017). In this study, we isolated one gene encoding *LfFNS* which was significantly upregulated under drought stress. This result likely explains the increased flavone content, including luteolin-7-*O*-glucoside and luteolin-7-*O*-glucuronide, in *L. fischeri* under drought stress. In contrast, three *LfF3H* and four *LfF3'H* genes were considerably downregulated under drought stress. This may decrease the content of flavonols, including hyperoside and 2″-acetylhyperoside, in *L. fischeri*. These findings are similar to those of previous studies. For instance, most of the flavonoid concentrations, including apigenin, caffeic acid, chlorogenic acid, luteolin-7-*O*-glycoside, luteolin, kaempferol, 1,3-dicaffeoylquinic acid, and rutin, and the expression of the relevant genes, such as *PAL*, *CHS*, *CHI*, *F3H*, *F3'H*, *F3'5'H*, and *FLS*, also increased after 21 days of exposure to drought in *Achillea pachycephala* Rech.f (Gharibi et al., 2019). Previously, it was also discovered that genes involved in flavonoid production were increased in buckwheat during drought stress (Hou et al., 2019). Anthocyanins are glycosides of anthocyanidins, and their glycosylation is carried out by a variety of enzymes, including UDP-glucose and anthocyanidin 3-*O*-glucosyltransferase (UF3GT), which is the most well-studied flavonoid glycosyltransferase (Deng and Lu, 2017). Based on our results, *LfA5GT* was upregulated under drought stress, which might explain the newly identified and enhanced anthocyanin content in *L. fischeri*. This is also in accordance with previous findings that an increase in the overall anthocyanin, flavonoid, and phenolic content may be closely related to the enhanced levels of flavonoid biosynthetic genes, including *CHS*, *CHI*,

F3H, *FNS*, *FLS*, *DFR*, and *ANS*, in wheat leaves under drought stress (Ma et al., 2014).

The presence of a benzene ring bonded to one or more hydroxyl or methoxy groups distinguishes phenolic acids from aromatic acids. These compounds can be classified into two classes based on their constitutive carbon skeletons: hydroxycinnamic acid and hydroxybenzoic acid. The major compounds in *L. fischeri* and the CQAs belong to the hydrocinnamic acid group. Although phenolic acids are abundantly dispersed in plants, it has been established that they are produced *via* the shikimate pathway; however, they are still unknown and have been revised (Deng and Lu, 2017). Three biosynthetic steps have been suggested for CQA formation in plants. The first involves caffeic acid coenzyme A and quinic acid catalysis by *HQT*. The second involves the production of CQAs by employing caffeoyl glucoside as an active intermediate, which is catalyzed by hydroxyl cinnamoyl D-glucose:quinate hydroxycinnamoyl transferase (*HCGQT*). Finally, CQAs were synthesized from *p*-coumaroyl quinic acid *via* a catalytic reaction involving *HCT* and *C3H*. Notably, the *HQT*-mediated pathway is the most important for CQA biosynthesis among the three pathways (Xu et al., 2022). The transcript levels of two *LfC3H* genes, *LfHCT* and *LfHQT* were downregulated only under drought stress in *L. fischeri*. These findings were consistent with the decreased levels of CQA and its derivatives.

Drought stress has a negative impact on both the quantitative and qualitative aspects of growth, production, and agricultural output. It interferes with the crop's normal physiological functioning by reducing nutrient delivery and causing cellular toxicity. It aslo leads to membrane destabilization, harm to the photosynthetic apparatus, and oxidative stress (Yadav et al., 2021). Plants undergo considerable molecular and physiological changes as a result of drought stress, and global transcriptional regulation is regarded as the most basic molecular response of plants to adapt to and deal with drought stress (Qiu et al., 2017). These changes include the transcriptomic, proteomic and metabolomics modifications of plants, which leads to cellular biosynthesis and breakdown activities. To maintain growth and yield during water shortage, thus, plants activate various strategies such as increased synthesis of secondary metabolites, phytohormones, ROS signaling, plant hydraulic status, and osmotic adjustment (Jogawat et al., 2021). Of these, the biosynthesis of secondary metabolites induced by drought stress has been investigated. Flavonoids and polyphenols are well known as adaptive natural substances that enable to plants to scavenge ROS under drought stress (Treml and Šmejkal, 2016). As a consequence, numerous crop plants were shown to have enhanced biosynthetic pathways associated with their accumulations. A variety of biotechnological and bioinformatics methods, including transcriptomics, metabolomics, proteomics, have also been demonstrated to significantly stimulate these metabolites between drought-toleranct and drought-sensitive varieties or cultivars (Yadav et al., 2021). In previous study, expression level of 4CL, a key gene in the catechin production, was decreased by drought treatment in *Camellia sinensis*. It suggested that increased polyphenol accumulation, including isoflavonoids and catechins, is related to resistance in this plant (Cheruiyot et al., 2008). In potato, expression levels of genes including flavonone-3-hydroxylase, flavonol synthase, and β-carotene synthase, which are crucial for the flavonoids, carotenoids, and other phenolic compounds production, were enhanced during drought stress. The transcript levels of these genes impacts how well some potato cultivars tolerate drought (Min et al., 2008). Moreover, it was discovered that flavonoid biosynthetic genes, as well as genes linked to lignin, were induced in response to drought stress in cotton. This ultimately results in the polyphenols and xanthophylls accumulation (Ranjan et al., 2012). Above this, various plants have indicated varied increases in flavonoids and polyphenols or the gene expression involved in their biosynthesis under water deficit (Yadav et al., 2021). Based on our findings (Table 2 and Figure 4), we suggest that the upregulated *LfFNS* (TRINITY_DN31661_c0_g1_i1) and *LfA5GT1* (TRINITY_DN782_c0_g1_i1), which are associated with the biosynthesis of flavones and anthocyanins, respectively, are candidate drought-responsive genes in *L. fischeri*. In addition, we observed two downregulated genes, *LfHCT* (TRINITY_DN31661_c0_g1_i1) and *LfHQT4* (TRINITY_DN15180_c0_g1_i1), which led to a decrease in CQAs and may play a crucial role in drought response in *L. fischeri*.

In summary, the drought response in *L. fischeri* is a complicated process that involves several morphological and molecular changes at all levels. Two flavones and two anthocyanins were newly isolated and were significantly elevated under drought stress. In contrast, the contents of CQAs and flavonols was decreased. These results indicate that the biosynthesis of CQAs and flavonols was downregulated, whereas flavones and anthocyanin biosynthesis were upregulated in response to drought stress. Furthermore, we used RNA-Seq analysis to assess transcriptional changes to obtain insights into the molecular response to drought stress in *L. fischeri*. A total of 2,105 hits for 516 different transcripts were found to be drought-responsive genes, and their expression levels were distributed in intricate patterns, spanning up- and downregulation under drought stress. Additionally, KEGG enrichment analysis showed that DEGs involved in phenylpropanoid biosynthesis were found in the highest number of both up-and downregulated DEGs. In particular, several DEGs were associated with flavonoid biosynthesis. Based on the regulation of phenylpropanoid biosynthetic genes, we propose that the upregulated *LfFNS* (TRINITY_DN31661_c0_g1_i1) and *LfA5GT1* (TRINITY_DN782_c0_g1_i1) genes are potential drought-responsive genes that correspond to the high concentrations of flavones and anthocyanins in *L. fischeri* under drought stress. Moreover, the downregulated genes, *LfHCT* (TRINITY_DN31661_c0_g1_i1) and *LfHQT4* (TRINITY_DN15180_c0_g1_i1), which resulted in a decrease in CQAs, may play a key role in the *L. fischeri* drought response. In particular, only one or two BLASTP hits for the translated transcript *LfHCT* (TRINITY_DN31661_c0_g1_i1) were obtained in the proteomes of the different Asteraceae species. This could imply that the HCT gene is important for the regulation of CQA and CQA derivatives in this species. In addition, it is interesting that two transcripts were identified as being *L. fischeri*-speicific at the amino acid sequence level, such as TRINITY_DN22101_c0_g1_i2 and TRINITY_DN26907_c0_g1_i1 encoding *LfC4H* and *LfF3H*, respectively. These findings contribute to a better understanding of drought-responsive changes in morphological, physiological, biochemical, and molecular processes, specifically the regulatory mechanisms of certain genes involved in phenylpropanoid biosynthesis in mitigating drought severity in *L. fischeri*.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/,

SRR11585851; https://www.ncbi.nlm.nih.gov/, SRR11585852; https://www.ncbi.nlm.nih.gov/, SRR11585853; https://www.ncbi.nlm.nih.gov/, SRR11585854; https://www.ncbi.nlm.nih.gov/, SRR11585855; https://www.ncbi.nlm.nih.gov/, SRR11585856.

Author contributions

SMK contributed to the conception and design of the study. YJP and DYK performed the experiments and analyzed the data. YJP and DYK drafted the manuscript. S-CH and JC performed the bioinformatics analysis. SYK, TQT, and JM critically reviewed the manuscript and helped with the data interpretation. SMK supervised the research project and the experiments. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by an intramural grant (2Z06670) from the Korea Institute of Science and Technology (KIST), Republic of Korea.

Conflict of interest

Author 'DYK is employed by Euseed Inc.

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.2023.1140509/full#supplementary-material

References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Nat. Prec.* 1–10. doi: 10.1038/npre.2010.4282.1

Archana,, Ram, S., and Singh, V. (2020). Accumulation of anthocyanins: An aadaptive feature in sugarcane leaf under cold stress. *Agriways* 8, 88–91. doi: 10.38112/agw.2020.v08i02.003

Azam Ansari, M., Chung, I.-M., Rajakumar, G., A Alzohairy, M., Almatroudi, A., Gopiesh Khanna, V., et al. (2019). Evaluation of polyphenolic compounds and pharmacological activities in hairy root cultures of ligularia fischeri turcz. f. spiciformis (Nakai). *Molecules* 24, 1586. doi: 10.3390/molecules24081586

Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Staton, S. E., Cottret, L., et al. (2017). The sunflower genome provides insights into oil metabolism, flowering and asterid evolution. *Nature* 546, 148–152. doi: 10.1038/nature22380

Castellarin, S. D., Matthews, M. A., Di Gaspero, G., and Gambetta, G. A. (2007). Water deficits accelerate ripening and induce changes in gene expression regulating flavonoid biosynthesis in grape berries. *Planta* 227, 101–112. doi: 10.1007/s00425-007-0598-8

Cheng, L., Han, M., Yang, L.-M., Li, Y., Sun, Z., and Zhang, T. (2018). Changes in the physiological characteristics and baicalin biosynthesis metabolism of scutellaria baicalensis georgi under drought stress. *Ind. Crops Prod* 122, 473–482. doi: 10.1016/j.indcrop.2018.06.030

Cheruiyot, E. K., Mumera, L. M., Ng'etich, W. K., Hassanali, A., Wachira, F., and Wanyoko, J. K. (2008). Shoot epicatechin and epigallocatechin contents respond to water stress in tea [Camellia sinensis (L.) o. kuntze]. *Biosci. Biotechnol. Biochem.* 72, 1219–1226. doi: 10.1271/bbb.70698

Cirillo, V., D'amelia, V., Esposito, M., Amitrano, C., Carillo, P., Carputo, D., et al. (2021). Anthocyanins are key regulators of drought stress tolerance in tobacco. *Biology* 10, 139. doi: 10.3390/biology10020139

Cox, M. P., Peterson, D. A., and Biggs, P. J. (2010). SolexaQA: At-a-glance quality assessment of illumina second-generation sequencing data. *BMC Bioinform.* 11, 1–6. doi: 10.1186/1471-2105-11-485

Deng, Y., and Lu, S. (2017). Biosynthesis and regulation of phenylpropanoids in plants. *CRC Crit. Rev. Plant Sci.* 36, 257–290. doi: 10.1080/07352689.2017.1402852

Ebrahimian, E., and Bybordi, A. (2012). Influence of ascorbic acid foliar application on chlorophyll, flavonoids, anthocyanin and soluble sugar contents of sunflower under conditions of water deficit stress. *J. Food Agric. Environ.* 10, 1026–1030.

Egilla, J. N., Davies, F. T., and Drew, M. C. (2001). Effect of potassium on drought resistance of hibiscus rosa-sinensis cv. leprechaun: Plant growth, leaf macro-and micronutrient content and root longevity. *Plant Soil* 229, 213–224. doi: 10.1023/A:1004883032383

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., and Basra, S. (2009). Plant drought stress: Effects, mechanisms and management. *Agron. Sustain. Dev.* 29, 185–212. doi: 10.1051/agro:2008021

Fecka, I., and Turek, S. (2008). Determination of polyphenolic compounds in commercial herbal drugs and spices from lamiaceae: thyme, wild thyme and sweet marjoram by chromatographic techniques. *Food Chem.* 108, 1039–1053. doi: 10.1016/j.foodchem.2007.11.035

Gharibi, S., Tabatabaei, B. E. S., Saeidi, G., Talebi, M., and Matkowski, A. (2019). The effect of drought stress on polyphenolic compounds and expression of flavonoid biosynthesis related genes in achillea pachycephala rech. f. *Phytochemistry* 162, 90–98. doi: 10.1016/j.phytochem.2019.03.004

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Guo, R., Shi, L., Jiao, Y., Li, M., Zhong, X., Gu, F., et al. (2018). Metabolic responses to drought stress in the tissues of drought-tolerant and drought-sensitive wheat genotype seedlings. *AoB Plants* 10, ply016. doi: 10.1093/aobpla/ply016

Guvenalp, Z., Ozbek, H., Karadayi, M., Gulluce, M., Kuruuzum-Uz, A., Salih, B., et al. (2015). Two antigenotoxic chalcone glycosides from mentha longifolia subsp. longifolia. *Pharm. Biol.* 53, 888–896. doi: 10.3109/13880209.2014.948633

Hao, W., Manivannan, A., Yuze, C., and Jeong, B. R. (2018). Effect of different cultivation systems on the accumulation of nutrients and phytochemicals in ligularia fischeri. *Hortic. Plant J.* 4, 24–29. doi: 10.1016/j.hpj.2018.01.004

Hou, Z., Yin, J., Lu, Y., Song, J., Wang, S., Wei, S., et al. (2019). Transcriptomic analysis reveals the temporal and spatial changes in physiological process and gene expression in common buckwheat (Fagopyrum esculentum moench) grown under drought stress. *Agron.* 9, 569. doi: 10.3390/agronomy9100569

Ilyas, M., Nisar, M., Khan, N., Hazrat, A., Khan, A. H., Hayat, K., et al. (2021). Drought tolerance strategies in plants: a mechanistic approach. *J. Plant Growth Regul.* 40, 926–944. doi: 10.1007/s00344-020-10174-5

Jia, X., Sun, C., Li, G., Li, G., and Chen, G. (2015). Effects of progressive drought stress on the physiology, antioxidative enzymes and secondary metabolites of radix astragali. *Acta Physiol. Plant* 37, 1–14. doi: 10.1007/s11738-015-2015-4

Jogawat, A., Yadav, B., Lakra, N., Singh, A. K., and Narayan, O. P. (2021). Crosstalk between phytohormones and secondary metabolites in the drought stress tolerance of crop plants: A review. *Physiol. Plant* 172, 1106–1132. doi: 10.1111/ppl.13328

Kim, H.-J., Choi, H.-K., Chung, M.-Y., Park, J.-H., Chung, S., Lee, S.-H., et al. (2019). Ethanol extract of ligularia fischeri inhibits the lipopolysaccharide-induced inflammatory response by exerting anti-histone acetyltransferase activity to negatively regulate p65. *J. Med. Food* 22, 1127–1135. doi: 10.1089/jmf.2019.4412

Kim, S. M., Jeon, J.-S., Kang, S. W., Jung, Y.-J., Ly, L. N., and Um, B.-H. (2012). Content of antioxidative caffeoylquinic acid derivatives in field-grown ligularia fischeri (Ledeb.) turcz and responses to sunlight. *J. Agric. Food Chem.* 60, 5597–5603. doi: 10.1021/jf300976y

Kim, S.-M., Kang, S.-W., and Um, B.-H. (2010). Extraction conditions of radical scavenging caffeoylquinic acids from gomchui (Ligularia fischeri) tea. *J. Korean Soc Food Sci. Nutr.* 39, 399–405. doi: 10.3746/jkfn.2010.39.3.399

Kvam, V. M., Liu, P., and Si, Y. (2012). A comparison of statistical methods for detecting differentially expressed genes from RNA-seq data. *Am. J. Bot.* 99, 248–256. doi: 10.3732/ajb.1100340

Laforest, M., Martin, S. L., Bisaillon, K., Soufiane, B., Meloche, S., and Page, E. (2020). A chromosome-scale draft sequence of the Canada fleabane genome. *Pest Manag Sci.* 76, 2158–2169. doi: 10.1002/ps.5753

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lee, S., Han, S., Kim, H. M., Lee, J. M., Mok, S.-Y., and Lee, S. (2011). Isolation and identification of phytochemical constituents from taraxacum coreanum. *J. Korean Soc Appl. Biol. Chem.* 54, 73–78. doi: 10.3839/jksabc.2011.010

Li, X.-L., Xiang, L., Wang, X.-H., Qin, P., Zhang, M.-S., and Ren, M.-J. (2020). Biotic and abiotic stress-responsive genes are stimulated to resist drought stress in purple wheat. *J. Integr. Agric.* 19, 33–50. doi: 10.1016/S2095-3119(19)62659-6

Liu, B., Yan, J., Li, W., Yin, L., Li, P., Yu, H., et al. (2020). Mikania micrantha genome provides insights into the molecular mechanism of rapid growth. *Nat. Commun.* 11, 1–13. doi: 10.1038/s41467-019-13926-4

Ma, D., Sun, D., Wang, C., Li, Y., and Guo, T. (2014). Expression of flavonoid biosynthesis genes and accumulation of flavonoid in wheat leaves in response to drought stress. *Plant Physiol. Biochem.* 80, 60–66. doi: 10.1016/j.plaphy.2014.03.024

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Min, F., Li-Ping, J., San-Wen, H., Kai-Yun, X., Qing-Chang, L., and Dong-Yu, Q. (2008). Effects of drought on gene expressions of key enzymes in carotenoid and flavonoid biosynthesis in potato. *Acta Hortic. Sin.* 35, 535.

Nakabayashi, R., Mori, T., and Saito, K. (2014). Alternation of flavonoid accumulation under drought stress in arabidopsis thaliana. *Plant Signal Behav.* 9, e29518. doi: 10.4161/psb.29518

Narayan, M., and Venkataraman, L. (2000). Characterisation of anthocyanins derived from carrot (Daucus carota) cell culture. *Food Chem.* 70, 361–363. doi: 10.1016/S0308-8146(00)00104-7

Nogués, S., Allen, D. J., Morison, J. I., and Baker, N. R. (1998). Ultraviolet-b radiation effects on water relations, leaf development, and photosynthesis in droughted pea plants. *Plant Physiol.* 117, 173–181. doi: 10.1104/pp.117.1.173

Ondov, B. D., Bergman, N. H., and Phillippy, A. M. (2011). Interactive metagenomic visualization in a web browser. *BMC Bioinform.* 12, 1–10. doi: 10.1186/1471-2105-12-385

Pappi, P., Nikoloudakis, N., Fanourakis, D., Zambounis, A., Delis, C., and Tsaniklidis, G. (2021). Differential triggering of the phenylpropanoid biosynthetic pathway key genes transcription upon cold stress and viral infection in tomato leaves. *Horticulturae* 7, 448. doi: 10.3390/horticulturae7110448

Park, C. H., Ahn, M. J., Hwang, G. S., An, S. E., and Whang, W. K. (2016). Cosmeceutical bioactivities of isolated compounds from ligularia fischeri turcz leaves. *Appl. Biol. Chem.* 59, 485–494. doi: 10.1007/s13765-016-0189-0

Piao, X.-L., Mi, X.-Y., Tian, Y.-Z., Wu, Q., Piao, H.-S., Zeng, Z., et al. (2009). Rapid identification and characterization of antioxidants from ligularia fischeri. *Arch. Pharm. Res.* 32, 1689–1694. doi: 10.1007/s12272-009-2204-z

Puente-Garza, C. A., Meza-Miranda, C., Ochoa-Martínez, D., and García-Lara, S. (2017). Effect of *in vitro* drought stress on phenolic acids, flavonols, saponins, and antioxidant activity in agave salmiana. *Plant Physiol. Biochem.* 115, 400–407. doi: 10.1016/j.plaphy.2017.04.012

Qiu, Z., Yuan, M., He, Y., Li, Y., and Zhang, L. (2017). Physiological and transcriptome analysis of he-Ne laser pretreated wheat seedlings in response to drought stress. *Sci. Rep.* 7, 1–12. doi: 10.1038/s41598-017-06518-z

Ranjan, A., Nigam, D., Asif, M. H., Singh, R., Ranjan, S., Mantri, S., et al. (2012). Genome wide expression profiling of two accession of g. herbaceum l. @ in response to drought. *BMC Genom.* 13, 1–18. doi: 10.1186/1471-2164-13-94

Rekha, K., Sivasubramanian, C., and Thiruvengadam, M. (2015). Evaluation of polyphenol composition and biological activities of two samples from summer and winter seasons of ligularia fischeri var. spiciformis nakai. *Acta Biol. Hung.* 66, 179–191. doi: 10.1556/018.66.2015.2.5

Reyes-Chin-Wo, S., Wang, Z., Yang, X., Kozik, A., Arikit, S., Song, C., et al. (2017). Genome assembly with *in vitro* proximity ligation data and whole-genome triplication in lettuce. *Nat. Commun.* 8, 1–11. doi: 10.1038/ncomms14953

Scaglione, D., Reyes-Chin-Wo, S., Acquadro, A., Froenicke, L., Portis, E., Beitel, C., et al. (2016). The genome sequence of the outbreeding globe artichoke constructed *de novo* incorporating a phase-aware low-pass sequencing strategy of F1 progeny. *Sci. Rep.* 6, 1–17. doi: 10.1038/srep19427

Shang, Y. F., Kim, S. M., Song, D. G., Pan, C. H., Lee, W. J., and Um, B. H. (2010). Isolation and identification of antioxidant compounds from ligularia fischeri. *J. Food Sci.* 75, C530–C535. doi: 10.1111/j.1750-3841.2010.01714.x

Sharma, A., Shahzad, B., Rehman, A., Bhardwaj, R., Landi, M., and Zheng, B. (2019). Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. *Molecules* 24, 2452. doi: 10.3390/molecules24132452

Shen, Q., Zhang, L., Liao, Z., Wang, S., Yan, T., Shi, P., et al. (2018). The genome of artemisia annua provides insight into the evolution of asteraceae family and artemisinin biosynthesis. *Mol. Plant* 11, 776–788. doi: 10.1016/j.molp.2018.03.015

Shojaie, B., Mostajeran, A., and Ghanadian, M. (2016). Flavonoid dynamic responses to different drought conditions: amount, type, and localization of flavonols in roots and shoots of arabidopsis thaliana l. *Turk. J. Biol.* 40, 612–622. doi: 10.3906/biy-1505-2

Singh, D., Singh, C. K., Taunk, J., Tomar, R. S. S., Chaturvedi, A. K., Gaikwad, K., et al. (2017). Transcriptome analysis of lentil (Lens culinaris medikus) in response to seedling drought stress. *BMC Genet.* 18, 1–20. doi: 10.1186/s12864-017-3596-7

Sourour, A., Afef, O., Mounir, R., and Mongi, B. Y. (2017). A review: Morphological, physiological, biochemical and molecular plant responses to water deficit stress. *Int. J. Eng. Sci.* 6, 1–4. doi: 10.9790/1813-0601010104

Takahashi, F., Kuromori, T., Sato, H., and Shinozaki, K. (2018). *Regulatory gene networks in drought stress responses and resistance in plants* (Singapore: Springer). Springer, Singapore.

Takahashi, F., Kuromori, T., Urano, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2020). Drought stress responses and resistance in plants: From cellular responses to long-distance intercellular communication. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.556972

Toupchi Khosrowshahi, Z., Salehi-Lisar, S. Y., Ghassemi-Golezani, K., and Motafakkerazad, R. (2019). Effect of polyamines on antioxidative responses of safflower (Carthamus tinctorius) under drought stress. *J. Plant Prod. Sci.* 26, 157–171. doi: 10.22069/JOPP.2019.14942.2337

Treml, J., and Šmejkal, K. (2016). Flavonoids as potent scavengers of hydroxyl radicals. *Compr. Rev. Food Sci. Food Saf.* 15, 720–738. doi: 10.1111/1541-4337.12204

Wilkinson, S., Kudoyarova, G. R., Veselov, D. S., Arkhipova, T. N., and Davies, W. J. (2012). Plant hormone interactions: innovative targets for crop breeding and management. *J. Exp. Bot.* 63, 3499–3509. doi: 10.1093/jxb/ers148

Xie, W. D., Weng, C. W., Li, X., and Row, K. H. (2010). Eremophilane sesquiterpenoids from ligularia fischeri. *Helv. Chim. Acta* 93, 1983–1989. doi: 10.1002/hlca.201000010

Xu, J., Zhu, J., Lin, Y., Zhu, H., Tang, L., Wang, X., et al. (2022). Comparative transcriptome and weighted correlation network analyses reveal candidate genes involved in chlorogenic acid biosynthesis in sweet potato. *Sci. Rep.* 12, 1–11. doi: 10.1038/s41598-022-06794-4

Yadav, B., Jogawat, A., Rahman, M. S., and Narayan, O. P. (2021). Secondary metabolites in the drought stress tolerance of crop plants: A review. *Gene Rep.* 23, 101040. doi: 10.1016/j.genrep.2021.101040

Yadav, R. K., Sangwan, R. S., Sabir, F., Srivastava, A. K., and Sangwan, N. S. (2014). Effect of prolonged water stress on specialized secondary metabolites, peltate glandular trichomes, and pathway gene expression in artemisia annua l. *Plant Physiol. Biochem.* 74, 70–83. doi: 10.1016/j.plaphy.2013.10.023

Yang, J.-L., Wang, R., and Shi, Y.-P. (2011). Phytochemicals and biological activities of ligularia species. *Nat. Prod. Bioprospect.* 1, 1–24. doi: 10.1007/s13659-011-0003-y

Yuan, Y., Liu, Y., Wu, C., Chen, S., Wang, Z., Yang, Z., et al. (2012). Water deficit affected flavonoid accumulation by regulating hormone metabolism in scutellaria baicalensis georgi roots. *PLoS One* 7, e42946. doi: 10.1371/journal.pone.0042946

Zahedi, S. M., Karimi, M., and Venditti, A. (2021). Plants adapted to arid areas: Specialized metabolites. *Nat. Prod. Res.* 35, 3314–3331. doi: 10.1080/14786419.2019.1689500

Zahir, A., Abbasi, B. H., Adil, M., Anjum, S., and Zia, M. (2014a). Synergistic effects of drought stress and photoperiods on phenology and secondary metabolism of silybum marianum. *Appl. Biochem. Biotechnol.* 174, 693–707. doi: 10.1007/s12010-014-1098-5

Zhao, N., Cui, S., Li, X., Liu, B., Deng, H., Liu, Y., et al. (2021). Transcriptome and co-expression network analyses reveal differential gene expression and pathways in response to severe drought stress in peanut (Arachis hypogaea l.). *Front. Genet.* 12. doi: 10.3389/fgene.2021.672884

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 February 2023
DOI 10.3389/fpls.2023.1133912

OPEN ACCESS

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

REVIEWED BY
Jiabao Ye,
Yangtze University, China
Lei Wang,
Kunming Institute of Botany (CAS), China

*CORRESPONDENCE
Hao Hu
haohu@mail.hzau.edu.cn
Caiyun Wang
wangcy@mail.hzau.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 29 December 2022
ACCEPTED 06 February 2023
PUBLISHED 20 February 2023

CITATION
Xu Z, Zeng T, Li J, Zhou L, Li J, Luo J, Zheng R, Wang Y, Hu H and Wang C (2023) TcbZIP60 positively regulates pyrethrins biosynthesis in *Tanacetum cinerariifolium*. *Front. Plant Sci.* 14:1133912. doi: 10.3389/fpls.2023.1133912

TcbZIP60 positively regulates pyrethrins biosynthesis in *Tanacetum cinerariifolium*

Zhizhuo Xu[1†], Tuo Zeng[1,2†], Jiawen Li[1], Li Zhou[1], Jinjin Li[1], Jing Luo[1], Riru Zheng[1], Yuanyuan Wang[1], Hao Hu[1*] and Caiyun Wang[1*]

[1]Key Laboratory for Biology of Horticultural Plants, Ministry of Education, College of Horticulture & Forestry Sciences, Huazhong Agricultural University, Wuhan, China, [2]School of Life Sciences, Guizhou Normal University, Guiyang, China

Pyrethrins, synthesized in the perennial plant *Tanacetum cinerariifolium*, are a class of terpene mixtures with high insecticidal activity and low human toxicity, which are widely used in plant-derived pesticides. Numerous studies have identified multiple pyrethrins biosynthesis enzymes, which can be enhanced by exogenous hormones such as methyl jasmonate (MeJA). However, the mechanism by which hormone signaling regulates pyrethrins biosynthesis and the potential involvement of certain transcription factors (TFs) remain unclear. In this study, we found that the expression level of a TF in *T. cinerariifolium* was significantly increased after treatment with plant hormones (MeJA, abscisic acid). Subsequent analysis identified this TF as a member of the basic region/leucine zipper (bZIP) family and was thus named *TcbZIP60*. TcbZIP60 was localized in the nucleus, suggesting that it is involved in the transcription process. The expression profiles of *TcbZIP60* were similar to those of pyrethrins synthesis genes in different flower organs and at different flowering stages. Furthermore, TcbZIP60 could directly bind to the E-box/G-box motifs in the promoters of the pyrethrins synthesis genes *TcCHS* and *TcAOC* to activate their expression. Transient overexpression of *TcbZIP60* increased the expression levels of pyrethrins biosynthesis genes, leading to the significant accumulation of pyrethrins. Silencing of *TcbZIP60* significantly downregulated pyrethrins accumulation and the expression of related genes. Overall, our results reveal a novel TF, TcbZIP60, that regulates both the terpenoid and jasmonic acid pathways of pyrethrins biosynthesis in *T. cinerariifolium*.

KEYWORDS

***Tanacetum cinerariifolium*, pyrethrins biosynthesis, exogenous hormone, transcription factor, TcbZIP60**

1 Introduction

Tanacetum cinerariifolium is an economically important horticultural ornamental plant in the Asteraceae family because it produces a class of insecticidal compounds called pyrethrins (Casida, 1973). Pyrethrins are widely used as plant-derived pesticides owing to their broad-spectrum and highly effective insecticidal activity, easy decomposition into harmless substances under light exposure, and low toxicity to mammals (Casida and Quistad, 1995). *T. cinerariifolium* has been grown as an environmentally friendly commercial crop in more than a dozen countries in Africa, Asia, Europe, and South America (Lybrand et al., 2020).

Natural pyrethrins are composed of six types of monoterpene esters with similar structures, including pyrethrin I and II, jasmolin I and II, and cinerin I and II (Ramirez et al., 2013). Pyrethrins are obtained by esterification between an acid moiety (pyrethric acid or chrysanthemic acid) and an alcohol moiety (pyrethrolone, cinerolone, or jasomolone) (Freemont et al., 2016; Khan et al., 2017). Monoterpene acids are derived from the methylerythritol-4-phosphate pathway. The first step of the reaction is the formation of chrysanthemyl diphosphate catalyzed by chrysanthemyl diphosphate synthase (CDS) from dimethylallyl diphosphate, which in turn is catalyzed by CDS to trans-chrysanthemol. CDS is a bifocal enzyme that is also known as chrysanthemol synthase (CHS) (Rivera et al., 2001; Yang et al., 2014; Hu et al., 2018; Lybrand et al., 2020). Subsequently, pyrethric acid is formed under the catalysis of alcohol dehydrogenase and aldehyde dehydrogenase (ALDH). Ketols are derived from the oxylipin alcohol pathway. In this way, linolenic acid is used as a substrate to form ketol through lipoxygenase, allene oxide synthase, allene oxide cyclase (AOC), oxo-phytodienoic acid reductase, jasmone hydroxylase, and pyrethrolone synthase (Song et al., 1993; Creelman and Mullet, 1997; Tijet and Brash, 2002; Ramirez et al., 2013; Li et al., 2018). Finally, pyrethrins are synthesized by the two precursor compounds under the catalysis of GDSL lipase-like protein (GLIP) (Kikuta et al., 2012). Although pyrethrins synthesis pathways have been largely clarified, there are relatively few studies on the mechanisms of pyrethrins synthesis at the transcriptional level.

Although pyrethrins are insecticidal compounds, on average, the pyrethrins content of the leaves of plants is only approximately 0.1% (dry weight), which is much lower than that of the flowers (1–2% dry weight) (Varga et al., 2021). A variety of methods have been evaluated to increase the yield of pyrethrins, including optimizing the cultivation mode or modifying the breeding strategy (Li et al., 2014). However, these strategies have not led to a major increase in the pyrethrins content of the leaves. Pyrethrins biosynthesis of the alcohol moiety derives from the jasmonic acid (JA) pathway (Hu et al., 2018). In particular, methyl jasmonate (MeJA) stimulates the overaccumulation of many secondary metabolites in plants. In a previous study, MeJA treatment was shown to induce the accumulation of pyrethrins for a short period, possibly *via* regulating multiple pyrethrins biosynthesis-related genes (Li et al., 2018). However, persistently treating MeJA to plants is not a robust strategy to maintain pyrethrins production.

In the process of plant evolution, many new transcription factor (TF) families have emerged during adaptation to changing environments or in response to exogenous hormones, further affecting the synthesis of secondary metabolites in plants by regulating the expression of target genes (Li et al., 2020; Chen et al., 2022). Therefore, the application of various hormones to plants is a feasible strategy to discover the critical TFs involved in pyrethrins biosynthesis. Basic leucine zipper (bZIP) is one of the most diverse TF families of plants, playing an important role in plant stress signal transduction, pathogen defense, and flower development (Wigge et al., 2005; Tang and Page, 2013; Sagor et al., 2015; Chen et al., 2021). All bZIP TFs possess a conserved bZIP domain, which is usually composed of 60–80 amino acids, including a basic DNA-binding region and a leucine (Leu) zipper domain that can recognize and combine cis-acting elements such as E-box (5′-CANNTG-3′), G-box (5′-CACGTG-3′), and ACGT-box (5′-ACGT-3′) in the gene promoters (Jakoby et al., 2002; Wolfgang et al., 2018).

Many studies have confirmed that bZIP TFs are involved in the response to signaling pathways and abiotic/biotic stress, including light signaling, abscisic acid (ABA) signaling, drought, and pathogen infections (Uno et al., 2000; Sornaraj et al., 2016; Wolfgang et al., 2018). Moreover, bZIP TFs also regulate the biosynthesis of numerous secondary metabolites in plants. For example, in *Artemisia annua*, the bZIP TF AaHY5 directly affects artemisinin biosynthesis through interaction with *AaCOP1* (Hao et al., 2019). In addition, AabZIP1 was reported to activate *ADS* and *CYP71AV1* gene expression under ABA treatment to promote artemisinin biosynthesis (Zhang et al., 2015). Along with a role in the biosynthesis of terpenoids, bZIP TFs also positively or negatively regulate flavonoids and alkaloids biosynthesis by activating the target genes in each biosynthesis pathway (Sibéril et al., 2001; Zhang et al., 2011; Pal et al., 2015; Mao et al., 2021). However, the downstream target genes of TFs are mainly involved in consecutive steps of metabolites synthesis or the same precursor-derived pathway. bZIP TFs regulating genes independently in different biosynthesis pathways are rarely reported.

In this study, we found that MeJA treatment activates reporter gene expression driven by pyrethrins biosynthesis-related gene promoters to increase the pyrethrins content of *T. cinerariifolium*. Based on these findings and the MeJA-treated transcriptome, we successfully identified a novel bZIP TF, named TcbZIP60. Further experiments demonstrated that TcbZIP60 responds to not only MeJA but also to a variety of other plant hormone signals, and can directly bind to the promoter region of the pyrethrins biosynthesis genes *TcCHS* and *TcAOC*. The transient overexpression and virus-induced gene silencing (VIGS) of *TcbZIP60* in *T. cinerariifolium* leaves further confirmed the positive role of TcbZIP60 in regulating the biosynthesis of both terpenoid- and JA-derived pyrethrins.

2 Materials and methods

2.1 Plant materials and growth conditions

Tanacetum cinerariifolium 'W99' plants grown in the flower base of Huazhong Agricultural University, Wuhan, China were used in this study. This cultivar has the advantages of easily rooting cuttings, rapid growth, and high pyrethrins content and, therefore, was suitable

for the present experiments. W99 seedlings were subcultured in half-strength Murashige and Skoog medium for 1 month (25°C, 16 h light/8 h dark) prior to transient treatment with MeJA or ABA. The rooted cuttings were sprayed with 5 ml of 2 mM MeJA/ABA solution as a single foliar application. Leaves were sampled in triplicate at 0 (control), 2, 4, 6, 8, 12, and 24 h after treatment, immediately placed in liquid nitrogen, and stored at −80°C until further analysis.

Flower heads and leaves were harvested at seven flowering stages: S1, well-developed closed buds; S2, ray floret limb in vertical position; S3, ray floret limb in horizontal position and disc florets of the outermost whorl open; S4, three whorls of disc florets open; S5, all disc florets open; S6, termed the early overblown condition, the disc floret color is faded but the ray florets remain intact; S7, termed the late overblown condition, the disc florets retain little color and the ray florets are desiccated. Three biological replicates were collected for each sample, which were immediately frozen in liquid nitrogen and stored at −80°C for further analysis.

2.2 Exogenous MeJA effect on *TcCHS* and *TcGLIP* promoter activity

Our previously established transgenic chrysanthemum (*Chrysanthemum* × *morifolium*) harboring *TcCHS*-promoter-driven GFP and tobacco (*Nicotiana tabacum*) harboring *TcGLIP*-promoter-driven GUS were treated with MeJA (Sultana et al., 2015). The plants were sprayed with 5 ml of 300 µM MeJA dissolved in 0.8% ethanol as a single foliar application. Leaves were collected in triplicate at 0 (control) and 12 h after treatment, and were immediately frozen in liquid nitrogen. The GFP expression level was determined by quantitative real-time PCR (qRT-PCR) analysis. The GUS activity was measured as previously described (Luo et al., 2013).

2.3 Gene cloning and bioinformatic analysis

Total RNA was extracted from the collected samples using the standard phenol–chloroform extraction method. The first-strand cDNA was synthesized using the EasyScript® One-step gDNA Removal and cDNA Synthesis SuperMix Kit (TransGen Biotech, Beijing, China), in accordance with the manufacturer's instructions, using the total RNA extracts as the template. The *TcbZIP60* gene sequence of *T. cinerariifolium* was identified from the inflorescence transcriptome database determined by our laboratory. The Primer Premier software was used to design specific primers for amplifying the open reading frame (ORF) fragment. The primers are listed in Supplementary Table S1. The *cis*-acting elements of associated genes (*TcCHS*, *TcAOC*, *TcALDH*, and *TcGLIP*) were determined using the PlantCARE database. The *cis*-regulatory E-box and G-box elements were determined using PlantPAN 3.0 (Chow et al., 2019).

2.4 Real-time quantitative PCR

Total RNA was extracted using the Ultrapure RNA Kit (CWBIO, Beijing, China), and the EasyScript One-step gDNA Removal and cDNA Synthesis SuperMix Kit (TransGen Biotech) was used to reverse-transcribe the RNA into cDNA. A qRT-PCR analysis was performed using a LightCycler® 96 Real-Time PCR System (Roche, Basel, Switzerland) in accordance with the manufacturer's instructions, with SYBR Premix Ex Taq II (Takara, Kusatsu, Japan) and sequence-specific primers (listed in Supplementary Table S1). Glyceraldehyde-3-phosphate dehydrogenase (GADPH) was used as the internal reference gene (Li et al., 2019). The relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method with three biological and three technical replicates (Livak and Schmittgen, 2001). Student's two-tailed *t*-test was used to determine statistical significance. Differences at $P < 0.05$ were considered to be significant and those at $P < 0.01$ were considered to be highly significant.

2.5 Phylogenetic analysis of TcbZIP60

The ORF of *TcbZIP60* was cloned from *T. cinerariifolium* cDNA using the sequence-specific primers TcbZIP60-F and TcbZIP60-R (Supplementary Table S1). *Arabidopsis* sequences homologous to the *TcbZIP60* sequence were identified by a Basic Local Alignment Search Tool (BLAST) search against The Arabidopsis Information Resource (TAIR) database, and homologous sequences from other plant species were identified by a BLAST search against the National Center for Biotechnology Information (NCBI) database. A phylogenetic tree was constructed with the maximum likelihood method using MEGA X software with 1000 bootstrap replications (Kumar et al., 2018). The GenBank accession numbers of the genes used in the analysis are listed in Supplementary Tables S2 and S3.

2.6 Subcellular localization of TcbZIP60

The *TcbZIP60* coding sequence was amplified by PCR using sequence-specific primers (Supplementary Table S1). The gene was cloned using the ClonExpress® II One Step Cloning Kit (Vazyme Biotech, Nanjing, China). The full-length sequence of *TcbZIP60* (without the stop codon) was spliced into the pSuper-1300 GFP vector. The recombinant product was transferred into *Escherichia coli* strain DH5α, then sequenced, and the correct plasmid was extracted and constructed. The resultant plasmid was introduced into *Agrobacterium tumefaciens* strain GV3101, and then co-infiltrated into *Nicotiana benthamiana* leaves together with the RFP-NLS plasmid (a nuclear marker). The empty vector was used as a control. After 72 h of weak light exposure, fluorescence signals were observed with a confocal laser scanning microscope (TCS-SP8, Leica, Wetzlar, Germany).

2.7 Yeast one-hybrid assay

The full-length cDNA of *TcbZIP60* was cloned into the pGADT7 vector used for homologous recombination, and the promoter sequences of *TcCHS* and *TcAOC* were cloned separately into the pHis2.1 vector. The resultant pGADT7:*TcbZIP60* plasmid was co-transformed into yeast strain Y187, together with pHis2.1:*TcCHS* or pHis2.1:*TcAOC*, using the Super Yeast Transformation Kit (Coolaber, Beijing, China). The transformed yeast cells were selected on DDO

(SD/−Leu/−Trp) medium and interreacted with TDO (SD/−Leu/−Trp/−His) medium at 30°C for 3 days.

2.8 Dual-luciferase reporter assay

To generate reporter constructs, and in accordance with the ORF sequence of the cloned *TcbZIP60* gene and the map of the plant expression vector, the downstream sequences of the gene promoter region (for *TcCHS* and *TcAOC*) were cloned into the linearized pGreenII0800-LUC vector. To generate effector constructs, the coding sequence of *TcbZIP60* was cloned into the linearized pGreenII62-SK vector under the control of the Cauliflower mosaic virus (CaMV) *35S* promoter. The resultant vectors were transiently co-expressed in *N. benthamiana* leaves. Luminescence was detected using the LB 985 Nightshade system (Berthold, Bad Wildbad, Germany). Introduction of the pGreen-62-SK empty vector and the pGreen-62-SK/pGreenII-0800-LUC : *TcCHS*/*TcAOC* constructs served as negative controls. Three biological replicates per treatment were measured.

The CDS of *TcbZIP60* was ligated into pGreenII62-SK vector to generate an effector plasmid, while *TcCHS*/*TcAOC* were fused into the vector pGreenII0800-LUC to produce the reporter plasmids. The effector, each of the two reporter constructs were co-transformed into *A. tumefaciens* GV3101, respectively. Transient expression assay in *N. benthamiana* leaves was as described previously (Geng and Liu, 2018). The activities of firefly luciferase (LUC) and Renilla luciferase (REN) were measured using the Dual-Luciferase® Reporter Assay System (Promega, WI, USA). The promoter activity was expressed as the ratio of LUC to REN (Supplementary Figure S1).

2.9 Electrophoretic mobility shift assay

The ORF without the *TcbZIP60* terminator codon was used to generate the pET6×HN-C vector protein, which was fused into the N-terminal frame of 6×His, and the vector pET6×HN-C-TcbZIP60 was then transformed into *E. coli* strain Rosetta (DE3). For the induced recombinant protein, 0.5 mM IPTG was used, and the cultures were incubated at 18°C for 16 h. Then, Ni^{2+}–nitrilotriacetic acid was used to purify the recombinant proteins. For the electrophoretic mobility shift assay (EMSA), the promoter fragments of *TcCHS* and *TcAOC* containing E-box or G-box *cis*-regulatory elements labeled with fluorescein amidite (FAM) as probes, the same but unlabeled DNA fragments, and *cis*-element mutant DNA fragments were used as competitors in the assay. After performing the EMSA assays, FAM-labeled DNA was detected from the chemiluminescent signal.

2.10 Transient overexpression of TcbZIP60 in *T. cinerariifolium* leaves

The full-length *TcbZIP60* coding sequence was cloned into the *Hin*dIII-linearized pGreenII62-SK vector downstream of the CaMV *35S* promoter using gene-specific primers (Supplementary Table S1). The experimental method followed a previously described procedure (Jung et al., 2015). The pGreenII62-SK : *TcbZIP60* vector was co-transformed together with the helper plasmid pSoup19 into *A. tumefaciens* strain GV3101. The pSoup19-transformed *Agrobacterium* cells were inoculated into YEB liquid medium containing 100 mg/l kanamycin and cultured on a rotating shaker at 28°C for 12 h. The supernatant was removed and resuspended in MES (containing 100 μM acetosyringone) to attain the final optical density (OD_{600} = 0.6). Leaves of *T. cinerariifolium* were placed in a 500 ml beaker containing a suspension of *A. tumefaciens*, and the beaker was then placed in a vacuum chamber at 0.23 ATM for 5 min. The soaked leaves were then dried with filter paper and stored in a petri dish lined with moist filter paper in the base. After 4 days of culture, leaves were sampled for qRT-PCR and high-performance liquid chromatography (HPLC) analyses. To determine the effect of transient expression, transient overexpression of GUS in *T. cinerariifolium* leaves was verified by staining with X-Gluc reagent (Supplementary Figure S2).

2.11 VIGS assay

The Tobacco rattle virus (TRV)-based vectors pTRV1 and pTRV2 were used for the VIGS assay (Supplementary Table S1). The pTRV2:*TcbZIP60* and pTRV1 plasmids were transformed into chemically active *A. tumefaciens* strain GV3101 cells using a liquid nitrogen freeze–thaw method. The positive transformant cells were inoculated into YEB liquid medium containing 100 mg/l kanamycin, incubated overnight on a shaker at 28°C, then reactivated in an infiltrating buffer (pH = 5.6) containing 10 mM $MgCl_2$, 10 mM MES, and 100 μM acetoeugenone, and adjusted to OD_{600} = 0.6. After standing for 3 h, *Agrobacterium* cells containing the pTRV2 plasmid and pTRV1 plasmid were mixed (1:1, v/v), and the bacterial solution was injected through the adaxial epidermis of *T. cinerariifolium* leaves using a needle-free syringe. The leaves were then incubated in a culture room in the dark for 3 days and under light for 11 days. The pTRV1 and pTRV2 vectors were used as the control group. The method followed a previously described procedure (Senthil-Kumar and Mysore, 2014). The VIGS assay was conducted with three biological replicates. The empty vectors were used as controls. After 14 days, the samples were subjected to qRT-PCR and HPLC analyses.

2.12 Pyrethrins quantitation by HPLC

Transiently transformed leaf samples were dried for 48 h in an oven at 50°C to constant dry weight. The dried leaves were ground to a fine powder, then 100 mg powder was placed in a screw-capped glass tube, and dissolved in 600 μl *n*-hexane. After vortex-oscillation for 30 s, the sample was extracted in an ultrasonic water bath for a further 10 min, followed by vortex-oscillation for 30 s. The extracted samples were filtered with a 0.22 μm filter and analyzed by HPLC. The pyrethrins content was determined using a Waters HPLC system equipped with a photodiode array detector as described previously (Hu et al., 2018). Three biological replicates were analyzed for each sample, and commercial pyrethrum extract (Sigma-Aldrich, St. Louis,

MO, USA) was used as the standard. Pyrethrins standard solution mother liquor was obtained by absorbing 3.0 µl pyrethrins standard and dissolving it in a small amount of *n*-hexane, and finally the volume was standardized to 1.00 ml. The mother liquor of 200, 100, 50, 25, 12.5, and 6.25 µl was standardized to 250 µl with methanol to obtain the pyrethrins standard sample solutions with concentrations of 1.15, 0.575, 0.288, 0.144, 0.072, and 0.036 mg/ml, respectively. After absorbing 20 µl of each standard sample solution, the samples were analyzed by HPLC and a standard curve for the pyrethrins standard samples was generated with Microsoft Excel.

2.13 Statistical analysis

All experiments were repeated using at least three biological replicates. The data were analyzed with one-way ANOVA and Student's *t*-test using SPSS 18 software. Significant differences were determined with Student's *t*-test, with $p < 0.05$ considered to be statistically significant.

3 Results

3.1 Exogenous MeJA positively regulates pyrethrins biosynthesis

Pyrethrins are derived from two independent pathways: the terpenoid biosynthesis pathway and the JA biosynthesis pathway (Figure 1A). Most pyrethrins biosynthesis enzymes have been identified to date, including those encoded by the key regulatory genes *CHS*, *AOC*, *GLIP*, and *ALDH* (Kikuta et al., 2012; Xu et al., 2018a). Therefore, we first analyzed the promoter elements of these key genes involved in pyrethrins synthesis using the PlantCARE and PlantPAN3.0 databases (Figure 1B), revealing several hormone-responsive elements, including those responding to MeJA, ABA, and salicylic acid (SA). This analysis indicated that MeJA treatment might play a role in pyrethrins synthesis.

Using our previously established transgenic platform for CHSpro-green fluorescent protein (GFP) and GLIPpro-b-glucuronidase (GUS) reporter genes (Sultana et al., 2015), we

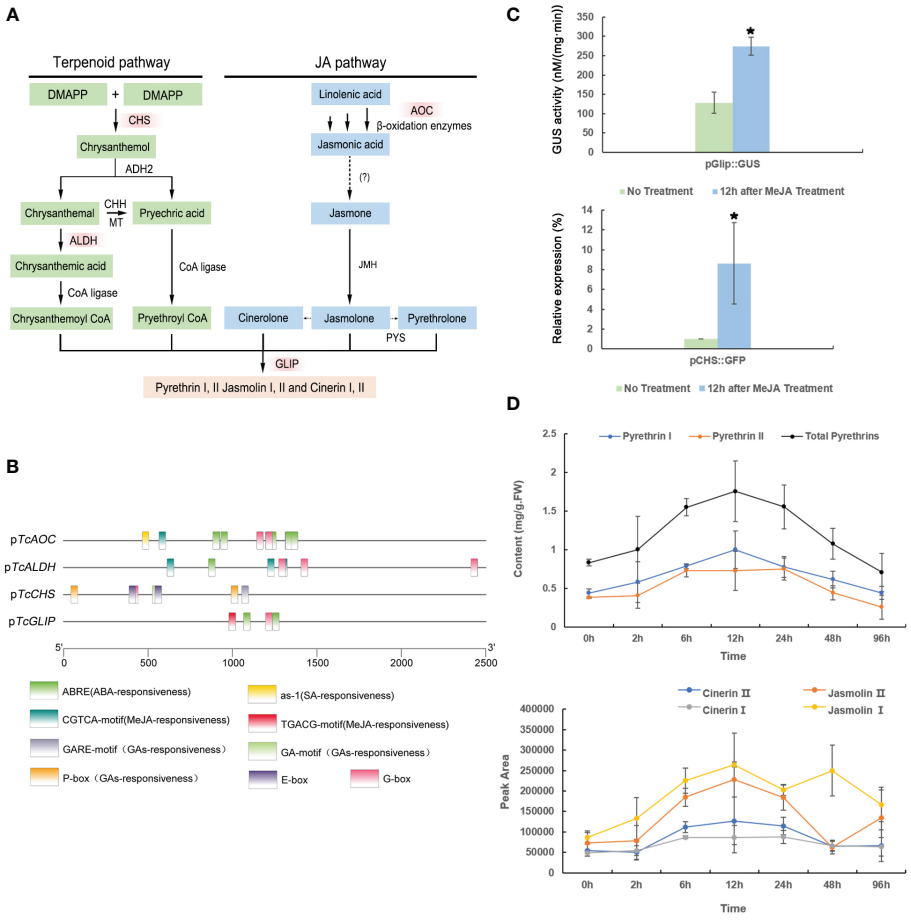


FIGURE 1
The pathway of pyrethrin biosynthesis and bioanalysis of promoters. **(A)** The biosynthesis of pyrethrins. The red boxes show the key genes involved in the pathway. **(B)** Analysis of the cis-regulatory elements in promoters of the key pyrethrin biosynthesis genes. **(C)** Effect of methyl jasmonate (MeJA) treatment on *CHS* and *GLIP* promoter expression. MeJA induced the pCHS::GFP gene in transgenic chrysanthemum. *GFP* gene expression was detected by real-time PCR at 12 h after spraying with 300 µM MeJA. MeJA induced the pGlip::GUS gene in transgenic tobacco. GUS activity was determined at 12 h after spraying 300 µM MeJA. Data are presented as mean ± SE. *P < 0.05 (ANOVA followed by Duncan's multiple range test). **(D)** Effect of MeJA treatment on pyrethrins production in *Tanacetum cinerariifolium* plants. Leaves were collected at 2, 6, 12, 24, 48, and 96 h after spraying 300 µM MeJA. Samples at 0 h were considered the control. The pyrethrins content was quantified by high-performance liquid chromatography using pyrethrin external standards. Data indicate mean ± SE (n = 3).

evaluated the MeJA-induced upregulation on the *CHS* and *GLIP* promoters individually. After treatment with 300 μM MeJA, the specific *GFP* expression level was measured by real-time polymerase chain reaction (PCR) and GUS activity was measured. Consequently, the *GFP* gene driven by the *CHS* promoter and the *GUS* gene driven by the *GLIP* promoter were both significantly activated at 12 h after MeJA treatment (Figure 1C), confirming that the *CHS* and *GLIP* promoters are MeJA-inducible, which will ultimately lead to enhanced pyrethrins production.

Indeed, in our previous study, we found that the pyrethrins content increased for a short time under high-concentration (2 mM) treatment of MeJA (Zeng et al., 2022). However, treatment of such a high concentration of MeJA leads to a dramatic decline in the pyrethrins content after 6 h due to disturbed homeostasis or metabolites feedback. Thus, in the present study, we treated *T. cinerariifolium* plants with a reduced concentration of MeJA of 300 μM. As expected, *T. cinerariifolium* plants showed a higher yield of pyrethrins after MeJA treatment, including all six constituents. The pyrethrins content increased by 2-fold to 0.99 mg/g per fresh weight at 12 h after MeJA treatment, and this accumulation level was maintained for 4 days (Figure 1D). These results strongly suggested that pyrethrins biosynthesis genes are MeJA-inducible and that the optimal working concentration of MeJA treatment could maintain a high pyrethrins yield for several days.

3.2 Cloning and characterization of bZIP60 in *T. cinerariifolium*

Considering that many E-box and G-box elements that bind to members of the bZIP TF family were identified in the promoter regions of pyrethrins biosynthesis genes, we searched for genes encoding putative bZIP TFs in the transcriptome data of *T. cinerariifolium*. Since we found that MeJA treatment can activate the *CHS* and *GLIP* promoter-driven reporter genes and improve pyrethrins production, we selected Unigene27160, which showed the highest expression level induced by MeJA, as the candidate bZIP TF (Figure 2A).

To identify the Unigene27160 gene in the *T. cinerariifolium* genome, the sequence was input as a query object in The Arabidopsis Information Resource (TAIR). In the phylogenetic comparison between Unigene27160 and members of the *Arabidopsis* bZIP TF family, Unigene27160 clustered with Arabidopsis *AtbZIP60* and *AtHY5* in the same evolutionary branch with high homology, suggesting that these genes exhibit similar functions (Figure 2B, blue box). Unigene27160 showed the closest relationship with *AtbZIP60* (Figure 2C). Moreover, 10 bZIP60 proteins from different plant species were selected for further comparative sequence analysis. These proteins clustered into five branches, represented by boxes with different colors in Figure 2D. Unigene27160 clustered with NabZIP60, NbbZIP60, and AtbZIP60, and contained a conserved bZIP domain similar to that of other bZIP proteins (Figure 2E). Thus, Unigene27160 was named *TcbZIP60*, and its open reading frame region was cloned for further analysis.

3.3 Expression profiles and subcellular localization of TcbZIP60

Since *TcbZIP60* expression was most strongly induced by MeJA, we confirmed this result by real-time PCR and further investigated whether other hormones could also induce the expression of *TcbZIP60*. As expected, treatment with both MeJA and ABA strongly induced *TcbZIP60* expression, and high-level expression was maintained for 24 h (Figures 3A, B). To confirm that TcbZIP60 is involved in pyrethrins biosynthesis, the relative expression levels of *TcbZIP60* and other pyrethrins biosynthesis genes were analyzed at different flowering stages. The expression profiles were similar, with high expression detected in the first two stages (Figure 3C). Based on these results, we speculated that the TF TcbZIP60 regulates pyrethrins synthesis through interaction with the pyrethrins synthase gene.

To determine the subcellular localization of TcbZIP60, the coding sequence of *TcbZIP60* was fused into the GFP framework under control of the *CaMV35S* promoter. When TcbZIP60-GFP was instantaneously expressed in *Nicotiana benthamiana* leaves, strong and specific fluorescence was observed in the nucleus (Figure 3D). In addition, the co-localization of the GFP signal and red fluorescent protein (as a nuclear localization signal) resulted in a fused yellow fluorescence signal (Figure 3E). These results suggested that TcbZIP60 is a nuclear-localized protein, further supporting its potential role as a TF.

3.4 TcbZIP60 directly binds to the *TcCHS* and *TcAOC* promoters

The binding sites of bZIP TFs were predicted in the promoter sequence of the pyrethrins synthesis genes (Figure 1B). In *T. cinerariifolium*, similar expression profiles were observed between *TcbZIP60* and pyrethrins synthesis genes during different flowering stages (Figure 3C). Therefore, we further speculated that TcbZIP60 can regulate the expression of these genes by directly binding to the corresponding promoters. To test this hypothesis, we performed a yeast monohybrid (Y1H) experiment. The promoters of pyrethrins synthesis genes were cloned and inserted into the pHis2.1 vector to generate reporter genes. TcbZIP60 was fused to the GAL4 activation domain to generate the effector construct pGADT7-*TcbZIP60* (Figure 4A).

In the interaction experiments with several pyrethrins synthesis gene promoters, yeast clones were observed on solid synthesized Leu, Trp, and His (SD/-Leu-Trp-His) media only when pGADT7-*TcbZIP60* was transferred into yeast cells expressing pHis2.1-*TcCHS* and pHis2.1-*TcAOC*, but not in those expressing pHis2.1. This suggested that TcbZIP60 binds directly to the promoters of *TcCHS* and *TcAOC* (Figure 4B). In addition, we hypothesized that TcbZIP60 directly binds to the motifs on the promoters of the two genes, which was further verified by electrophoretic mobility shift assays (EMSAs).

We predicted that TcbZIP60 binds to the potential E-box and G-box sites on the two promoters, and designed the corresponding

10.3389/fpls.2023.1133912

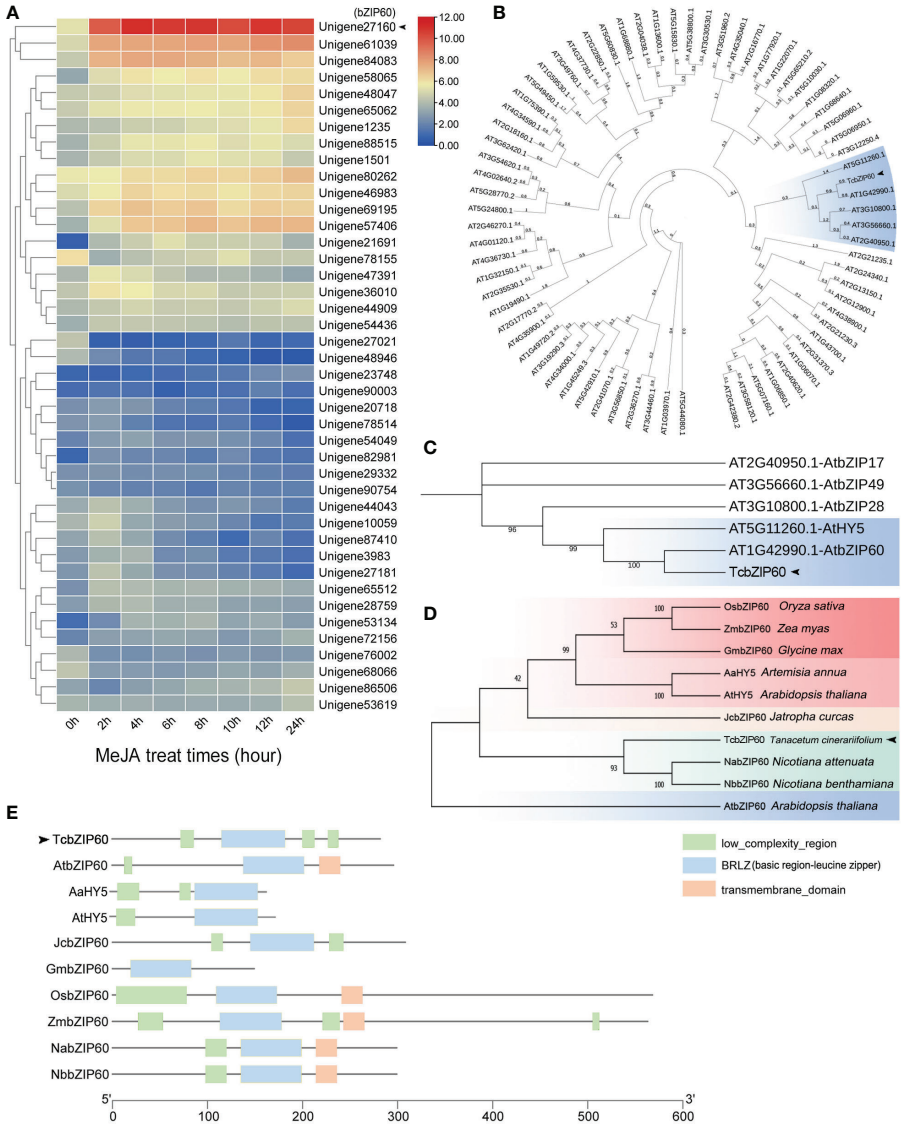


FIGURE 2
Expression patterns, phylogenetic analysis, and domain of the bZIP transcription factor family. **(A)** Heat maps of differentially expressed genes in the transcriptome with methyl jasmonate (MeJA) treatment to the *T. cinerariifolium* leaves at different times. The Pearson distance measure and Ward clustering algorithm were used. **(B)** Phylogenetic analysis of TcbZIP60 and *Arabidopsis thaliana* bZIP family protein sequences obtained from the PfamA domain database. **(C)** Partial enlargement of the phylogenetic tree in **(B)**. **(D)** Phylogenetic analysis of TcbZIP60 and bZIP family protein sequences of 12 other plants obtained from the NCBI database. **(E)** BRLZ domain of TcbZIP60 and its homologs in other species.

mutation sites (Figure 4C). The migration bands were detected in the presence of the purified and concentrated TcbZIP60 protein and the labeled probe of the E-box containing the *TcCHS* promoter or the G-box containing the *TcAOC* promoter (Figure 4D). When 50 times the concentration of the cold probe (unlabeled probe) was added, the intensity of the migration band decreased. When the mutant probe (unlabeled probe) was added, the intensity of the migration band recovered.

To further verify the interaction between TcbZIP60 and the *TcCHS* or *TcAOC* promoter in plants, we performed a transient dual-luciferase (dual-LUC) assay with a reporter structure (*TcCHS*/*TcAOC* Pro : LUC) and effector structure (Figure 4E). Strong LUC activity was observed in tobacco leaves co-transformed by *TcCHS*/*TcAOC* Pro : LUC and 35S:*TcbZIP60*. However, almost no fluorescence signal was detected in tobacco leaves co-transformed with *TcCHS*/*TcAOC* Pro : LUC and an empty carrier (Figure 4F). These results suggested that the TF TcbZIP60 directly activates promoters of the *TcCHS* and *TcAOC* genes and regulates pyrethrins biosynthesis in *T. cinerariifolium*.

3.5 Activation of pyrethrins biosynthesis by TcbZIP60

Previous studies demonstrated that bZIP TFs in *A. annua* bind to gene promoters and regulate artemisinin biosynthesis (Hao et al., 2019; Lv et al., 2019). Similarly, we confirmed that TcbZIP60 can directly bind to the promoters of the pyrethrins biosynthesis genes

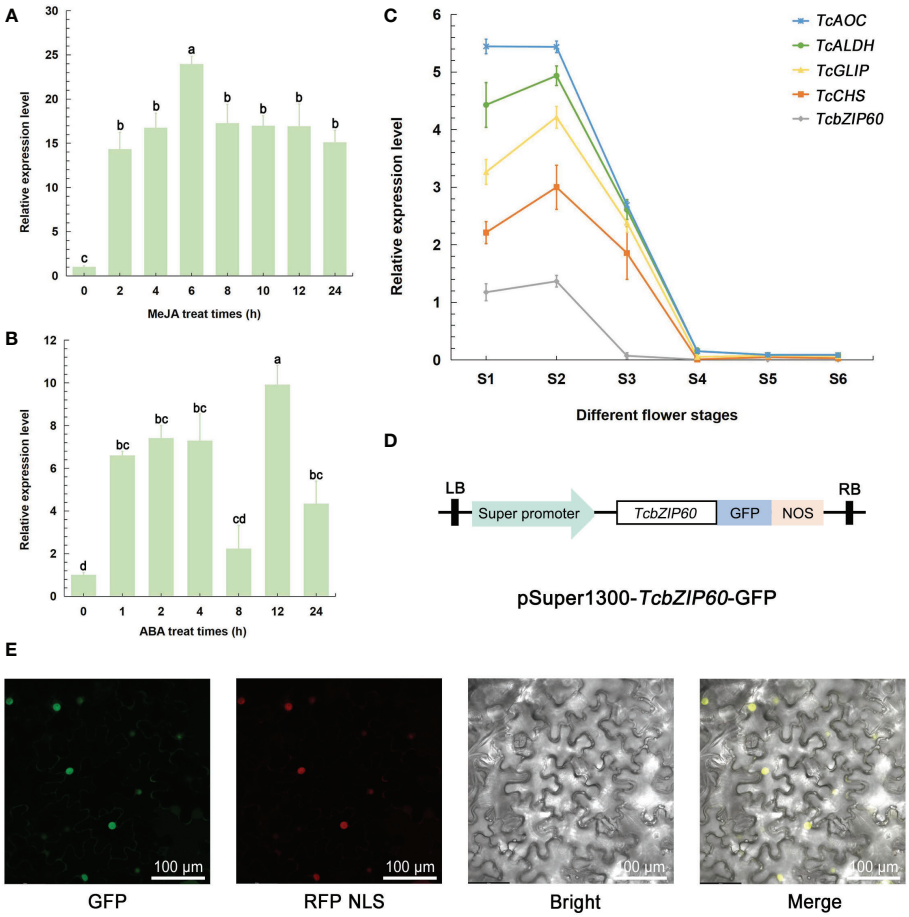


FIGURE 3
Expression profile and subcellular localization of TcbZIP60. **(A)** RT-qPCR analysis of the transcript abundance of *TcbZIP60* induced in the leaves of 1-month-old tissue culture seedlings treated with methyl jasmonate (MeJA); three biological replicates were established. **(B)** RT-qPCR analysis of the transcript abundance of *TcbZIP60* induced in the leaves of 1-month-old tissue culture seedlings treated with abscisic acid (ABA); three biological replicates were established. **(C)** RT-qPCR analysis of *TcbZIP60* transcript abundance from stages S1 to S6. **(D)** Diagram of the pSuper1300-*TcbZIP60*-GFP construct. **(E)** Subcellular localization of TcbZIP60 in *N. benthamiana* leaves. Scale bars = 100 μm.

TcCHS and *TcAOC*. Therefore, to further verify the role of TcbZIP60 in regulating pyrethrins biosynthesis, we constructed *TcbZIP60*-overexpressing plants (*TcbZIP60*-OE) and *TcbZI60*-suppressed plants (pTRV2-*TcbZIP60*) in *T. cinerariifolium*. Transient overexpression and transient RNA interference-mediated silencing of TcbZIP60 were performed in the leaves of *T. cinerariifolium*, and the expression of the pyrethrins biosynthesis genes was detected by reverse transcription-quantitative PCR after transformation. Compared with wild-type (Mock) plants of pGreenII 62-SK, the transcription level of *TcbZIP60* in *TcbZIP60*-OE plants significantly increased by 49.65 times, and the expression levels of the pyrethrins synthesis genes *TcCHS* and *TcAOC* also increased by 1.76 times and 1.88 times, respectively. The other two pyrethrins biosynthesis genes, *TcALDH* and *TcGLIP*, were also upregulated with *TcbZIP60* overexpression (Figure 5A). By contrast, in the *TcbZIP60*-silenced plants, the transcription level of *TcbZIP60* decreased by 0.8-fold (Figure 5B) and the expression levels of four pyrethrins synthesis genes also decreased (Figure 5C).

Subsequently, we measured the contents of pyrethrins in the transient overexpression and RNA interference plants. We used the pyrethrins standard to compare and identify the six main components of pyrethrins in the sample as shown by the arrow and the black line in Figures 5E, F, whereas the red and green lines represent the HPLC results of the experimental group and the control group, respectively. Compared with the pGreenII 62-SK Mock plants in the blank control, the total amount of pyrethrins in the overexpression plants significantly increased by 1.239 times (Figures 5B, E) and decreased by 0.273 times in the *TcbZIP60* interference plants (Figures 5D, F). These results showed that TcZIP60 can positively regulate pyrethrins accumulation by activating pyrethrins biosynthesis genes.

4 Discussion

Pyrethrins represent unique metabolites produced in pyrethrums that can provide plants with an effective endogenous chemical defense against insect and fungal diseases (Zeng et al., 2021). Since the 19th century, pyrethrins products have been gradually used as household and agricultural insecticides, and as insect-borne disease prevention agents (Orenstein, 1913; Lange and Akesson, 1973; Katsuda, 1999). However, in the extraction process for agricultural production, pyrethrins are obtained from the dry flowers of *T. cinerariifolium*

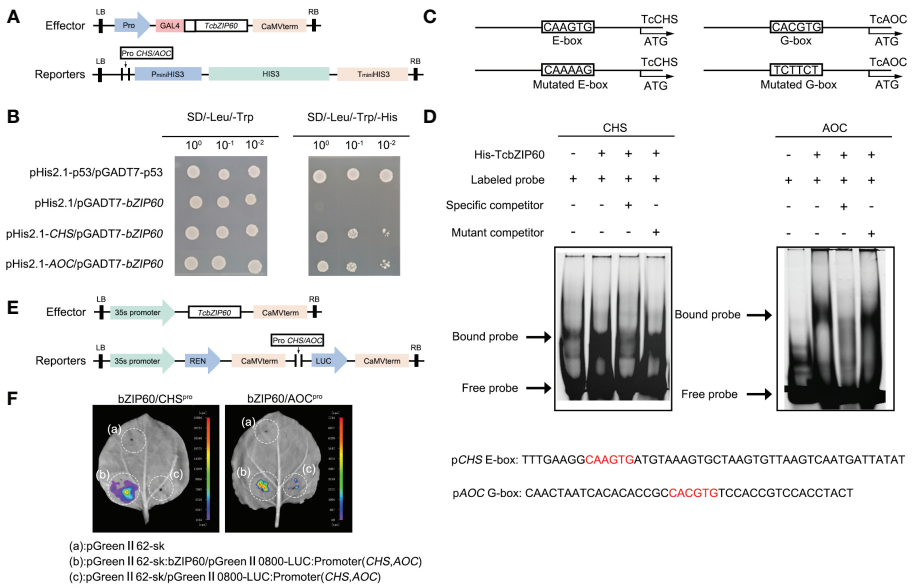


FIGURE 4
TcbZIP60 directly binds to and activates the promoters of the *TcCHS* and *TcAOC* genes. **(A)** Diagram of vectors used in the yeast one hybrid (Y1H) assay. Reporters: the promoter of *TcCHS*/*TcAOC* was introduced into the pHis2.1 vector. Effector: *TcbZIP60* was introduced into the pGADT7 vector. **(B)** Y1H assay showing TcbZIP60 binding to *TcCHS*/*TcAOC* promoter fragments. **(C)** Schematic diagram of the *TcCHS*/*TcAOC* promoter showing the potential TcbZIP60 binding sites. The predicted E-box/G-box and mutated E-box/G-box sites and sequences are indicated with black boxes. **(D)** Electrophoretic mobility shift assay (EMSA) of interactions between TcbZIP60 and the promoters of the *TcCHS* and *TcAOC* genes. The FAM-labeled probe (10 μM) was competed with a 50-fold excess of unlabeled wild-type cold probes and unlabeled mutant-type cold probes. **(E)** Diagram of dual-LUC vectors. Effector: *TcbZIP60* was introduced into the pGreen-62-SK vector, which was driven by the 35S promoter. Reporters: the promoter of *TcCHS*/*TcAOC* was introduced into the pGreen II-0800-LUC vector, driving the translation of LUC. **(F)** Dual-LUC reporter assay showing TcbZIP60 binding to the *TcCHS*/*TcAOC* promoters and promotion of LUC expression.

and the yield of pyrethrins per plant is low, only accounting for 0.10–1.35% of the dry weight of flowers (Varga et al., 2021). Therefore, increasing the amount of pyrethrins produced in the leaves is the key to breaking through the bottleneck of production.

Previously, we reported that MeJA can upregulate the expression of pyrethrins biosynthesis genes as well as induce pyrethrins accumulation (Zeng et al., 2022). However, the high working concentration of MeJA used in the previous study cannot be sustained to constituently increase the pyrethrins content. In general, under stress/MeJA treatment, plant growth and metabolites synthesis will be halted to invest in defense. In the present study, treatment of a lower concentration of MeJA maintained the increase in the pyrethrins content in the leaves for more than 3 days. This was considered to be due to activation of the promoters of pyrethrins biosynthesis genes by MeJA. To further investigate the mechanism by which MeJA regulates pyrethrins biosynthesis, we identified the TF TcbZIP60 by the global expression profile, which was confirmed to be induced by exogenous MeJA, suggesting that TcbZIP60 might be involved in MeJA-regulated pyrethrins biosynthesis.

In addition to MeJA, ABA also plays an important role in regulating the biosynthesis of terpenoids (Jing et al., 2009). TFs involved in mediating ABA stimulation of terpenoids biosynthesis, including AaHY5, AaTGA6, and AabZIP1, respond to JA, SA, and ABA treatment and regulate artemisinin biosynthesis in *A. annua* (Hao et al., 2019; Lv et al., 2019; Shu et al., 2022). Similarly, we found that the transcription level of *TcbZIP60* was regulated by ABA treatment. Cross-talk between ABA and MeJA is a universal phenomenon, and some TFs involved in ABA signaling also participate in MeJA signaling in Arabidopsis and tobacco (Lackman et al., 2011). For example, AaMYC2, which is involved in JA signaling, was suggested to function as the bridge of JA signaling to the ABA signaling pathway (Abe et al., 2003). *AaMYC2* expression could be upregulated by ABA-induced AabZIP1 through direct interaction to subsequently activate *AaALDH1* transcription, thereby governing artemisinin biosynthesis (Shu et al., 2022). In *T. cinerariifolium*, TcMYC2 was also identified as a MeJA-induced TF that positively regulates pyrethrins biosynthesis by upregulating *TcCHS*, *TcAOC*, and *TcGLIP* gene expression (Zeng et al., 2022). Thus, it is strongly speculated that TcbZIP60 might also interact with TcMYC2 to co-regulate pyrethrins biosynthesis. This could explain why we detected upregulation of *TcGLIP* expression even when its promoter was not directly activated by TcbZIP60. Therefore, whether more TFs in *T. cinerariifolium* function as a network to cooperatively regulate pyrethrins biosynthesis is worthy of further study.

Through the development of metabolic engineering, there has been significant progress in producing pyrethrins in other model plants by introducing pyrethrins biosynthesis genes into heterologous systems. However, to date, only the precursors of pyrethrins have been synthesized in tobacco and tomato (Xu et al., 2018b; Xu et al., 2019), because not all of the genes in the full pathway have been cloned. *T. cinerariifolium* remains the only available resource for pyrethrins production. Because pyrethrins biosynthesis is derived from two independent pathways involving multiple enzymes,

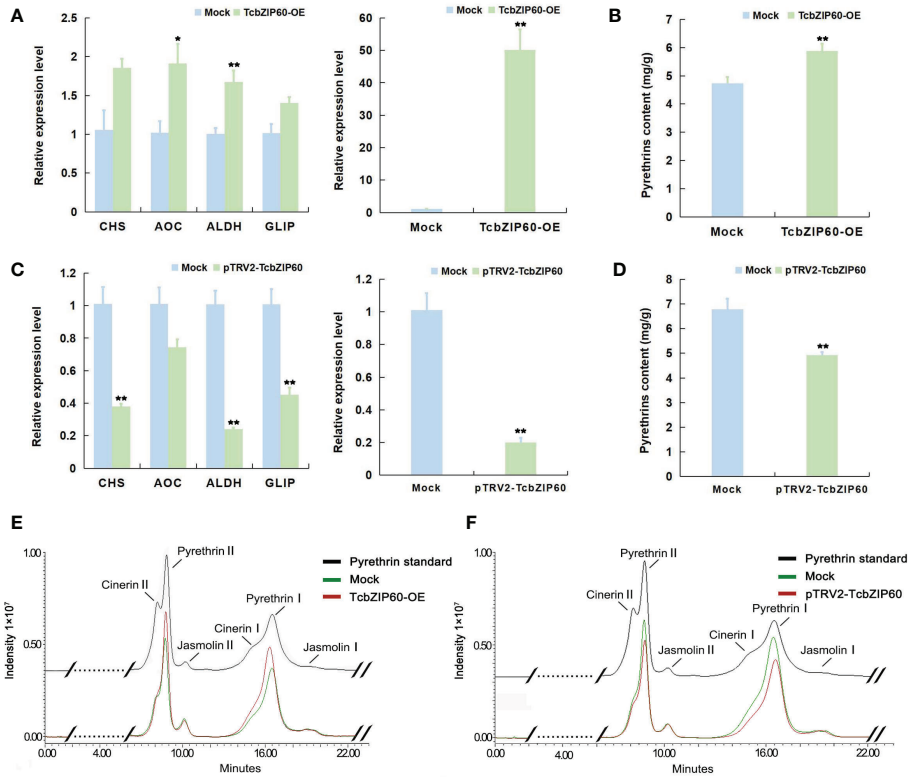


FIGURE 5
Transient overexpression and TRV-VIGS treatment of *TcbZIP60* in *T. cinerariifolium* leaves. **(A)** Relative expression levels of pyrethrin biosynthesis genes in *T. cinerariifolium* leaves at 4 days after transient overexpression of *TcbZIP60*. Mock: transient overexpression of the empty vector; *TcbZIP60*-OE: transient overexpression of pGreen-62-SK-*TcbZIP60*. The value for the Mock was set to 1. **(B)** Quantity of pyrethrins in *T. cinerariifolium* leaves after transient *TcbZIP60* overexpression at 4 days. The ordinate represents the pyrethrins content measured by high-performance liquid chromatography (HPLC) and the abscissa represents the six major constituent pyrethrins. **(C)** Relative expression levels of pyrethrins biosynthesis genes in *T. cinerariifolium* leaves at 4 days after TRV-VIGS treatment of *TcbZIP60* at 14 days. Mock: VIGS empty pTRV2 vector; pTRV2-*TcbZIP60*: *TcbZIP60* silencing of leaves. The value for Mock was set to 1. **(D)** Quantity of pyrethrins in *T. cinerariifolium* leaves after TRV-VIGS treatment of *TcbZIP60*. The ordinate represents pyrethrins content measured by HPLC and the abscissa represents the six major constituent pyrethrins. **(E)** Representative chromatogram of total pyrethrins showing the peaks of six different constituents (cinerarin II, pyrethrin II, jasmoline II, cinerarin I, pyrethrin I, and jasmoline I) in a standard sample (black line), *TcbZIP60*-OE (red line), and Mock (green line). **(F)** Representative chromatogram of total pyrethrins showing the peaks of six different constituents in a standard sample (black line), Mock (green line), and pTRV2-*TcbZIP60* (red line). Asterisks indicate that the value is significantly different from that of the control (**P<0.01, *P < 0.05).

overexpression of one or two enzymes would not be efficient in increasing pyrethrins biosynthesis, and it is even harder to co-express enzymes from different pathways simultaneously. Members of the bZIP TF family typically target the key steps in the biosynthesis of secondary metabolites. bZIP TFs contain a conserved alkaline region consisting of 16 amino acid residues with a constant N-X7-R/K motif, which can recognize and bind to the specific elements on the DNA sequence on the promoters, thereby affecting the multiple genes involved in the synthesis of secondary metabolites. Thus, overexpression of bZIP TFs leads to the improvement of secondary metabolites production. For example, in *A. annua*, AabZIP1 binds to both the *AaADS* and *AaCYP71AV1* promoters, and increases the expression of *AaADS* and *AaCYP71AV1* as well as the artemisinin content (Zhang et al., 2015). Overexpression of *AaHY5* and other upstream transcriptional regulatory genes significantly increased the transcriptional levels of the downstream genes *AaADS*, *AaCYP71AV1*, *AaDBR2*, and *AaALDH1*, and increased the synthesis and accumulation of secondary metabolites (Hao et al., 2019). In *Oryza sativa*, bZIP72 binds to the *AOC* promoter G-box, which strengthens the *AOC* transcription level and endogenous JA level (Wang et al., 2020). In *T. cinerariifolium*, *CHS* and *AOC* are upstream genes that are more likely to become rate-limiting factors for pyrethrins synthesis (Figure 1A). AOC is considered to be the first key enzyme in JA synthesis, affecting the overall biosynthesis rate (Yoeun et al., 2018), and CHS is the first key enzyme in the biosynthesis of pyrethrins (Hu et al., 2018). Our results demonstrated that TcbZIP60 could bind to the E-box and G-box cis-elements on the promoters of both *TcCHS* and *TcAOC*, and activated the transcription of these genes *in vitro*. Moreover, to further confirm the function of TcbZIP60 in pyrethrins production, we transiently overexpressed the *TcbZIP60* gene in *T. cinerariifolium* leaves, which upregulated the expression of the main pyrethrins biosynthesis genes *TcCHS*, *TcAOC*, *TcALDH*, and *TcGLIP*, consequently leading to the significant accumulation of pyrethrins. As expected, interference of *TcbZIP60* expression in *T. cinerariifolium* decreased the production of pyrethrins.

Based on these results, we propose a model of the role of TcbZIP60 in the regulation of pyrethrins biosynthesis in *T.*

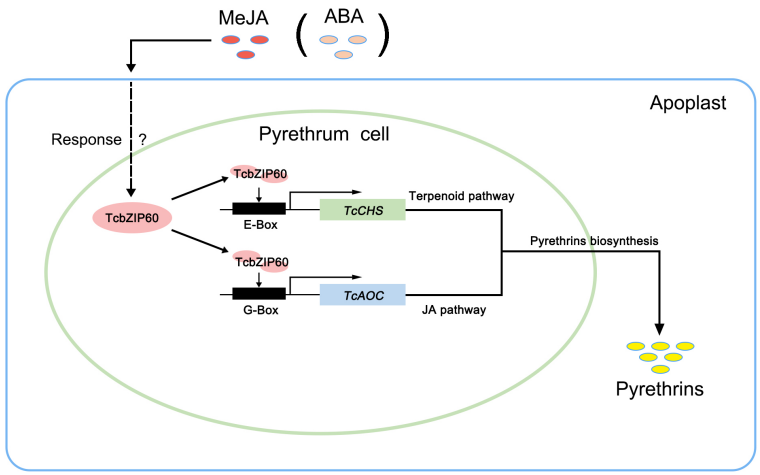


FIGURE 6
Model illustrating the involvement of TcbZIP60 in regulating pyrethrins biosynthesis in *T. cinerariifolium*. The arrows indicate TcbZIP60, regulated by methyl jasmonate (MeJA) or abscisic acid (ABA) signals, activating the *TcCHS* and *TcAOC* genes by binding to their promoters and subsequently regulating pyrethrins biosynthesis. JA, jasmonic acid.

cinerariifolium (Figure 6). This study thus identified TcbZIP60 as a novel and important positive regulator that could be used in engineering pyrethrins synthesis to improve its production in *T. cinerariifolium*.

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, TZ, JiaL, CW, and HH conceived the project. ZX, TZ, JinjL, and HH conducted the experiments. ZX, TZ, JiaL, LZ, JinjL, RZ, JLU, YW, CW, and HH discussed the experiment. ZX, TZ, JiaL, and HH analyzed and interpreted the data. ZX and HH prepared the figures. ZX and HH prepared the manuscript draft. ZX and TZ wrote the article with input from all authors. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Key National R & D Projects During the 14th Five-Year Plan Period (2019YFD1001500), China Postdoctoral Science Foundation (2022T150243, 2018 M64072), National Natural Science Foundation of China (31902051, 32160718), and Natural Science Research Project of Guizhou (KY [2022]170, ZK [2022]301).

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.2023.1133912/full#supplementary-material

References

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell.* 15, 63–78. doi: 10.1105/tpc.006130

Casida, J. E. (1973). *Pyrethrum: The natural insecticide* (New York: Academic Press).

Casida, J. E., and Quistad, G. B. (1995). *Pyrethrum flowers: Production, chemistry, toxicology, and uses* (New York: Oxford University Press).

Chen, S. Y., Ma, T., Song, S. R., Li, X. L., Fu, P. N., Wu, W., et al. (2021). Arabidopsis downy mildew effector HaRxLL470 suppresses plant immunity by attenuating the DNA-

binding activity of bZIP transcription factor HY5. *New Phytol.* 230, 1562–1577. doi: 10.1111/nph.17280

Chen, K., Tang, W., Zhou, Y., Chen, J., Xu, Z., Ma, R., et al. (2022). AP2/ERF transcription factor GmDREB1 confers drought tolerance in transgenic soybean by interacting with GmERFs. *Plant Physiol. Biochem.* 170, 287–295. doi: 10.1016/j.plaphy.2021.12.014

Chow, C. N., Lee, T. Y., Hung, Y. C., Li, G.-Z., Tseng, K.-C., Liu, Y.-H., et al. (2019). PlantPAN3. 0: A new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Res.* 47, D1155–D1163. doi: 10.1093/nar/gky1081

Creelman, R. A., and Mullet, J. E. (1997). Biosynthesis and action of jasmonates in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 48, 355–381. doi: 10.1146/annurev.arplant.48.1.355

Freemont, J. A., Littler, S. W., Hutt, O. E., Mauger, S., and Duggan, P. J. (2016). Molecular markers for pyrethrin autoxidation in stored pyrethrum crop: analysis and structure determination. *J. Agric. Food Chem.* 64, 7134–7141. doi: 10.1021/acs.jafc.6b02959

Geng, J., and Liu, J. H. (2018). The transcription factor CsbHLH18 of sweet orange functions in modulation of cold tolerance and homeostasis of reactive oxygen species by regulating the antioxidant gene. *J. Exp. Bot.* 69, 2677–2692. doi: 10.1093/jxb/ery065

Hao, X., Zhong, Y., Nützmann, H.-W., Fu, X., Yan, T., Shen, Q., et al. (2019). Light-induced artemisinin biosynthesis is regulated by the bZIP transcription factor AaHY5 in *Artemisia annua*. *Plant Cell Physiol.* 60, 1747–1760. doi: 10.1093/pcp/pcz084

Hu, H., Li, J. J., Delatte, T., Vervoort, J., Gao, L., Verstappen, F., et al. (2018). Modification of chrysanthemum odour and taste with chrysanthemol synthase induces strong dual resistance against cotton aphids. *Plant Biotechnol. J.* 16, 1434–1445. doi: 10.1111/pbi.12885

Jakoby, M., Weisshaar, B., Drge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., et al. (2002). bZIP transcription factors in arabidopsis. *Trends Plant Sci.* 7, 106–111. doi: 10.1016/S1360-1385(01)02223-3

Jing, F., Zhang, L., Li, M., Tang, Y., Wang, Y., Wang, Y., et al. (2009). Abscisic acid (ABA) treatment increases artemisinin content in artemisia annua by enhancing the expression of genes in artemisinin biosynthetic pathway. *Biologia.* 64, 319–323. doi: 10.2478/s11756-009-0040-8

Jung, S.-K., McDonald, K. A., and Dandekar, A. M. (2015). Effect of leaf incubation temperature profiles on *Agrobacterium tumefaciens*-mediated transient expression. *Biotechnol. Progr.* 31, 783–790. doi: 10.1002/btpr.2077

Katsuda, Y. (1999). Development of and future prospects for pyrethroid chemistry. *Pesticide Sci.* 55, 775–782. doi: 10.1002/(sici)1096-9063(199908)55:8<775::Aid-ps27>3.0.Co;2-n

Khan, S. A., Verma, P., Banerjee, S., Chaterjee, A., Tandon, S., Kalra, A., et al. (2017). Pyrethrin accumulation in elicited hairy root cultures of *Chrysanthemum cinerariaefolium*. *Plant Growth Regul.* 81, 365–376. doi: 10.1007/s10725-016-0213-8

Kikuta, Y., Ueda, H., Takahashi, M., Mitsumori, T., Yamada, G., Sakamori, K., et al. (2012). Identification and characterization of a GDSL lipase-like protein that catalyzes the ester-forming reaction for pyrethrin biosynthesis in *Tanacetum cinerariifolium*- a new target for plant protection. *Plant J.* 71, 183–193. doi: 10.1111/j.1365-313X.2012.04980.x

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lackman, P., González-Guzmán, M., Tilleman, S., Carqueijeiro, I., Pérez, A. C., Moses, T., et al. (2011). Jasmonate signaling involves the abscisic acid receptor PYL4 to regulate metabolic reprogramming in arabidopsis and tobacco. *Proc. Natl. Acad. Sci.* 108, 5891–5896. doi: 10.1073/pnas.1103010108

Lange, W. H., and Akesson, N. B. (1973). "Chapter 14," in *Pyrethrum for control of agricultural insects* (Amsterdam: Elsevier Inc).

Li, J., Li, Y., Wang, Z., and Wang, C. (2014). Effect of three growth regulators on pyrethrin content and the traits related to pyrethrin yield. *Chin. J. Trop. Crops.* 35, 1067–1070.

Li, J. J., Liu, H., Yang, C., Wang, J., Yang, G. J., Si, P., et al. (2020). Genome-wide identification of MYB genes and expression analysis under different biotic and abiotic stresses in *Helianthus annuus* l. *Indust. Crops Prod.* 143, 111924. doi: 10.1016/j.indcrop.2019.111924

Li, W., Lybrand, D. B., Zhou, F., Last, R. L., and Pichersky, E. (2019). Pyrethrin biosynthesis: the cytochrome P450 oxidoreductase CYP82Q3 converts jasmolone to pyrethrolone. *Plant Physiol.* 181, 934–944. doi: 10.1104/pp.19.00499

Li, W., Zhou, F., and Pichersky, E. (2018). Jasmone hydroxylase, a key enzyme in the synthesis of the alcohol moiety of pyrethrin insecticides. *Plant Physiol.* 177, 1498–1509. doi: 10.1104/pp.18.00748

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.* 408, 402–408. doi: 10.1006/meth.2001.1262

Luo, J., Ma, N., Pei, H., Chen, J., Li, J., and Gao, J. (2013). A DELLA gene, RhGAI1, is a direct target of EIN3 and mediates ethylene-regulated rose petal cell expansion *via* repressing the expression of RhCesA2. *J. Exp. Botany.* 64, 5075–5084. doi: 10.1093/jxb/ert296

Lv, Z., Guo, Z., Zhang, L., Zhang, F., Jiang, W., Shen, Q., et al. (2019). Interaction of bZIP transcription factor TGA6 with salicylic acid signaling modulates artemisinin biosynthesis in artemisia annua. *J. Exp. Botany.* 70, 3969–3979. doi: 10.1093/jxb/erz166

Lybrand, D. B., Xu, H., Last, R. L., and Pichersky, E. (2020). How plants synthesize pyrethrins: Safe and biodegradable insecticides. *Trends Plant Sci.* 25, 1240–1251. doi: 10.1016/j.tplants.2020.06.012

Mao, Z., Jiang, H., Wang, S., Wang, Y., Yu, L., Zou, Q., et al. (2021). The MdHY5-MdWRKY41-MdMYB transcription factor cascade regulates the anthocyanin and proanthocyanidin biosynthesis in red-fleshed apple. *Plant Sci.* 306, 110848. doi: 10.1016/j.plantsci.2021.110848

Orenstein, A. J. (1913). Mosquito catching in dwellings in the prophylaxis of malaria. *Am. J. Public Health* 3, 106. doi: 10.2105/AJPH.3.2.106

Pal, T., Malhotra, N., Chanumolu, S. K., and Chauhan, R. S. (2015). Next-generation sequencing (NGS) transcriptomes reveal association of multiple genes and pathways contributing to secondary metabolites accumulation in tuberous roots of *Aconitum heterophyllum* wall. *Planta.* 242, 239–258. doi: 10.1007/s00425-015-2304-6

Ramirez, A. M., Saillard, N., Yang, T., Franssen, M. C., Bouwmeester, H. J., and Jongsma, M. A. (2013). Biosynthesis of sesquiterpene lactones in pyrethrum (*Tanacetum cinerariifolium*). *PloS One* 8, e65030. doi: 10.1371/journal.pone.0065030

Rivera, S. B., Swedlund, B. D., King, G. J., Bell, R. N., Hussey, C. E., Shattuck-Eidens, D. M., et al. (2001). Chrysanthemyl diphosphate synthase: Isolation of the gene and characterization of the recombinant non-head-to-tail monoterpene synthase from *Chrysanthemum cinerariaefolium*. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4373–4378. doi: 10.1073/pnas.071543598

Sagor, G., Chawla, P., Kim, D. W., Berberich, T., Kojima, S., Niitsu, M., et al. (2015). The polyamine spermine induces the unfolded protein response *via* the MAPK cascade in arabidopsis. *Front. Plant Sci.* 6, 687. doi: 10.3389/fpls.2015.00687

Senthil-Kumar, M., and Mysore, K. S. (2014). Tobacco rattle virus-based virus-induced gene silencing in *Nicotiana benthamiana*. *Nat. Protoc.* 9, 1549–1562. doi: 10.1038/nprot.2014.092

Shu, G., Tang, Y., Yuan, M., Wei, N., Zhang, F., Yang, C., et al. (2022). Molecular insights into AabZIP1-mediated regulation on artemisinin biosynthesis and drought tolerance in *Artemisia annua*. *Acta Pharmaceut. Sin. B.* 12, 1500–1513. doi: 10.1016/j.apsb.2021.09.026

Sibéril, Y., Benhamron, S., Memelink, J., Giglioli-Guivarc'h, N., Thiersault, M., Boisson, B., et al. (2001). *Catharanthus roseus* G-box binding factors 1 and 2 act as repressors of strictosidine synthase gene expression in cell cultures. *Plant Mol. Biol.* 45, 477–488. doi: 10.1023/A:1010650906695

Song, W. C., Funk, C. D., and Brash, A. R. (1993). Molecular cloning of an allene oxide synthase: a cytochrome-p450 specialized for the metabolism of fatty-acid hydroperoxides. *Proc. Natl. Acad. Sci. U. S. A.* 90, 8519–8523. doi: 10.1073/pnas.90.18.8519

Sornaraj, P., Luang, S., Lopato, S., and Hrmova., M. (2016). Basic leucine zipper (bZIP) transcription factors involved in abiotic stresses: A molecular model of a wheat bZIP factor and implications of its structure in function. *Biochim. Biophys. Acta Gen. Subj.* 1860, 46–56. doi: 10.1016/j.bbagen.2015.10.014

Sultana, S., Hu, H., Gao, L., Mao, J., Luo, J., Jongsma, M. A., et al. (2015). Molecular cloning and characterization of the trichome specific chrysanthemyl diphosphate/chrysanthemol synthase promoter from *Tanacetum cinerariifolium*. *Scient. Horticult.* 185, 193–199. doi: 10.1016/j.scienta.2015.01.032

Tang, W., and Page, M. (2013). Transcription factor AtbZIP60 regulates expression of Ca2+-dependent protein kinase genes in transgenic cells. *Mol. Biol. Rep.* 40, 2723–2732. doi: 10.1007/s11033-012-2362-9

Tijet, N., and Brash, A. R. (2002). Allene oxide synthases and allene oxides. *Prostagland. Other Lipid Mediat.* 68-9, 423–431. doi: 10.1016/s0090-6980(02)00046-1

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2000). Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11632–11637. doi: 10.1073/pnas.190309197

Varga, F., Jeran, N., Satovic, Z., Biosic, M., and Grdisa, M. (2021). High diversity of natural *Dalmatian pyrethrum* based on pyrethrin composition at intra- and interpopulation level. *Phytochemistry.* 192, 11. doi: 10.1016/j.phytochem.2021.112934

Wang, Y., Hou, Y., Qiu, J., Wang, H., Wang, S., Tang, L., et al. (2020). Abscisic acid promotes jasmonic acid biosynthesis *via* a 'SAPK10-bZIP72-AOC' pathway to synergistically inhibit seed germination in rice (*Oryza sativa*). *New Phytol.* 228, 1336–1353. doi: 10.1111/nph.16774

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., et al. (2005). Integration of spatial and temporal information during floral induction in arabidopsis. *Science.* 309, 1056–1059. doi: 10.1126/science.1114358

Wolfgang, D. L., Snoek, B. L., Berend, S., and Christoph, W. (2018). The arabidopsis bZIP transcription factor family-an update. *Curr. Opin. Plant Biol.* 45, 36–49. doi: 10.1016/j.pbi.2018.05.001

Xu, H., Li, W., Schilmiller, A. L., van Eekelen, H., de Vos, R. C. H., Jongsma, M. A., et al. (2019). Pyrethric acid of natural pyrethrin insecticide: complete pathway elucidation and reconstitution in *Nicotiana benthamiana*. *New Phytol.* 223, 751–765. doi: 10.1111/nph.15821

Xu, H., Lybrand, D., Bennewitz, S., Tissier, A., Last, R. L., and Pichersky, E. (2018b). Production of trans-chrysanthemic acid, the monoterpene acid moiety of natural pyrethrin insecticides, in tomato fruit. *Metab. Eng.* 47, 271–278. doi: 10.1016/j.ymben.2018.04.004

Xu, H., Moghe, G. D., Wiegert-Rininger, K., Schilmiller, A. L., Barry, C. S., Last, R. L., et al. (2018a). Coexpression analysis identifies two oxidoreductases involved in the biosynthesis of the monoterpene acid moiety of natural pyrethrin insecticides in *Tanacetum cinerariifolium*. *Plant Physiol.* 176, 524–537. doi: 10.1104/pp.17.01330

Yang, T., Gao, L., Hu, H., Stoopen, G., Wang, C., and Jongsma, M. A. (2014). Chrysanthemyl diphosphate synthase operates in planta as a bifunctional enzyme with chrysanthemol synthase activity. *J. Biol. Chem.* 289, 36325–36335. doi: 10.1074/jbc.M114.623348

Yoeun, S., Cho, K., and Han, O. (2018). Structural evidence for the substrate channeling of rice allene oxide cyclase in biologically analogous nazarov reaction. *Front. Chem.* 6, 500. doi: 10.3389/fchem.2018.00500

Zeng, T., Li, J.-W., Xu, Z.-Z., Zhou, L., Li, J.-J., Yu, Q., et al. (2022). TcMYC2 regulates pyrethrin biosynthesis in *Tanacetum cinerariifolium*. *Hortic. Res.* 9, uhac178. doi: 10.1093/hr/uhac178

Zeng, T., Li, J.-W., Zhou, L., Xu, Z.-Z., Li, J.-J., Hu, H., et al. (2021). Transcriptional responses and GCMS analysis for the biosynthesis of pyrethrins and volatile terpenes in *Tanacetum coccineum*. *Int. J. Mol. Sci.* 22, 13005. doi: 10.3390/ijms222313005

Zhang, F., Fu, X., Lv, Z., Lu, X., Shen, Q., Zhang, L., et al. (2015). A basic leucine zipper transcription factor, AabZIP1, connects abscisic acid signaling with artemisinin biosynthesis in *Artemisia annua*. *Mol. Plant* 8, 163–175. doi: 10.1016/j.molp.2014.12.004

Zhang, Y., Zheng, S., Liu, Z., Wang, L., and Bi, Y. (2011). Both HY5 and HYH are necessary regulators for low temperature-induced anthocyanin accumulation in arabidopsis seedlings. *J. Plant Physiol.* 168, 367–374. doi: 10.1016/j.jplph.2010.07.025

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 27 February 2023
DOI 10.3389/fpls.2023.1098280

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Mingzhi Zhu,
Hunan Agricultural University, China
Qi Tang,
Hunan Agricultural University, China

*CORRESPONDENCE
Likai Chen
chenlk@gzucm.edu.cn

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 14 November 2022
ACCEPTED 14 February 2023
PUBLISHED 27 February 2023

CITATION
Wang X, Zhong L, Zou X, Gong L, Zhuang J, Zhang D, Zheng H, Wang X, Wu D, Zhan R and Chen L (2023) GC-MS and UHPLC-QTOFMS-assisted identification of the differential metabolites and metabolic pathways in key tissues of *Pogostemon cablin*. *Front. Plant Sci.* 14:1098280. doi: 10.3389/fpls.2023.1098280

GC-MS and UHPLC-QTOFMS-assisted identification of the differential metabolites and metabolic pathways in key tissues of *Pogostemon cablin*

Xiaobing Wang[1], Liting Zhong[1], Xuan Zou[1], Lizhen Gong[1], Jiexuan Zhuang[1], Danhua Zhang[1], Hai Zheng[2], Xiaomin Wang[1], Daidi Wu[1], Ruoting Zhan[1,3] and Likai Chen[1,3]*

[1]Research Center of Chinese Herbal Resource Science and Engineering, Guangzhou University of Chinese Medicine, Key Laboratory of Chinese Medicinal Resource from Lingnan (Guangzhou University of Chinese Medicine), Ministry of Education, Joint Laboratory of National Engineering Research Center for the Pharmaceutics of Traditional Chinese Medicines, Guangzhou, China, [2]School of Pharmaceutical Sciences, Guangdong Food and Drug Vocational College, Guangzhou, China, [3]Maoming Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Maoming, Guangdong, China

Pogostemon cablin is an important aromatic medicinal herb widely used in the pharmaceutical and perfume industries. However, our understanding of the phytochemical compounds and metabolites within *P. cablin* remains limited. To our knowledge, no integrated studies have hitherto been conducted on the metabolites of the aerial parts of *P. cablin*. In this study, twenty-three volatile compounds from the aerial parts of *P. cablin* were identified by GC-MS, predominantly sesquiterpenes. Quantitative analysis showed the highest level of patchouli alcohol in leaves (24.89 mg/g), which was 9.12 and 6.69-fold higher than in stems and flowers. UHPLC-QTOFMS was used to analyze the non-volatile compounds of leaf, stem and flower tissues. The differences in metabolites between flower and leaf tissues were the largest. Based on 112, 77 and 83 differential metabolites between flower-leaf, flower-stem and leaf-stem, three tissue-specific biomarkers of metabolites were identified, and the differential metabolites were enriched in several KEGG pathways. Furthermore, labeling differential metabolites in the primary and secondary metabolic pathways showed that flowers accumulated more lipids and amino acids, including proline, lysine and tryptophan; the leaves accumulated higher levels of terpenoids, vitamins and flavonoids, and stems contained higher levels of carbohydrate compounds. Based on the role of acetyl coenzyme A, the distribution and possible exchange mechanism of metabolites in leaves, stems and flowers of *P. cablin* were mapped for the first time, laying the groundwork for future research on the metabolites in *P. cablin* and their regulatory role.

KEYWORDS

metabolism pathway, metabolite biomarker, metabolomics, phytochemical, *Pogostemon cablin*

Introduction

Pogostemon cablin (*P. cablin*) is a member of the Lamiaceae family, native to tropical areas such as Malaysia, and widely cultivated in Guangdong, Guangxi and Hainan provinces in China (Hu et al., 2006). The aerial parts of *P. cablin* consists of leaves, stems and flowers. The leaves and stems of *P. cablin* have a strong aromatic taste after maceration, substantiating that it is rich in volatile components, such as sesquiterpenes. In addition, leaves and stems are rich in non-volatile substances, such as flavonoids and alkaloids. However, the metabolites in *P. cablin* flowers have been rarely reported since they rarely bloom. *P. cablin* is not only used as the fixative of various perfumes but also as the raw material of various cosmetics and oral hygiene products (Sandes et al., 2016). Interestingly, it is also used in traditional Chinese medicine to treat colds, nausea, diarrhea, headache and fever (Liu et al., 2017), suggesting that the market demand is huge. One study found that there are many active components in *P. cablin* (Hu et al., 2006), and specific compounds have been gradually revealed in recent years, including terpenoids (Chen et al., 2019), flavonoids (Miyazawa et al., 2000), phytosterols, organic acids (Xie et al., 2022), phenols, alkaloids and glycosides (Swamy and Sinniah, 2015). The accumulation and exchange of these metabolites are reportedly closely related to the biological function of *P. cablin*; however, the underlying process remains unknown.

Patchouli alcohol represents a major bioactive and aromatic compound in *P. cablin*, widely used as an indicator of sample quality. Besides, patchouli alcohol is a natural sesquiterpenoid. In plants, terpenoids are usually produced by the mevalonate pathway (MVA) and methylerythritol phosphate (MEP) pathway. The cytoplasmic MVA pathway is the main pathway for synthesizing sesquiterpenes, and its original substrate is acetyl-CoA. It has been established that AACT primarily catalyzes acetyl-CoA to produce acetoacetyl-CoA, which is critical for the biosynthesis of the steroid backbone (Wang et al., 2017). Isopentenyl diphosphate is an intermediate substance linking the MVA and MEP pathways (Henry et al., 2018). The MVA pathway eventually produces sesquiterpenes and triterpenes, while the MEP pathway produces diterpenes and other steroids (Ahmed et al., 2016; Schwarz et al., 2018). However, the tissue-specific differences in the metabolic pathways responsible for synthesizing the active ingredients in *P. cablin* remain unknown.

The past decade has witnessed unprecedented medical advances with the advent of metabolomics which can be harnessed to study the accumulation of medicinally active ingredients in different conditions. Metabolomic analyses rely on research methods, such as gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), and ultra-high-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry (UHPLC-QTOFMS), and each method has unique advantages and disadvantages (Patterson et al., 2010). In a comprehensive study of lettuce metabolites in different cultivation environments using GC-MS and LC-MS, the researchers found that the level of amino acids such as lysine, phenylalanine, tryptophan and valine was significantly increased in hydroponically grown leaf lettuce, while soil-cultivation derived leaf lettuce samples contained significantly higher levels of fatty-acid derived alcohols and lettuce specific sesquiterpene lactones. In addition, the difference in metabolite content endowed lettuce with different nutritional components and tastes (Tamura et al., 2018). Moreover, other vegetables have been subjected to similar analyses. For example, the quality of three different cultivars of cynara scolymus (artichoke) was analyzed; UHPLC and qTOF-MS results showed the content of caffeic acid derivatives, flavonoids and fatty acid was different among cultivars (Farag et al., 2013). Ultrahigh-pressure liquid chromatography-high-resolution mass spectrometry (UHPLC-HRMS) has been used to evaluate the freshness of egg products, and 31 compounds have been identified as useful markers of egg freshness (Cavanna et al., 2018). Metabolomics methods have also been used to study the differences in metabolites produced under stress conditions in different tissues. Interestingly, an integrated GC-MS and LC-MS analysis revealed that disease-resistant tomatoes contained significantly higher concentrations of acyl sugars (Firdaus et al., 2013). Moreover, stress- related substances have been documented on the waxy cuticles of wheat leaves and stems. GC-MS has been used to analyze the metabolite composition of wheat epidermal wax and found that leaves contained more primary alcohol than stems (Lavergne et al., 2018). In addition, an integrated GC-MS and LC-MS metabonomics study of pumpkins showed a significant difference in the chemical composition of male and female nectaries (Chatt et al., 2018). A recent multivariate statistical analysis of Corni Fructus based on UHPLC-QTOFMS identified 17 different compounds between raw and processed products, revealing the relationship between compounds and the color of Corni Fructus and the crucial compounds for color (Qian et al., 2022).

Herein, GC-MS and UHPLC-QTOFMS were used to study the volatile and non-volatile substances in the leaves, stems and flowers of *P. cablin* and their metabolic pathways. We carried out a tissue-specific analysis of the primary and secondary metabolites, and the metabolites in the flowers of this species were analyzed for the first time. Next, metabolomics and pathway analyses were performed to assess the phytochemical panorama and metabolism pathway connectivity. We reported the relative metabolite contents of different tissues and characterized the metabolic networks present in *P. cablin*. Based on the above contents, this study mapped the distribution and possible exchange mechanism of metabolites in key tissues of *P. cablin*, which provides the foothold for further research on metabolites and the regulation of this important aromatic medicinal plant.

Materials and methods

Plant materials

P. cablin used in this study was collected at the flowering stage from Shizhen Mountain at Guangzhou University of Chinese Medicine (23.03°N, 113.23°E) in Guangdong province, China.

The plants were cultured in a growth chamber in our laboratory. For GC-MS analysis, *P. cablin* leaves, stems and flowers were collected and dried in an oven at 40 °C. The same samples were collected again and immediately placed in liquid nitrogenfor UHPLC-QTOFMS analysis. All samples were kept in our laboratory.

Analysis of volatile compounds in *P. cablin* tissues using GC-MS

0.25 g sample powder of leaves, stems and flowers was accurately weighed. Extracted twice with 50 mL ethyl acetate for 20 min each time using an ultrasonic cleaner, and then concentrate the filtrate by rotating evaporation. The concentrate was dissolved in hexane and diluted in a 5 mL volumetric flask. 0.22 μM organic membranes were used to filter the solution, and the filtrate was transferred to a new sample injection bottle for GC-MS analysis using an Agilent 7890B gas chromatograph with 5977A inert mass selective detector (Agilent, United States). The injection volume was 1 μL, and helium was the carrier gas. The Agilent HP-5MS column (30 m × 0.25 mm × 0.25 μm film thickness) was used for separation. The GC oven temperature was programmed at an initial temperature of 50 °C for 2 min with an increase of 20 °C/min to 130 °C and increased to 150 °C at a rate of 2 °C/min for 5 min. The temperature was then increased to 230°C at 20 °C/min. Quadrupole and ion source temperatures are set to 150 °C and 230 °C. The NIST14/Wiley275 mass spectral library was used for metabolite identification. The external standard method was used to quantify the content of patchouli alcohol in leaves, stems and flowers. Three biological replicates and two technical replicates were performed for each tissue in this study.

Extraction of non-volatile compounds in *P. cablin* tissues

50 mg samples were extracted from 1 mL of extraction solution containing an internal standard (*V* methanol: *V* acetonitrile: *V* water = 2:1:1, which was kept at -20 °C before extraction). The sample was homogenized in a ball mill for 4 minutes at 45 Hz and then ultrasound treated for 5 minutes (incubated in ice water). After homogenization for 3 times, the sample was incubated at -20 °C for 1 h to precipitate proteins. Then, the sample was centrifuged at 12000 rpm for 15 min at 4 °C and the supernatant (500 μL) was transferred into EP tubes. After drying the extracts in a vacuum concentrator without heating, the sample was reconstituted in 200 μL extraction liquid (*V* acetonitrile: *V* water = 1:1). The dry extract was vortexed for 30 s and sonicated 10 min (4 °C water bath). After centrifuging for 15 min at 12000 rpm at 4 °C, the supernatant (60 μL) was transferred into a fresh 2 mL LC/MS glass vial for the UHPLC-QTOFMS analysis. Six biological replicates were performed for each tissue in this study.

Analysis of non-volatile compounds in *P. cablin* tissues using UHPLC-QTOFMS

LC-MS/MS analyses were performed using an UHPLC system (1290, Agilent Technologies) with a UPLC BEH amide column (1.7 μm × 2.1 mm × 100 mm, Waters) coupled to TripleTOF 5600 (Q-TOF, AB Sciex). The mobile phase consisted of 25 mM NH_4OAc and 25 mM NH_4OH in water (pH=9.75) (A) and acetonitrile (B) was carried with elution gradient as follows: 0 min, 95% B; 7 min, 65% B; 9 min, 40% B; 9.1 min, 95% B; 12 min, 95% B, which was delivered at 0.5 mL min^{-1}. The injection volume was 2 μL. The Triple TOF mass spectrometer was used for its ability to acquire MS/MS spectra on an information-dependent basis during an LC/MS experiment. Using this mode, the acquisition software (Analyst TF 1.7, AB Sciex) continuously evaluates the full scan survey MS data as it collects and triggers the acquisition of MS/MS spectra depending on preselected criteria. During each cycle, 12 precursor ions with an intensity greater than 100 were chosen for fragmentation at collision energy (CE) of 30 V (15 MS/MS events with each product ion accumulation for 50 msec). ESI source conditions were set as follows: Ion source gas 1 as 60 Psi, Ion source gas 2 as 60 Psi, Curtain gas as 35 Psi, source temperature 650 °C, Ion Spray Voltage Floating (ISVF) 5000 V or -4000 V in positive or negative modes, respectively.

Data preprocessing and annotation

MS raw data files were converted to the mzXML format using ProteoWizard, and processed by R package XCMS (version 3.2). The preprocessing results generated a data matrix that included the retention time, mass-to-charge ratio (m/z) values, and peak intensity. The R package CAMERA was used for peak annotation after XCMS data processing. An in-house MS2 database was applied for metabolites identification. A data matrix, including the metabolite name (putatively identified by UHPLC-QTOFMS), sample information (six biological repeats for each sample), and raw abundance (peak area for each putatively identified metabolite), was generated, uploaded to MetaboAnalyst 3.0, and analyzed according to the instructions provided. On the Data Normalization page, we selected “Normalization by sum” for sample normalization, “None” for data transformation, and “Auto-scaling” for data scaling. The filtered data were loaded into Metabo-Analyst 3.0 for multivariate statistical analyses, including principal components analysis (PCA) and partial least squares discriminant analysis (PLS-DA). Discriminating compounds were then validated using a false discovery rate (FDR) test with $P < 0.05$, and the thresholds of fold-change ≥ 2 and q-value ≤ 0.05 for metabolites were considered to indicate differentially expressed metabolites. Differentially metabolites were analyzed by hierarchical clustering, and similarity assessment for clustering was based on the Euclidean distance coefficient. KEGG (http://www.kegg.jp/) was used to analyze the metabolic pathway.

10.3389/fpls.2023.1098280

Results

Determination of the volatile compounds in *P. cablin* leaves, stems, and flowers using GC-MS

It is well-acknowledged that *P. cablin* rarely blooms. Fortunately, in the present study, *P. cablin* blossomed in spring, and the characteristics of leaf, stem and flower tissues were recorded (Figure 1A). The leaf of *P. cablin* is broadly ovate, with irregular tooth fissure. The stem is covered with villus; the verticillaster has ten to more flowers, forming a spike, and the corolla is purple. Then the tissues were collected, and GC-MS was selected to detect and quantify the volatile compounds present in *P. cablin*. Differences were observed in the volatile compounds in the three tissues of *P. cablin*. 11, 2 and 7 specific volatile compounds (Figure 1B) and 6 common compounds (Figure 1C) were identified in leaf, stem and flower tissues, respectively. 6 common compounds were as follows: α-patchoulene, β-patchoulene, seychellene, α-guaiene, caryophyllene and patchouli alcohol. As shown in Figure 1, δ-Guaiene was detected in leaves and flowers, while selinene and 1-ethynyl-2-methyl-1(E)-cyclododecene were found in stems and flowers. Furthermore, 10 compounds, including β-elemene, humulene, γ-patchoulene, valencene, α-gurjunene, aciphyllene, norpatchoulenol, spathulenol, globulol and humulane-1,6-dien-3-ol, were specifically detected in leaves. Moreover, 4 compounds, including 2-vinylfuran, pogostol, neophytadiene and phytol, were exclusively detected in flower tissues. All compounds found in leaves and stems were sesquiterpenes or derivatives, while two diterpenes (neophytadiene and phytol) were identified in flowers.

The volatile compounds released from all three selected tissues suggested that patchouli alcohol is the primary compound in *P. cablin*. To accurately measure the content of patchouli alcohol in leaves, stems and flowers, the calibration curve generated with the patchouli alcohol standard was linear in the 0.01~1.14 mg/mL range, and the correlation coefficient was 0.99. GC-MS revealed the highest levels of patchouli alcohol in the leaves (24.89 ± 0.40 mg/g DW). On the contrary, the content of patchouli alcohol in stems (2.72 ± 0.07 mg/g DW) and flowers (3.71 ± 0.07 mg/g DW) was lower than in leaves (Table S1; Figure 1D, E).

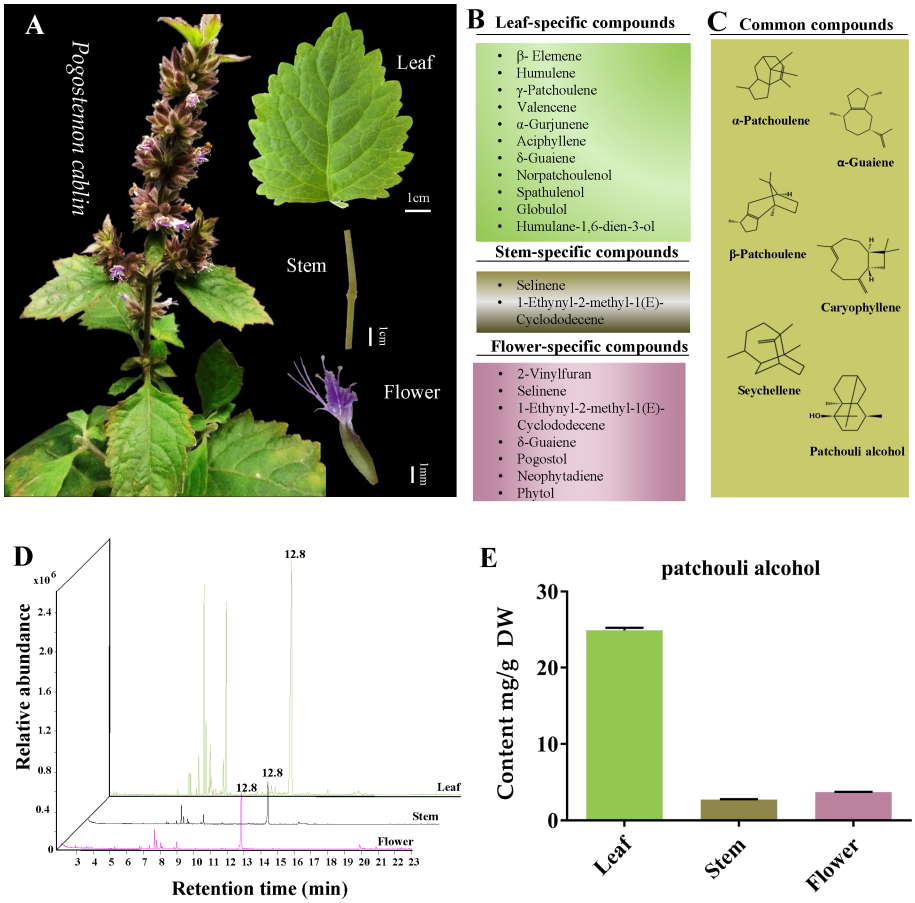


FIGURE 1
The tissue characteristics and volatile components analysis of *Pogostemon cablin*. **(A)** The phenotypic characteristics of *P. cablin* plants and their leaf, stem and flower. **(B)** Specific compounds in leaf, stem and flower. **(C)** Common compounds in leaf, stem and flower. **(D)** Gas chromatography-mass spectrometer (GC-MS) chromatograms of samples from the leaf, stem and flower, the retention time of patchouli alcohol is 12.8 min. **(E)** The content of patchouli alcohol in leaf, stem and flower. Error bars represent 95% confidence intervals. DW stands for dry weight.

Metabolite assessment in leaves, stems, and flowers using UHPLC-QTOFMS

To analyze the non-volatile components and metabolism of *P. cablin*, UHPLC-QTOFMS was used to evaluate the abundance of metabolites in the leaves, stem and flower tissues of *P. cablin*. A total of 2817 and 3250 relative retention time-quantitative peaks were generated by positive and negative ion modes analyses in *P. cablin* (Figure S1). The raw abundance of each metabolite was calculated based on the peak area. Data were normalized using an internal standard (2-chloro-L-phenylalanine, Shanghai Hengbai Biotech Co., Ltd.), and the mean of the normalized peak area data was log2 transformed. After data normalization (Figure S2), principal component analysis (PCA) was performed. The PCA plot exhibited an obvious separation pattern among tissues in both ion modes (Figure 2). The first and second principal component (PC1 and PC2) scores of three different tissues were 36.4% and 22.5%, respectively (Figure 2A). As shown in Figure 2B, PC1 and PC2 accounted for 55.6% and 7.8% of the total variance within the flower-leaf group in positive ion mode (POS) and 47.1% and 10.6% in negative ion mode (NEG). Furthermore, PC1 and PC2 of the flower-stem group were 43.8% and 11.1% in POS mode and 37.9% and 12.7% in NEG mode, while the PC1 and PC2 of the leaf-stem group were 39.1% and 11.9% in the POS mode and 39.6% and 13.5% in the NEG mode. The results showed significant differences between groups, suggesting significant differences in metabolites among the flower-leaf, flower-stem, and leaf-stem groups.

Metabolite variation assessed by PLS-DA, UVA, and clustering analysis

To further analyze the differences of metabolites in leaves, stems and flowers of *P. cablin*, partial least squares discriminant analysis (PLS-DA) was used to examine differences in secondary metabolism. Unlike PCA, PLS-DA is a supervised discriminant analysis method that uses partial least squares regression to simulate the relationship between metabolite expression levels. Features with variable importance in the projection (VIP) scores greater than 1 are important for the projection of the first principal component of the PLS-DA model and can be used as the criteria for feature screening. In our study, metabolites with a VIP >1 in the PLS-DA model were identified, and the top 15 most important metabolite compounds ranked by VIP scores are shown in Figure 3. By comparing the top 15 candidate compounds in the positive and negative ion modes between tissues, it was found that L-glutamine and raffinose had high VIP scores in leaf, stem and flower, which could be used to distinguish the three tissues. Compounds more

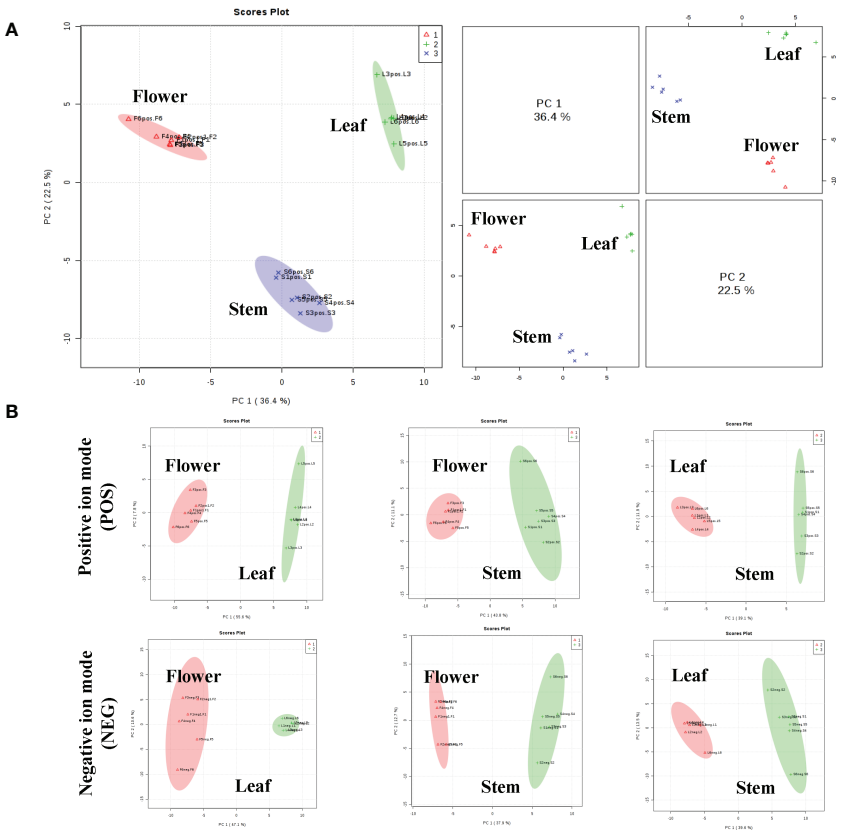


FIGURE 2
The principal component analysis (PCA) scatter plots between the three tissues of *P. cablin*. **(A)** PCA scatter plot of leaf (green circle), stem (blue circle) and flower (red circle) in positive ion mode (POS). **(B)** The PCA between any two tissues in *P. cablin* in positive and negative ion modes, respectively. The PCA was depends on the LC-MS data from all groups in positive and negative ion modes.

10.3389/fpls.2023.1098280

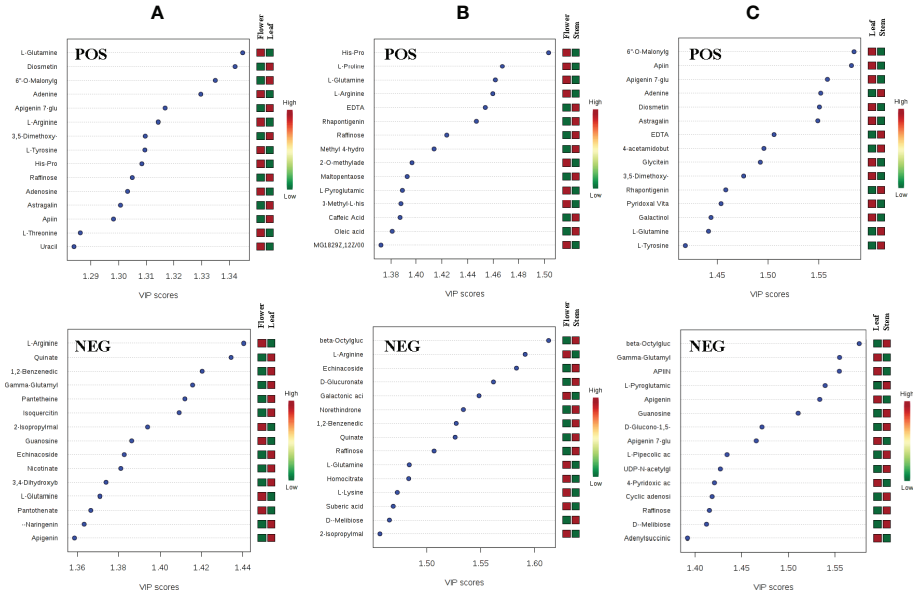


FIGURE 3
The VIP scores of compounds with high principal component ranking in the flower-leaf group **(A)**, flower-stem group **(B)** and leaf-stem group **(C)** of *P. cablin* in positive and negative ion modes. The color intensity represents the relative abundance of metabolites detected.

important in leaf than stem and flower included gamma-glutamylcysteine, guanosine, apigenin, diosmetin, 6"-o-malonylglycitin, adenine, apigenin 7-glucoside, 3,5-dimethoxy-4-hydroxycinnamic acid, L-tyrosine, astragalin, and apiin. Compounds more important in flower than leaf and stem included L-arginine, quinate, 1,2-benzenedicarboxylic acid, 2-isopropylmalic acid, echinacoside, and His-Pro. Moreover, compounds more important in the stem than in the leaves and flowers included beta-octylglucoside, D-melibiose, EDTA, rhapontigenin, and L-pyroglutamic acid (Figure 3).

Univariate analysis (UVA) can objectively reflect the significance of metabolite changes between the two samples. In this study, a P-value less than 0.05 and | Log2 (FC) |>1 were used for screening differential compounds. The results were illustrated in volcano plots (Figure S3). 65 and 61 differential compounds were identified between flower and leaf tissues using the POS and NEG modes (Figure S3A); 48 and 44 differential compounds were identified between flower and stem tissues using the POS and NEG modes (Figure S3B), while 46 and 49 differential compounds were identified between leaf and stem tissues using the POS and NEG modes, respectively (Figure S3C). Subsequently, the differentially abundant metabolites obtained in any two tissues by POS and NEG modes were analyzed by hierarchical clustering (Figure S4). Each shade represents the relative expression level of a certain metabolite in each tissue sample. Overall, significant metabolite differences were observed between leaf, stem and flower groups. By combining the results of the POS and NEG modes, 112, 77 and 83 different metabolites between the flower-leaf (Table S2), flower-stem(Table S3), and leaf-stem (Table S4) groups were identified respectively, indicating that the difference in metabolite abundance between flower and leaf tissues was significantly. Among all metabolites examined, 33 compounds from leaves were found to be different in both the flower and stem groups, while no significant differences were found between flowers and stems. The 33 compounds differentially expressed in leaves can be used as potential metabolite biomarkers of *P. cablin* leaves. Similarly, 32 and 12 potential specific metabolite biomarkers for flowers and stems were found in this study. In addition, 22 common compounds were found in leaf, stem and flower tissues, including raffinose and adenine, which exhibited highly concentrations in the three tissues (Table 1 and Figure S5).

Pathway identification of differential metabolites

To further understand the metabolic pathways and specific biological functions of leaf, stem and flower metabolites in *P. cablin*, Pubchem and the Kyoto encyclopedia of genes and genomes (KEGG) databases were used in this study. Pathway significant enrichment analysis takes KEGG pathway as a unit, and applies hypergeometric test to find out the pathway that is significantly enriched in the differential metabolite compared with the background of the whole metabolome. KEGG, PubChem and other authoritative metabolite databases were used to analyze the differential metabolites and KEGG analysis was conducted using the screening criteria P-value less than 0.05 and | Log2 (FC) |>1. After obtaining the matching information of different metabolites, we searched the pathway database and analyzed the metabolic pathways. 50 metabolic pathways were significantly dysregulated in paired tissue comparisons (Figure 4), and 29 exhibited differences between leaf, stem and flower tissues (Table S5). 7 metabolic pathways were differentially expressed in leaf tissues, including isoquinoline alkaloid biosynthesis, phenylalanine metabolism, one

10.3389/fpls.2023.1098280

TABLE 1 The specific and common metabolite compounds of different tissues detected by UHPLC-QTOFMS in *P. cablin*.

NO.	Flower DEM compounds (32) [a]	Leaf DEM compounds (33) [b]	Stem DEM compounds (12) [c]	Common compounds (22) [d]
1	L-Arginine	Diosmetin	EDTA	L-Glutamine
2	L-Proline	6''-O-Malonylglycitin	Rhapontigenin	Adenine
3	Glyceryl Monolinoleate	Apigenin 7-glucoside	S-Adenosylmethionine	His-Pro
4	2-Amino-2-methyl-1,3-propanediol	3,5-Dimethoxy-4-hydroxycinnamic acid	Histamine	Raffinose
5	1-Methylhistidine	L-Tyrosine	DL-3-Hydroxybutyric acid	Apiin
6	3-Methyl-L-histidine	Adenosine	3-Methylindole	L-Threonine
7	3-Methylhistidine	Astragalin	D-Mannose	Uracil
8	Nicotinate	Glycitein	Xanthosine	Pantothenate
9	Arachidonic Acid peroxide free	6-Hydroxydopamine	L-Pipecolic acid	L-Pyroglutamic acid
10	Cytidine 2',3'-cyclic phosphate	4-Acetamidobutanoate	Galactonic acid	Caffeic Acid
11	1,2-Benzenedicarboxylic acid	Glycerophosphocholine	alpha-Linolenic acid	Maltopentaose
12	L-NG-Monomethylarginine	Riboflavin	D-galacturonic acid	Larixinic Acid
13	DL-Indole-3-lactic acid	Malvidin 3-O-glucoside cation		L-Tryptophan
14	Quercetin 3'-methyl ether	Phosphorylcholine		Allantoate/Allantoic acid
15	Methyl 4-hydroxybenzoate	Galactinol		D–Melibiose
16	S-Methyl-5'-thioadenosine	Lys-Asp		Echinacoside
17	1-Oleoyl-sn-glycero-3-phosphocholine	Stearidonic Acid		3,4-Dihydroxybenzoate Protocatechuic acid
18	Quinate	N-.alpha.-Acetyl-L-arginine		Apigenin
19	Pantetheine	1,2-dihexadecanoyl-sn-glycero-3-phosphocholine		L-Cysteic acid
20	Isoquercitin	Thioetheramide-PC		Adenylsuccinic acid
21	2-Isopropylmalic acid	Abscisic Acid cis,trans		Oleanolic acid
22	1-Palmitoyl-2-linoleoyl-sn-glycero-3-phosphate	Gamma-Glutamylcysteine		beta-Octylglucoside
23	L-Lysine	4-Pyridoxic acid		
24	Homocitrate	UDP-N-acetylglucosamine		
25	cis-6,9,12-Linolenic acid	Adipic acid		
26	Suberic acid	D-Glucono-1,5-lactone		
27	D-Glucuronate	Indapamide		
28	Coumestrol	Acetyl-DL-Valine		
29	cis-Aconitate	Maslinic Acid		
30	4-Acetoxyphenol	ketoisocaproic acid		
31	L-Asparagine	Folinic acid		
32	Cytidine 5'-monophosphate CMP	Phenylpyruvate		
33		11-Keto-.beta.-boswellic acid		

[a]significantly differentially expressed in flower compared with leaf and stem, [b]Significantly differentially expressed in leaf compared to flower and leaf and [c]significantly differentially expressed in stem compared with flower and leaf, [d]Common expressed in three types of tissue samples. DEM stands for differential expression of metabolites.

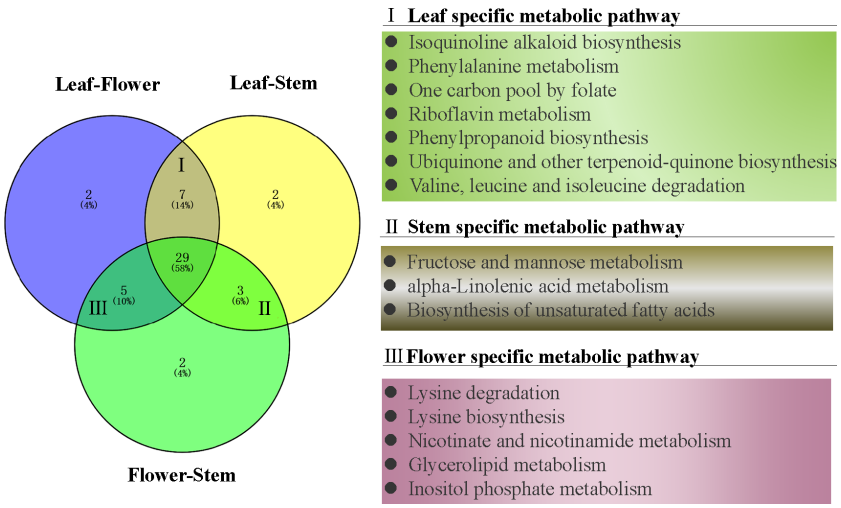


FIGURE 4
The venn diagram of differential metabolic pathways in *P. cablin*. The pathways were referenced in kyoto encyclopedia of genes and genomes (KEGG).

carbon pool by folate, riboflavin metabolism, phenylpropanoid biosynthesis, ubiquinone and other terpenoid-quinone biosynthesis and valine, leucine and isoleucine degradation. Similarly, 3 individual metabolic pathways in the stem, including fructose and mannose metabolism, alpha-linolenic acid metabolism, and biosynthesis of unsaturated fatty acids, were differentially expressed compared with flower or leaf tissues. 5 metabolic pathways were differentially expressed in the flower tissues compared with stem or leaf tissues, including lysine degradation, lysine biosynthesis, nicotinate and nicotinamide metabolism, glycerolipid metabolism and inositol phosphate metabolism (Figure 4).

Integration of pathways and differential metabolites from various tissues

The tissue-specific expression patterns of differentially expressed metabolites identified in the flower-leaf, flower-stem and stem-leaf groups were integrated and mapped onto the most relevant metabolic pathways, including both primary and secondary metabolic pathways (Figures 5-7). A comparison of leaf and flower tissues revealed that 112 metabolites were enriched in 43 pathways. As illustrated in Figure 5, the flower contained higher levels of amino acids (e.g., L-proline and L-lysine), lipids (e.g., PA, as well as cis-6, 9, 12-linolenic acid) and purines (e.g., adenine). In contrast, leaves contained higher levels of terpenoids (e.g., oleanolic acid), vitamins (e.g., folinic and sinapic acid), flavonoids (e.g., apigenin and naringenin), most carbohydrates (e.g., raffinose and D-melibiose) and organic acids (e.g., diosmetin). 77 differentially expressed metabolites were identified between stem and flower, and these metabolites were enriched in 39 metabolic pathways. As shown in Figure 6, the flower contained higher levels of amino acids (e.g., L-proline and L-lysine) and lipids (e.g., PA as well as cis-6, 9, 12-linolenic acid), stems contained higher levels of purines (e.g., adenine), flavonoids (e.g., apigenin and naringenin), and most carbohydrates (e.g., raffinose and D-melibiose). In addition, 83 differentially expressed metabolites between stem and leaf were enriched in 41 pathways. As shown in Figure 7, the leaf contained higher levels of terpenoids (e.g., oleanolic acid and maslinic acid), dipeptide and peptide amino acids (e.g., lysyl-aspartate and cysteic acid), vitamins (e.g., folinic acid and riboflavin) and flavonoids (e.g., apigenin and astragalin), the stem contained higher levels of purines (e.g., adenine), amino acids (e.g., L-isoleucine and L-glutamine), lipids (e.g., adenine) and carbohydrates (e.g., maltol and D-melibiose).

In conclusion, GC-MS and UHPLC-QTOFMS analysis revealed the accumulation and enrichment pathways of different metabolites among leaves, stems and flowers. Our findings suggest that the flowers accumulate more lipids and amino acids, including proline, lysine, arginine, asparagine, threonine, and tryptophan in *P. cablin*. Meanwhile, higher levels of terpenoids, vitamins and flavonoids accumulated in leaves, while higher levels of carbohydrates were found in stems. The oxidative decarboxylation of pyruvate and the catabolism of amino acids (such as lysine and tryptophan) produce acetyl-CoA, which is also a precursor involved in the biosynthesis of terpenoids, flavonoids and carbohydrates. The above results provide compelling evidence of the exchange of metabolites among leaf, stem and flower tissues in *P. cablin* and indicate a potential complex metabolic network in *P. cablin* (Figure 8).

Discussion

P. cablin is widely cultivated in many countries, including China, India, Malaysia, Indonesia, etc. It is rich in patchouli oil and represents the only natural source of essential oil. Moreover, *P. cablin* is one of the 20 essential oil plants most commonly traded in the international market, with huge market demand. Given the absence of a synthetic substitute, patchouli oil exhibits an increasingly important market value (Swamy et al., 2015). In addition to volatile substances such as essential oil, *P. cablin* is also rich in non-volatile compounds, such as flavonoids (Li et al.,

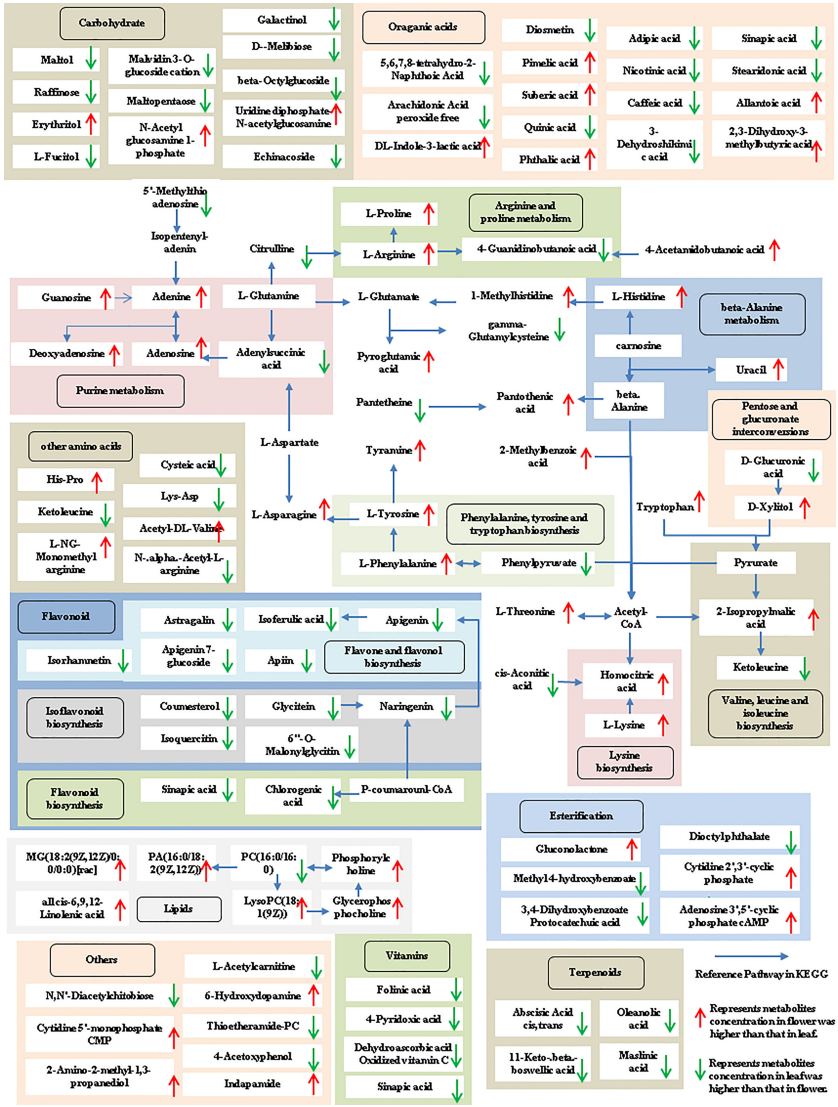


FIGURE 5
Metabolic pathways analysis in flower-leaf group of *P. cablin*. The 112 metabolites identified in flower-leaf group were mapped onto primary and secondary metabolism. The upward-pointing red arrows represent the value of |log2| (content flower/leaf) >1, which means the level of metabolites in flower is higher than that in leaf, the value of |log2|(content flower/leaf) <1 means the level of metabolites in leaf was higher than that in flower, and represented by downward-pointing green arrows.

2011). In recent years, most emphasis has been placed on the pharmacological activity and biosynthetic mechanism of metabolites in *P. cablin*, and there is no in-depth discussion on the distribution and exchange of compounds in key tissues of *P. cablin*, such as leaves, stems and flowers. Our research systematically analyzed the distribution and types of volatile and non-volatile substances in the leaves, stems and flowers of *P. cablin* and the metabolic pathways involved in them through GC-MS and UHPLC-QTOFMS. At the same time, the exchange of metabolites between different tissues of *P. cablin* was analyzed, providing a reference for future studies on the metabolic network of *P. cablin*.

Although the metabolites in the leaves and stems of *P. cablin* have been extensively reported, metabolites in the flowers of *P. cablin* have been understudied since they rarely bloom. Fortunately, in the present study, enough flowers were collected in spring. To conduct an integrated analysis of the differences in metabolite abundance in the three tissues, GC-MS was selected to analyze the volatile compounds. The results revealed that the leaf, stem and flower tissues of *P. cablin* were rich in various volatile components, and the leaves contained more volatile substances than the stems and flowers. It was found that patchouli alcohol is the main compound in the three tissues. The quantitative results showed that patchouli alcohol exhibited the highest levels in leaves (24.89 mg/g), 9.12 and 6.69-fold higher than in stems and flowers. It has been reported that glandular hair is the main site of volatile oil accumulation, and a large amount of patchouli alcohol in patchouli leaves is related to its rich epidermal hair (Guo et al., 2013), which warrants further study. In addition, as the representative of sesquiterpenes in *P. cablin*, patchouli alcohol is mainly generated through the MVA and MEP pathways, and the original substrate of

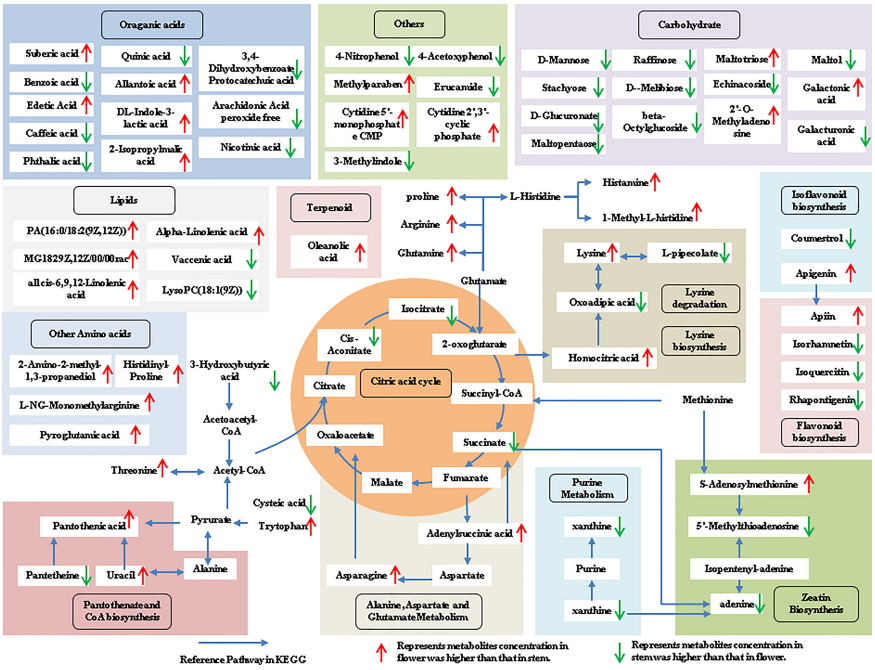


FIGURE 6
Metabolic pathways analysis in flower-stem group of *P. cablin*. The 77 metabolites identified in flower-stem group were mapped onto primary and secondary metabolism. The upward-pointing red arrows represent the value of |log2| (content flower/stem) >1, which means the level of metabolites in flower is higher than that in stem, the value of |log2|(content flower/stem) <1 represents metabolites concentration in stem was higher than that in flower, and represented by downward-pointing green arrows.

the MVA pathway is acetyl coenzyme A (Li et al., 2021). We have reason to believe that acetyl coenzyme A exists in three tissues and participates in the biosynthesis of volatile substances of patchouli, such as terpenoids.

In addition to volatile compounds, a variety of non-volatile compounds is present in *P. cablin*. It has been reported that the dried aerial parts of *P. cablin* contain 33 flavonoids, 21 organic acids and 9 phenylpropanoids, ect (Xie et al., 2022), but they are limited to the analysis of mixed tissues. UHPLC-QTOFMS can comprehensively predict the main active components in the leaves, stems and flowers of *P. cablin* under positive and negative ion scanning modes. PCA analysis between the flower-leaf, flower-stem and leaf-stem groups of *P. cablin* was conducted to obtain the scores of PC1 and PC2 in the positive and negative ion mode, respectively. Each tissue can be divided into different regions independently. It was found that the differences between groups were significant in the top-ranked components, indicating that there are significant differences in the main metabolites between the flower-leaf, flower-stem and leaf-stem. The Univariate analysis further identified the different metabolites among groups, including 33 compounds in leaves and 32 in flowers. However, there were fewer individual compounds in stems (n=12). These compound molecules can be used as biomarkers for three tissues of *P. cablin* for germplasm identification and quality evaluation, which may be attributed to the fact that *P. cablin* is dependent on asexual reproduction, resulting in a single genetic background.

KEGG analysis revealed flower-specific pathways, including lysine degradation, lysine biosynthesis, nicotinate and nicotinamide metabolism, glycerolipid metabolism and inositol phosphate metabolism. It was found that flowers contained high levels of lipids and amino acids, such as proline and lysine. However, there were no differences in proline abundance between leaves and stems. A study reported that the proline level in tomato flowers was 60 times higher than that in other tissues (Schwacke et al., 1999), consistent with our findings, indicating that the accumulation of proline is related to the growth and development of the reproductive organs of *P. cablin*. Indeed, as the reproductive organs of plants, flowers contain more lipids than other organs (Wang et al., 2016). Higher lipid levels in *P. cablin* indicate a certain resistance to environmental stress. In addition, pogostol, neophytadiene, and phytol were detected only in flower tissues and may be associated with metabolic pathways specific to flowers. Neophytadiene and phytol are diterpenoids, and little research has been done on their biosynthesis. In our previous study, OEL4 and OEL12 transgenic tobacco plants, which overexpressed *PcFPPS*, exhibited significantly higher levels of phytol and neophytadiene than wild-type lines (Wang et al., 2022). Based on the synthetic pathways involved in terpenoid biosynthesis, diterpenoid biosynthesis is widely thought to be related to MEP synthesis. Studies have found that L-lysine is involved in carotenoid biosynthesis (Henke et al., 2018), however, carotenoids, chlorophyll, phytol, tocopherol and prenylated hydrazine share a common precursor produced by the plastid MEP pathway. We hypothesize that lysine can form precursors of the MEP pathway to synthesize some tetraterpenoids and sterols,such as carotenoids and phytol.

The results of GC-MS revealed that leaves contain more sesquiterpenoids compared with flowers or stems, β-elemene,

10.3389/fpls.2023.1098280

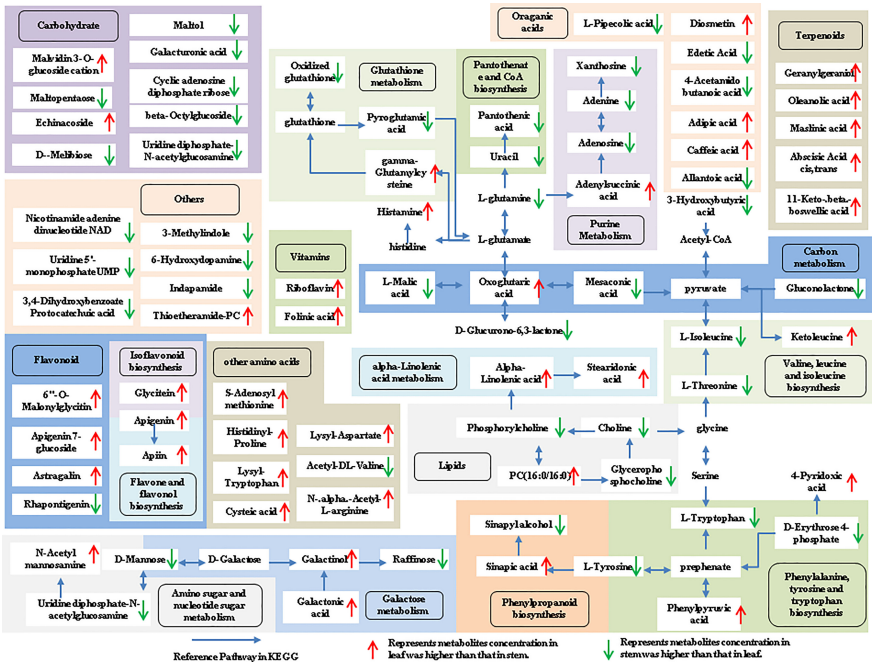


FIGURE 7
Metabolic pathways analysis in leaf-stem group of *P. cablin*. The 83 metabolites identified in leaf-stem group were mapped onto primary and secondary metabolism. The upward-pointing red arrows represent the value of |log2| (content leaf/stem) >1, which means the level of metabolites in leaf is higher than that in stem, the value of |log2|(content leaf/stem) <1 represents metabolites concentration in stem was higher than that in leaf, and represented by downward-pointing green arrows.

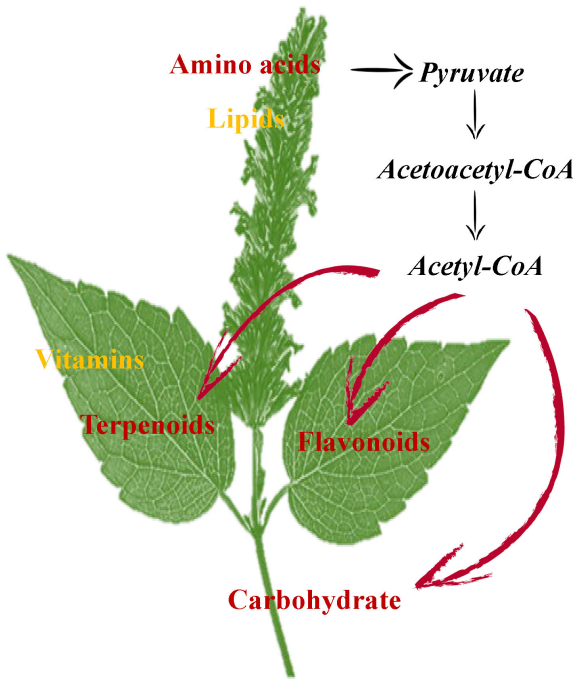


FIGURE 8
The possible mechanisms for the synthesis and accumulation of compounds in the leaves, stems and flowers of *P. cablin*. Flowers contain high levels of amino acids and lipids. Leaves contain high levels of terpenoids and flavonoids. Stems contain high levels of carbohydrates.

humulene, (+)spathulenol, valencene, alpha-Gurjunene and aciphyllene were found exclusively in leaves. It is well known that sesquiterpenes are mainly synthesized from acetyl coenzyme A *via* the MVA pathway (Wang et al., 2017), and acetyl coenzyme A is crucial for many metabolic and decomposition pathways, such as fatty acid biosynthesis(Tang et al., 2018). Moreover, Acetyl coenzyme A is a key precursor of the triterpene squalene. The accumulation of acetyl coenzyme A can be promoted by increasing the synthesis of pyruvate, which can effectively increase the content of squalene (Huang et al., 2018). Overwhelming evidence substantiates that the biosynthesis of acetyl-CoA derivatives can be enhanced by providing sufficient precursors (Meadows et al., 2016). Therefore, acetyl coenzyme A may be directly used to form terpenoids, thereby reducing the formation of ketone bodies, leading to the accumulation of terpenoids and the reduction of lipid content.

KEGG analysis revealed that metabolite content related to the citrate cycle, glyox-ylate and dicarboxylate metabolism, and galactose metabolism significantly differed among tissue types. As the main pathway of sugar aerobic oxidation, the tricarboxylic acid cycle of plants is a main energy source for plants. The key regulatory enzyme of the TCA cycle is citrate synthase (CS), which regulates energy generation during mitochondrial respiration by catalyzing the reaction between oxaloacetic acid (OAA) and acetyl coenzyme A (Ac-CoA) to generate citrate and CoA (Li et al., 2016). It is well-established that stem tissues maintain high levels of carbohydrates through various synthetic pathways. Stems are formed by tissues responsible for transportation, storage, and support, and

carbohydrates play important roles in each of these functions. Cellulose is a key component of the cell wall, which supports the elongation of the stem and is essential for plant anisotropic growth. (McFarlane et al., 2014). Moreover, maintaining high levels of carbohydrates in the stem represents another strategy to resist environmental stress. In a study on plant responses to desiccation, stem tissue accumulated sucrose and starch, unlike leaf tissue. Leaves died and were abscised to prevent water loss due to transpiration; however, stem photosynthesis was still observed, and stem tissues entered a dormant state until the period of water restriction ended (Liu et al., 2007). Overall, the present study's findings emphasize the differential metabolites in the leaves, stems and flowers of *P. cablin* and the metabolic pathways involved, providing novel insights to explain the differences in metabolic components caused by tissue differences. Our findings are expected to promote research on the metabolic network of *P. cablin*, the diversified application of *P. cablin* and the innovation, sustainable development and utilization of *P. cablin* germplasm resources.

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

LC, LZ, and RZ conceived and designed the experiments. LZ, DW, JZ, DZ, HZ, and XBW performed the experiments. LZ, XZ, LG, XMW, LC, DW, and HZ analyzed the data. LC, LZ, and XBW wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by Key-Area Research and Development Program of Guangdong Province, China (No. 2020B020221001), Project of south medicine innovation team in modern agricultural industry technology system of Guangdong Province (2022KJ148), Science and Technology Program of Guangzhou, China (No. 202102021279), and Universities Provincial-level Platform and Scientific Research Projects of Guangdong universities (No. 2020KQNCX173).

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.2023.1098280/full#supplementary-material

References

Ahmed, S., Zhan, C., Yang, Y., Wang, X., Yang, T., Zhao, Z., et al. (2016). The transcript profile of a traditional Chinese medicine, atractylodes lancea, revealing its sesquiterpenoid biosynthesis of the major active components. *PloS One* 11 (3), e0151975. doi: 10.1371/journal.pone.0151975

Cavanna, D., Catellani, D., Dall'Asta, C., and Suman, M. (2018). Egg product freshness evaluation: A metabolomic approach. *J. Mass Spectrom* 53 (9), 849–861. doi: 10.1002/jms.4256

Chatt, E. C., von Aderkas, P., Carter, C. J., Smith, D., Elliott, M., and Nikolau, B. J. (2018). Sex-dependent variation of pumpkin (Cucurbita maxima cv. big max) nectar and nectaries as determined by proteomics and metabolomics. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00860

Chen, X., Li, J., Wang, X., Zhong, L., Tang, Y., Zhou, X., et al. (2019). Full-length transcriptome sequencing and methyl jasmonate-induced expression profile analysis of genes related to patchoulol biosynthesis and regulation in pogostemon cablin. *BMC Plant Biol.* 19 (1), 266. doi: 10.1186/s12870-019-1884-x

Farag, M. A., El-Ahmady, S. H., Elian, F. S., and Wessjohann, L. A. (2013). Metabolomics driven analysis of artichoke leaf and its commercial products *via* UHPLC-q-TOF-MS and chemometrics. *Phytochemistry* 95, 177–187. doi: 10.1016/j.phytochem.2013.07.003

Firdaus, S., van Heusden, A. W., Hidayati, N., Supena, E. D., Mumm, R., de Vos, R. C., et al. (2013). Identification and QTL mapping of whitefly resistance components in solanum galapagense. *Theor. Appl. Genet.* 126 (6), 1487–1501. doi: 10.1007/s00122-013-2067-z

Guo, J., Yuan, Y., Liu, Z., and Zhu, J. (2013). Development and structure of internal glands and external glandular trichomes in pogostemon cablin. *PloS One* 8 (10), e77862. doi: 10.1371/journal.pone.0077862

Henke, N. A., Wiebe, D., Perez-Garcia, F., Peters-Wendisch, P., and Wendisch, V. F. (2018). Coproduction of cell-bound and secreted value-added compounds: Simultaneous production of carotenoids and amino acids by corynebacterium glutamicum. *Bioresour. Technol.* 247, 744–752. doi: 10.1016/j.biortech.2017.09.167

Henry, L. K., Thomas, S. T., Widhalm, J. R., Lynch, J. H., Davis, T. C., Kessler, S. A., et al. (2018). Contribution of isopentenyl phosphate to plant terpenoid metabolism. *Nat. Plants* 4 (9), 721–729. doi: 10.1038/s41477-018-0220-z

Hu, L. F., Li, S. P., Cao, H., Liu, J. J., Gao, J. L., Yang, F. Q., et al. (2006). GC-MS fingerprint of pogostemon cablin in China. *J. Pharm. BioMed. Anal.* 42 (2), 200–206. doi: 10.1016/j.jpba.2005.09.015

Huang, Y. Y., Jian, X. X., Lv, Y. B., Nian, K. Q., Gao, Q., Chen, J., et al. (2018). Enhanced squalene biosynthesis in yarrowia lipolytica based on metabolically engineered acetyl-CoA metabolism. *J. Biotechnol.* 281, 106–114. doi: 10.1016/j.jbiotec.2018.07.001

Lavergne, F. D., Broeckling, C. D., Cockrell, D. M., Haley, S. D., Peairs, F. B., Jahn, C. E., et al. (2018). GC-MS metabolomics to evaluate the composition of plant cuticular waxes for four triticum aestivum cultivars. *Int. J. Mol. Sci.* 19 (2), 249. doi: 10.3390/ijms19020249

Li, Y., Wang, H., Dai, F., Li, P., Jin, X., Huang, Y., et al. (2016). Label-free fluorescent enzymatic assay of citrate synthase by CoA-Au(I) co-ordination polymer and its

application in a multi-enzyme logic gate cascade. *Biosens Bioelectron.* 86, 1038–1046. doi: 10.1016/j.bios.2016.07.107

Li, Z.-J., Wang, Y.-Z., Wang, L.-R., Shi, T.-Q., Sun, X.-M., and Huang, H. (2021). Advanced strategies for the synthesis of terpenoids in yarrowia lipolytica. *J. Agric. Food Chem.* 69 (8), 2367–2381. doi: 10.1021/acs.jafc.1c00350

Li, K., Zhang, H., Xie, H., Liang, Y., Wang, X., and Ito, Y. (2011). Preparative isolation and purification of five flavonoids from pogostemon cablin benth by high-speed countercurrent chromatography and preparative high-performance liquid chromatography. *J. Liq Chromatogr. Relat. Technol.* 34 (15), 1617–1629. doi: 10.1080/10826076.2011.580486

Liu, F., Deng, C., Cao, W., Zeng, G., Deng, X., and Zhou, Y. (2017). Phytochemicals of pogostemon cablin (Blanco) benth. aqueous extract: Their xanthine oxidase inhibitory activities. *BioMed. Pharmacother.* 89, 544–548. doi: 10.1016/j.biopha.2017.01.040

Liu, Y., Zhang, T., Li, X., and Wang, G. (2007). Protective mechanism of desiccation tolerance in reaumuria soongorica: Leaf abscission and sucrose accumulation in the stem. *Sci. China C Life Sci.* 50 (1), 15–21. doi: 10.1007/s11427-007-0002-8

McFarlane, H. E., Doring, A., and Persson, S. (2014). The cell biology of cellulose synthesis. *Annu. Rev. Plant Biol.* 65, 69–94. doi: 10.1146/annurev-arplant-050213-040240

Meadows, A. L., Hawkins, K. M., Tsegaye, Y., Antipov, E., Kim, Y., Raetz, L., et al. (2016). Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature* 537 (7622), 694–697. doi: 10.1038/nature19769

Miyazawa, M., Okuno, Y., Nakamura, S.-I., and Kosaka, H. (2000). Antimutagenic activity of flavonoids from pogostemon cablin. *J. Agric. Food Chem.* 48, 642–647. doi: 10.1021/jf990160y

Patterson, A. D., Gonzalez, F. J., and Idle, J. R. (2010). Xenobiotic metabolism: A view through the metabolometer. *Chem. Res. Toxicol.* 23, 851–860. doi: 10.1021/tx100020p

Qian, Y.-J., Pi, W.-X., Zhu, G.-F., Wei, W., Lu, T.-L., and Mao, C.-Q. (2022). Quality evaluation of raw and processed corni fructus by UHPLC-QTOF-MS and HPLC coupled with color determination. *J. Pharm. Biomed. Anal.* 218, 114842. doi: 10.1016/j.jpba.2022.114842

Sandes, S. S., Zucchi, M. I., Pinheiro, J. B., Bajay, M. M., Batista, C. E., Brito, F. A., et al. (2016). Molecular characterization of patchouli (Pogostemon spp) germplasm. *Genet. Mol. Res.* 15 (1). doi: 10.4238/gmr.15017458

Schwacke, R., Grallath, S., E.Breitkreuz, K., Stransky, E., Stransky, H., Frommer, W. B., et al. (1999). LeProT1, a transporter for proline, glycine betaine, and gamma -amino butyric acid in tomato pollen. *Plant Cell* 11, 377–391. doi: 10.1105/tpc.11.3.377

Schwarz, P. N., Roller, L., Kulik, A., Wohlleben, W., and Stegmann, E. (2018). Engineering metabolic pathways in amycolatopsis japonicum for the optimization of the precursor supply for heterologous brasilicardin congeners production. *Synth Syst. Biotechnol.* 3 (1), 56–63. doi: 10.1016/j.synbio.2017.12.005

Swamy, M. K., Mohanty, S. K., Sinniah, U. R., and Maniyam, A. (2015). Evaluation of patchouli (Pogostemon cablinBenth.) cultivars for growth, yield and quality parameters. *J. Essential Oil Bearing Plants* 18 (4), 826–832. doi: 10.1080/0972060x.2015.1029989

Swamy, M. K., and Sinniah, U. R. (2015). A comprehensive review on the phytochemical constituents and pharmacological activities of pogostemon cablin benth.: An aromatic medicinal plant of industrial importance. *Molecules* 20 (5), 8521–8547. doi: 10.3390/molecules20058521

Tamura, Y., Mori, T., Nakabayashi, R., Kobayashi, M., Saito, K., Okazaki, S., et al. (2018). Metabolomic evaluation of the quality of leaf lettuce grown in practical plant factory to capture metabolite signature. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00665

Tang, Y., Zhou, J., Hooi, S. C., Jiang, Y. M., and Lu, G. D. (2018). Fatty acid activation in carcinogenesis and cancer development: Essential roles of long-chain acyl-CoA synthetases. *Oncol. Lett.* 16 (2), 1390–1396. doi: 10.3892/ol.2018.8843

Wang, Y., Ma, X., Zhang, X., He, X., Li, H., Cui, D., et al. (2016). ITRAQ-based proteomic analysis of the metabolic mechanisms behind lipid accumulation and degradation during peanut seed development and postgermination. *J. Proteome Res.* 15 (12), 4277–4289. doi: 10.1021/acs.jproteome.6b00345

Wang, X., Tang, Y., Huang, H., Wu, D., Chen, X., Li, J., et al. (2022). Functional analysis of pogostemon cablin farnesyl pyrophosphate synthase gene and its binding transcription factor PcWRKY44 in regulating biosynthesis of patchouli alcohol. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.946629

Wang, M., Wang, D., Zhang, Q., Chai, J., Peng, Y., and Cai, X. (2017). Identification and cytochemical immunolocalization of acetyl-CoA acetyltransferase involved in the terpenoid mevalonate pathway in euphorbia helioscopia laticifers. *Bot. Stud.* 58 (1), 62. doi: 10.1186/s40529-017-0217-3

Xie, B., Wu, X.-F., Luo, H.-T., Huang, X.-L., Huang, F., Zhang, Q.-Y., et al. (2022). Chemical profiling and quality evaluation of pogostemon cablin benth by liquid chromatography tandem mass spectrometry combined with multivariate statistical analysis. *J. Pharm. Biomed. Anal.* 209, 114526. doi: 10.1016/j.jpba.2021.114526

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 01 March 2023
DOI 10.3389/fpls.2023.1128001

OPEN ACCESS

EDITED BY
Weiwei Zhang,
Yangtze University, China

REVIEWED BY
Sofia D. Carvalho,
Independent researcher, Laramie, WY, United States
Xiaohua Li,
Hubei University of Chinese Medicine, China

*CORRESPONDENCE
Seok Hyun Eom
se43@khu.ac.kr

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

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 20 December 2022
ACCEPTED 16 February 2023
PUBLISHED 01 March 2023

CITATION
Lim YJ, Kwon S-J and Eom SH (2023) Red and blue light-specific metabolic changes in soybean seedlings.
Front. Plant Sci. 14:1128001.
doi: 10.3389/fpls.2023.1128001

Red and blue light-specific metabolic changes in soybean seedlings

You Jin Lim[1†], Soon-Jae Kwon[2†] and Seok Hyun Eom[1*]

[1]Department of Smart Farm Science, College of Life Sciences, Kyung Hee University, Yongin, Republic of Korea, [2]Advanced Radiation Technology Institute, Korea Atomic Energy Research Institute, Jeongeup, Republic of Korea

Red and blue artificial light sources are commonly used as photosynthetic lighting in smart farm facilities, and they can affect the metabolisms of various primary and secondary metabolites. Although the soybean plant contains major flavonoids such as isoflavone and flavonol, using light factors to produce specific flavonoids from this plant remains difficult because the regulation of light-responded flavonoids is poorly understood. In this study, metabolic profiling of soybean seedlings in response to red and blue lights was evaluated, and the isoflavone–flavonol regulatory mechanism under different light irradiation periods was elucidated. Profiling of metabolites, including flavonoids, phenolic acids, amino acids, organic acids, free sugars, alcohol sugars, and sugar acids, revealed that specific flavonol, isoflavone, and phenolic acid showed irradiation time-dependent accumulation. Therefore, the metabolic gene expression level and accumulation of isoflavone and flavonol were further investigated. The light irradiation period regulated kaempferol glycoside, the predominant flavonol in soybeans, with longer light irradiation resulting in higher kaempferol glycoside content, regardless of photosynthetic lights. Notably, blue light stimulated kaempferol-3-*O*-(2,6-dirhamnosyl)-galactoside accumulation more than red light. Meanwhile, isoflavones were controlled differently based on isoflavone types. Malonyl daidzin and malonyl genistin, the predominant isoflavones in soybeans, were significantly increased by short-term red light irradiation (12 and 36 h) with higher expressions of flavonoid biosynthetic genes, which contributed to the increased total isoflavone level. Although most isoflavones increased in response to red and blue lights, daidzein increased in response only to red light. In addition, prolonged red light irradiation downregulated the accumulation of glycitin types, suggesting that isoflavone's structural specificity results in different accumulation in response to light. Overall, these findings suggest that the application of specific wavelength and irradiation periods of light factors enables the regulation and acquisition of specialized metabolites from soybean seedlings.

KEYWORDS

isoflavone, flavonol, photosynthetic light, metabolic regulation, irradiation period

1 Introduction

Soybean is an economically important crop for oil and protein production and a predominant source of isoflavones in dietary supplements. In East Asian countries such as Korea and Japan, soybean sprouts are commonly used as culinary foods due to their short-term production and high nutritional value. During seed germination, many biochemical changes occur, resulting in the accumulation of various primary and secondary metabolites (Shu et al., 2008; Na Jom et al., 2011; Ma et al., 2022). Flavonoids, particularly isoflavones, which represent the health-promoting properties of soybeans, are synthesized throughout the whole soybean plant (Vacek et al., 2008; Sugiyama et al., 2017; Křížová et al., 2019). Isoflavone is a class of flavonoids that is abundant in soybean seeds and has a protective effect against breast cancer, prostate cancer, cardiovascular disease, and osteoporosis (Anthony et al., 1998; Arjmandi and Smith, 2002; Wei et al., 2012; Křížová et al., 2019). In soybean leaves, kaempferol glycosides are present as major flavonols (Ho et al., 2002; Kim et al., 2008). Kaempferol and its glycosides have health-promoting effects, such as antioxidant, antinociceptive, and anti-inflammatory properties (De Melo et al., 2009; Wang et al., 2018). Despite numerous studies on flavonoids, including isoflavones in soybeans, the regulation of specific flavonoid metabolism is still required for application on artificial controlled agricultural system (Shah and Smith, 2020; Sohn et al., 2021; Wang et al., 2022).

Recent research has focused on the application of artificial light to plants in controlled agricultural systems to produce high levels of phytochemicals (Bian et al., 2015; Zhang et al., 2020). Light is an important abiotic factor that directly or indirectly influences flavonoid biosynthesis in plants (Thoma et al., 2020; Naik et al., 2022). Red and blue artificial light sources are commonly used as photosynthetic lights in smart farm facilities, as they stimulate the production of phytochemicals and the consequent accumulation of various primary and secondary metabolites (Thoma et al., 2020). Red and blue light are the most effectively absorbed light spectra by photosynthetic pigments (Li et al., 2020). While the application of blue light and ultraviolet light to germinating soybeans promotes isoflavone accumulation, it also induces the accumulation of unsaturated fatty acids and free amino acids with a reduced sugar content (Azad et al., 2018; Ma et al., 2018; Ma et al., 2020; Lim et al., 2021). Although short-wavelength light, such as blue light and ultraviolet light, has been effectively used to increase isoflavone levels in soybean plants, red/far-red-mediated phytochrome response has also been reported to contribute to isoflavone level regulation (Kirakosyan et al., 2006).

The relationship between the light environment and flavonoid accumulation has been widely studied (Zoratti et al., 2014; Bian et al., 2015; Fu et al., 2016). Nevertheless, as the use of artificial light increases, studies on metabolic changes and acquisition of target metabolites under light-controlled systems are still required. The relationship between each metabolic pathway in a plant is intricate and strongly influenced by the external environment (Di Ferdinando et al., 2012). Reports on the regulation of flavonoid metabolic steps in response to light are highly dispersed (Neugart et al., 2021; Naik et al., 2022), making it difficult to apply light factors that can control specific metabolites in a certain crop. Moreover, light-induced metabolic changes are highly variable based on crop species and metabolite types (Bian et al., 2015; Thoma et al., 2020). In smart farms and plant factory facilities, the regulation of specialized metabolites, notably secondary metabolites and functional compounds, is essential for the production of high-value-added crops.

A metabolomic approach, including metabolite profiling and specific metabolic gene regulation, is considered a promising strategy for understanding the metabolic regulation by photosynthetic light (Kusano et al., 2011; Li et al., 2019). Although the soybean plant contains both isoflavone and flavonol as major flavonoids (Lim et al., 2021), using light factors to produce a specific flavonoid from this plant remains challenging because the regulation of light-responded isoflavone and flavonol remains unclear. Therefore, the aims of the present study are to evaluate the changes in metabolites in soybean seedlings in response to red and blue photosynthetic lights and to clarify the regulation of isoflavone and flavonol accumulation by different light irradiation periods.

2 Materials and methods

2.1 Cultivation of soybean seedlings under light treatment

The soybean seeds (*Glycine max* cv. Pungwon) were provided by Pulmuone Food Co. (Chungbuk, South Korea). The soybean seeds were soaked for 4 h with distilled water. The soaked seeds were grown in a 103.0 (diameter) x 78.6 mm (height) plant culture dish (SPL, Pocheon, Korea) with holes of 5 mm in diameter in the bottom under hydroponic culture supplying daily exchanged distilled water. The plant culture dishes were placed in a growth chamber with 23 ± 2°C. The soaked seeds were grown for 7 days in the dark growth chamber for the control group. In order not to affect seed germination by light, the light was irradiated for a total of 5 days after radicle emergence of initial 2 days. For evaluating the effect of LED red (650 nm) and blue (447 nm) light irradiation, 2-day-old soybean seedlings in darkness were transferred to either red or blue light and harvested after 5 d cultivation defined to long-term irradiation in this experiment. On the other hand, 5.5 and 6.5-day-old seedlings in darkness were transferred to the lights and harvested after 36 and 12 h cultivation, respectively, defined to short-term irradiation.

The light sources used in this experiment were red (650 nm) and blue (447 nm) LEDs (D9RBN10SC, Plant Husbandry, Suwon, Korea) (Supplementary Figure 1). The intensity of the red and blue LEDs was set to 50 μmol m^{-2} s^{-1} using a photo-radio meter (HD 2302.0; Delta OHM SRL, Marconi, Italy). Each treatment was performed in three replications, and 40 seeds were planted per replicate. The sprouts grown for 7 days were immediately put in liquid N_2 and stored at −80°C after harvest until the analysis of metabolic gene expression. For the analysis of chlorophyll, GC-MS, and HPLC, the harvested samples were lyophilized at −80°C using a freeze dryer (IlshinBioBase Co. Ltd., Dongducheon, Korea).

2.2 Measurement of seedling growth and total chlorophyll content

Sprout growth was measured by shoot and root length and dry weight. The dry weight was measured after the drying process using a freeze dryer (Ilshin Lab. Co., Yangju, South Korea). For chlorophyll analysis, the ground cotyledon of dried seedlings (10 mg) was immersed in 1 mL of 80% aqueous acetone (*v/v*). Chlorophyll was extracted in a shaking incubator for 24 h at 30°C. The extract was centrifuged at 12,000 rpm for 5 min. The absorbance of the supernatant was measured at 663 and 645 nm using a spectrophotometer (S-4100; SCINCO Co., Ltd., Seoul, South Korea). The total chlorophyll content was calculated using the following formula: total chlorophyll (mg g^{-1} DW) = ([8.02 × OD_{663}] + [20.2 × OD_{645}])/10 (where DW and OD indicate dry weight and optical density, respectively) (Mackinney, 1941).

2.3 Determination of metabolites by gas chromatography-mass spectrometry

Untargeted metabolites were analyzed by GC-MS. The samples were derivatized before GC-MS analysis. Derivatization and GC-MS processes were performed using previously published methods, with modifications (Gu et al., 2017; Park et al., 2022). Five milligrams of freeze-dried samples were mixed with 200 µL of 20,000 ppm methyl hydroxyl chloride amine in pyridine solution and treated with sonication for 10 min. The mixture was incubated at 30°C for 90 min for an oxygenation procedure. Fifty microliters of the oximated samples were mixed with 50 µL of a mixture of N,O-bis (trimethylsilyl)trifluoroacetamide and 1% trimethylchlorosilane solution for trimethylsilylation. As an internal standard, 10 µL of 500 ppm fluoranthene was added, and the mixture was vortex mixed and heated for 30 min at 60°C.

The separation of chemical compounds in the derivatized samples, as well as their identification, was performed using gas chromatography and analyzed using a mass selective detector (Agilent Technologies, Palo Alto, CA, USA) operating in selected ion monitoring (SIM) mode. GC (7890 B series, Agilent Technologies) was performed by connecting to a 5977 B MS (Agilent Technologies) equipped with a VF-5MS column (60 m length, 0.25 mm i.d., 0.25 µm film thickness, Agilent Technologies). Helium was the carrier gas at a flow rate of 1.5 mL min^{-1}. A split injection mode was executed at a ratio of 20:1 at 300°C. The GC oven temperature was held at 50°C for 2 min, raised to 180°C at a rate of 5°C min^{-1}, and then held at the temperature for 8 min. After that, the temperature was increased to 210°C at a rate of 2.5°C min^{-1} and to 320°C at a rate of 5°C min^{-1}. The final temperature was held for 10 min. Mass spectra were obtained at 70 eV through electron ionization. Data were acquired in scan mode. The peaks in the chromatogram were identified on the basis of their mass spectra selected using the database of NIST 17 library (http://www.nist.gov/srd/nist1.htm). Each compound was quantified using the internal standard ratio. The quality of prediction was assessed through match factor and reverse match factor with above 800 score, and probability (%) of the library hits. In addition, the retention index values supported the predicted identity. Triplicate analysis was performed.

2.4 Determination of flavonoids by HPLC

Extraction and HPLC analyses of isoflavones and kaempferol glycosides were performed using previously published methods (Lim et al., 2021). The standards of eight isoflavone—two aglycones (daidzein and genistein; LC Laboratories, Woburn, MA, USA), three β-glycosides (daidzin, glycitin, genistin; LC Laboratories), and three malonyl glycosides (malonyl daidzin, malonyl glycitin, malonyl genistin; GenDEPOT, Katy, Texas, USA)—were used with a serial concentration (2.5, 5, 10, and 20 mg/L) to calculate the standard curve for quantifying isoflavones in HPLC analysis based on an external standard method. Kaempferol glycosides were quantified with robinin (kaempferol-3-*O*-robinoside-7-*O*-rhanmoside), a kaempferol triglycoside with λmax 265.0 and 347.1 nm. The robinin standard isolated from kudzu and purified as described in Eom et al. (2018) was used with a serial concentration (1.25, 2.5, 5, and 10 mg/L) to calculate the standard curve for quantifying kaempferol glycosides in this experiment.

2.5 Transcript level analysis of genes involved in flavonoid biosynthesis by quantitative RT-PCR

Genes characterized in this experiment included PHENYLALANINE AMMONIA-LYASE 1 (*GmPAL1*), CINNAMATE-4-HYDROXYLASE (*GmC4H*), 4-COUMARATE : COENZYME A LIGASE (*Gm4CL*), CHALCONE SYNTHASE 1 to 8 (*GmCHS1–8*), CHALCONE REDUCTASE (*GmCHR*), CHALCONE ISOMERASE 1A TYPE II AND 1B1 TYPE II (*GmCHI1AII* and *1B1II*), ISOFLAVONE SYNTHASE 1 and 2 (*GmIFS1* and *GmIFS2*), FLAVANONE 3-HYDROXYLASE (*GmF3H*), and ISOFLAVONE REDUCTASE (*GmIFR*). The gene-specific primers used for qRT-PCR were obtained from Primer-BLAST in the National Center for Biotechnology Information databases (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and are listed in Supplementary Table 1. Total RNA was isolated from 7-day-old soybean seedlings using a Qiagen RNA isolation kit according to the manufacturer's protocol. Real-time PCR analysis was performed using first-strand cDNA as a template with the QuantiTect SyBR Green PCR kit (Clontech Laboratories Inc., Mountain View, CA, USA). The PCR threshold cycle number of each gene was normalized to the expression level of soybean EUKARYOTIC ELONGATION FACTOR 1-ALPHA (*GmELF1a*) as a reference gene. The normalized transcript levels were expressed as relative values of the dark-grown level. The relative expression level of the gene was calculated following the method of the $2^{-\Delta\Delta Ct}$ comparative Ct.

2.6 Statistical analysis

Each experimental treatment was performed in three replications, and the instrumental analysis was also triplicated per treatment. Two-way analysis of variance was performed to assess differences among light wavelength and irradiation period factors by Fisher's least significant difference (LSD) test using SAS software (Enterprise Guide 4.3 version; SAS Institute Inc., Cary, NC, USA). *Post-hoc* test was performed after ANOVA to assess significant differences between the light treatments using Tukey's studentized range (HSD) test at the level of $p < 0.05$. Metabolite data for the heatmap were normalized and plotted using Tbtools V1.098693. Circular heatmap was expressed as relative content among compounds and light treatments within each category and also expressed as relative content among light treatments within each compound.

3 Results

3.1 Seedling growth and chlorophyll content

Red and blue light strikingly affected shoot elongation (Figure 1; Table 1). As the irradiation period increased, shoot elongation was reduced, with blue light inhibiting it more than red light during the same irradiation period. While irradiation of both lights for 12 h slightly increased shoot length, long-term irradiation (120 h) of red and blue lights suppressed shoot length by 5% and 34% of dark-grown seedlings, respectively. In contrast, the long-term irradiation of either red or blue lights stimulated the root length compared to dark-grown seedlings. No significant difference was observed in root length between dark and short-term-exposed (12 and 36 h) seedlings.

The cotyledon dry weight was slightly increased as the red light irradiation period increased but was decreased by blue light compared to dark regardless of period. The hypocotyl dry weight followed a similar pattern as the shoot length. The total chlorophyll content (TCC) in cotyledon increased as the red and blue light irradiation periods increased. The TCC content remained unchanged for both lights for 12 h irradiation, but red light increased TCC relative to blue light at 36 and 120 h irradiation.

3.2 Metabolites profiles in soybean seedlings under red and blue lights

A total of 75 metabolites were identified in 8 categories—9 isoflavones, 6 flavonols, 2 phenolic acids, 16 amino acids, 10 organic acids, 21 free sugars, 6 alcohol sugars, and 5 sugar acids—by HPLC and GC-MS analysis. The circular heatmaps clearly show differences in the content among metabolites in each category and among the light treatment groups of each metabolite (Figure 2). The predominant metabolites and light-responded metabolites within each category were identified (Figure 2A). Notably, asparagine was dominant among the amino acids (61.8%–71.1% of total amino acids) in soybean seedling components. Fructose (25.2%–53.9% of total free sugars) and sucrose (20.3%–39.9%) were dominant among the free sugars. D-pinitol (82.7%–87.5% of total alcohol sugars) and galactaric acid (44.9%–49.2% of total sugar acids) were the dominant alcohol sugar and sugar acid, respectively. These dominant compounds were the most abundant under both light and dark conditions.

Figure 2B shows the different accumulation patterns of each metabolite following red and blue light irradiation. Even compounds in the same category showed different light responses depending on the presence or absence of light, light wavelength, and light irradiation period. In general, the light-induced accumulation pattern of metabolites was observed in the following three patterns:

FIGURE 1
Morphology of 7-day-old soybean seedlings grown under darkness and different exposure times of red (650 nm) and blue (447 nm) light.

TABLE 1 Length, dry weight, and total chlorophyll content in soybean seedlings grown under darkness, red (650 nm), and blue (447 nm) light.

Light	Time (h)	Length (cm)		Dry weight (mg)			TCC ($mg \cdot g^{-1}$ DW)
		Shoot	Root	Cotyledon	Hypocotyl	Root	
Dark		12.51 ± 0.26ab	7.20 ± 0.18b	547.1 ± 20.04ab	157.63 ± 5.35b	33.48 ± 0.95c	0.24 ± 0.01e
Red	12	13.40 ± 0.20a	7.22 ± 0.21b	520.01 ± 22.11ab	219.05 ± 6.04ab	50.34 ± 1.32bc	1.14 ± 0.09d
	36	12.31 ± 0.20ab	7.49 ± 0.17b	575.59 ± 11.34a	292.94 ± 75.56a	67.64 ± 3.95ab	3.10 ± 0.10a
	120	11.92 ± 0.89ab	9.43 ± 0.19a	583.23 ± 31.83a	226.93 ± 1.00ab	83.91 ± 3.90a	2.83 ± 0.06a
Blue	12	13.26 ± 0.24a	6.89 ± 0.25b	434.17 ± 2.53b	231.65 ± 8.98ab	64.67 ± 8.73ab	1.10 ± 0.06d
	36	11.46 ± 0.16b	6.65 ± 0.19b	448.78 ± 4.93b	216.41 ± 6.52ab	69.22 ± 4.89ab	1.76 ± 0.01c
	120	8.22 ± 0.33c	9.45 ± 0.21a	498.41 ± 36.06ab	189.34 ± 23.54b	67.39 ± 1.50ab	2.12 ± 0.07b

Different letters (a-e) in each column indicate significant differences at $p < 0.05$ by Tukey's studentized range (HSD) test. TCC, total chlorophyll content in cotyledon; DW, dry weight.

increase with short-term irradiation of red (12, 36 h) in this experimental design, increase with long-term irradiation of blue, or decrease following light irradiation. Certain sugar types, such as D-(-)-fructofuranose and D-xylose, were kept in higher accumulation under short-term red light irradiation (Figure 2B). Among the eight categories, flavonols were the most significant components that were increased by red and blue light irradiation. Also, it was observed that some specific isoflavones and phenolic acids were light induced metabolites. Therefore, to further reveal the isoflavone–flavonol regulatory mechanism in response to light, the metabolic gene expression level and accumulation of individual isoflavone and flavonol were investigated.

3.3 Changes in isoflavone and kaempferol glycoside

Figure 3 shows the changes in the content of eight isoflavones and six kaempferol glycosides. The contents of malonyl daidzin and malonyl genistin, which were dominant in soybean seedlings, significantly increased by 25.7% and 54.8% in 12 h and by 36.8% and 60.7% in 36 h red light irradiation compared to dark but did not change in long-term irradiation (Figure 3A). However, blue light accumulated two malonyl glycosides in different patterns. Malonyl daidzin content increased time-dependently by blue light with a maximum content in 120 h (121.2% of dark-grown seedlings), whereas malonyl genistin content was maintained after increasing at 12 h (140.5% of dark-grown seedlings). The accumulation patterns of seven isoflavones, except for daidzin and genistin, differed between the red and blue lights. Both red and blue light increased daidzin content but did not affect genistin content. Aglycone (daidzein and genistein) showed different patterns between red and blue light, although they were contained in small trace amounts in the seedlings. Red light promoted genistein accumulation more than daidzein in the short term. Conversely, in blue light, daidzein content increased in short-term irradiation, whereas genistein content increased in long-term irradiation. In particular, the content of glycitin-type isoflavones (glycitin and malonyl glycitin) gradually decreased as the red light irradiation period increased, whereas the content decreased for 12 h of blue light irradiation and gradually reversed following prolonged irradiation.

Kaempferol glycosides were controlled by a light irradiation period, exhibiting higher kaempferol glycoside content on longer light irradiation, regardless of photosynthetic lights (Figure 3B). Kaempferol glycoside contents peaked at 120 h irradiation, showing 1.62- to 2.98-fold higher in red light and 1.29- to 3.48-fold higher in blue light depending on kaempferol glycoside types than in dark-grown seedling. Kaempferol-3-*O*-glycosyl(1→2)-rhamnosyl(1→6)-galactoside (Kf1) and kaempferol-3-*O*-(2,6-dirhamnosyl)-galactoside (Kf2) were the predominant kaempferol glycosides in soybean seedlings, as previously reported (Lim et al., 2021). They did not respond to short-term irradiation of 12 h but thereafter significantly increased time-dependently. Five kaempferol glycosides were increased time-dependently by both red and blue light, except for kaempferol-3-*O*-diglucoside (Kf4). Kf4 did not increase after 36 h in either red or blue light irradiations. Notably, the accumulation of Kf2 was primarily stimulated by blue light compared to red light, with 1.7- and 1.2-fold higher content in 36 and 120 h irradiation of blue light than red light.

3.4 Changes in biosynthetic gene expression and accumulation of flavonoids

Figure 4 shows the expression levels of *GmPAL, GmC4H, Gm4CL, GmCHS1–8, GmCHR, GmCHI1AII* and *1B1II, GmIFS1–2, GmIFR*, and *GmF3H* genes involved in flavonoid biosynthesis. In red light, the expression levels of all genes except *GmCHS3/4* and *GmIFS1* gradually increased up to 36 h and then decreased at 120 h, which was similar to the accumulation patterns of malonyl daidzin and malonyl genistin contents. In blue light, the expression levels of most genes were downregulated time-dependently after upregulation in short-term irradiation (12 h), whereas those of *Gm4CL*, *GmCHS2/5*, and *GmIFS1* were gradually upregulated time-dependently.

Concerning *GmCHSs*, the *GmCHS1–5* expression levels were averagely higher in red light than in blue light. However, the *GmCHS6/7/8* expression levels responded to 12 *h* blue light irradiation, showing a high level similar to that of 36 h red light

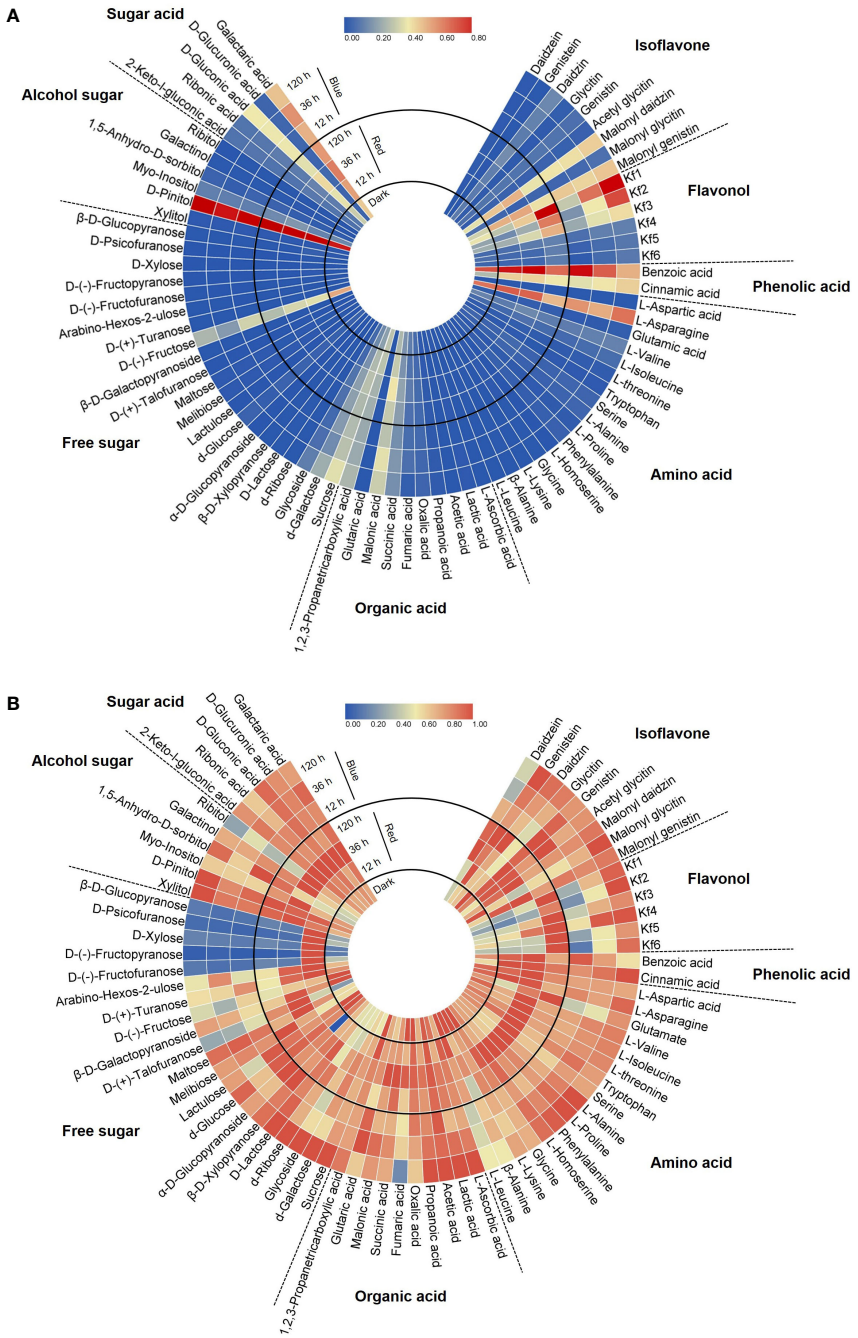


FIGURE 2
Circular heatmap based on the relative content in each category **(A)** and the relative content in each compound **(B)**. Metabolite data for the heatmap were normalized and plotted using Tbtools V1.098693. Kf1, kaempferol-3-*O*-glycosyl(1→2)-rhamnosyl(1→6)-galactoside; Kf2, kaempferol-3-*O*-(2,6-dirhamnosyl)-galactoside; Kf3, kaempferol-3-*O*-digalactoside; Kf4, kaempferol-3-*O*-diglucoside; Kf5, kaempferol-3-*O*-rhamnosyl-galactoside; Kf6, kaempferol-3-*O*-rutinoside.

irradiation. Among *CHSs*, *GmCHS5–6* expression levels were the highest in each 36 h of red light and 12 *h* of blue light, with 10- and 11-fold higher expression than in dark, respectively. Interestingly, *GmCHS4* expression was inhibited by both red and blue light and had a lower level than that in dark conditions. Genes involved in isoflavone synthesis, *GmIFS1–2*, displayed distinct patterns under red and blue light. Under red light irradiation, the *GmIFS2* expression level was highest at 36 h, whereas that of *GmIFS1* decreased time-dependently. Under blue light irradiation, *GmIFS1* was highly expressed compared to dark conditions, whereas *GmIFS2* remained unaltered by blue light and was upregulated only by red light.

Red and blue light significantly upregulated *GmF3H* relative to other genes. Red light significantly increased *GmF3H* after 36 h irradiation, with 17-fold higher expression compared to the dark. Most flavonoid biosynthetic genes were significantly more

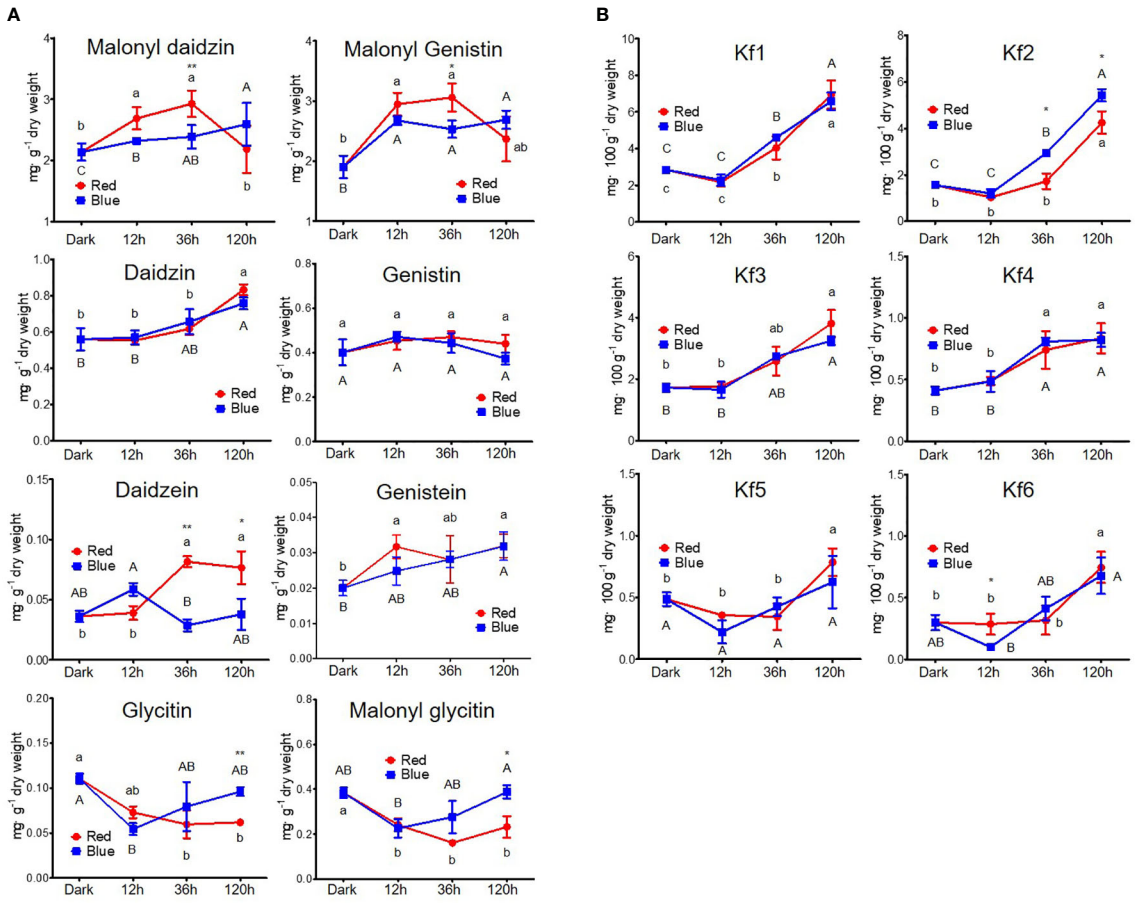


FIGURE 3
Content of flavonoids, including isoflavone **(A)** and kaempferol glycoside **(B)**. The following abbreviations indicate Kf1, kaempferol-3-*O*-glycosyl (1→2)-rhamnosyl(1→6)-galactoside; Kf2, kaempferol-3-*O*-(2,6-dirhamnosyl)-galactoside; Kf3, kaempferol-3-*O*-digalactoside; Kf4, kaempferol-3-*O*-diglucoside; Kf5, kaempferol-3-*O*-rhamnosyl-galactoside; Kf6, kaempferol-3-*O*-rutinoside; RE, robinin equivalent. Different lower and upper case letters (a–c/A-C) indicate significant differences at $p < 0.05$ by Tukey's studentized range (HSD) test between irradiation periods within each red and blue light, respectively. Asterisks indicate statistically significant differences (*$p < 0.05$; **$p < 0.01$) between red and blue light by Fisher's LSD test.

upregulated by short-term red light irradiation than by blue light. However, long-term irradiation of both lights led to the inhibition of flavonoid biosynthetic gene expression, resulting in a temporal discrepancy between gene expression and flavonoid accumulation.

3.5 Amino acids, phenolic acids, and free sugars involved in flavonoid synthesis

Figure 5 shows the quantification of flavonoid-related intermediate metabolites. Phenylalanine, a precursor for flavonoid synthesis, showed the highest content at 36 h in red light with a 1.3-fold higher content than in dark-grown seedlings, whereas it maximally increased at 120 h in blue light, presenting similar max content in red and blue lights. Tryptophan, competitively synthesized with phenylalanine, was not significantly affected by either light.

Benzoic and cinnamic acids were accumulated differently by red and blue light in soybean seedlings. Cinnamic acid, a precursor phenolic acid of flavonoids, increased by photosynthetic lights compared to dark conditions, regardless of irradiation period and light wavelength. Cinnamic acid was time-dependently increased by blue light, whereas it rapidly increased in red light short-term irradiation. The contents were 1.7-fold higher than in dark-grown seedlings in both 120 h of blue and 12 h of red light. Meanwhile, benzoic acid, in a competitive pathway with flavonoid synthesis, was considerably affected by the irradiation period. While 12 and 36 h red light irradiation stimulated benzoic acid accumulation, 120 h irradiation did not. However, in blue light, benzoic acid content increased by 20.6% at 12 h compared to dark conditions but decreased time-dependently after that, resulting in a 28% decrease in the content at 120 h.

D-galactose and D-glucose, which are sugar donors for flavonoid glycoside synthesis, accumulated differently under red and blue light. The concentrations of both sugars remained unchanged after 36 h of exposure to red light, while they increased significantly after 120 h of exposure to blue light. The upregulation of free sugar accumulation by photosynthetic light may have contributed to the increased flavonol glycoside content.

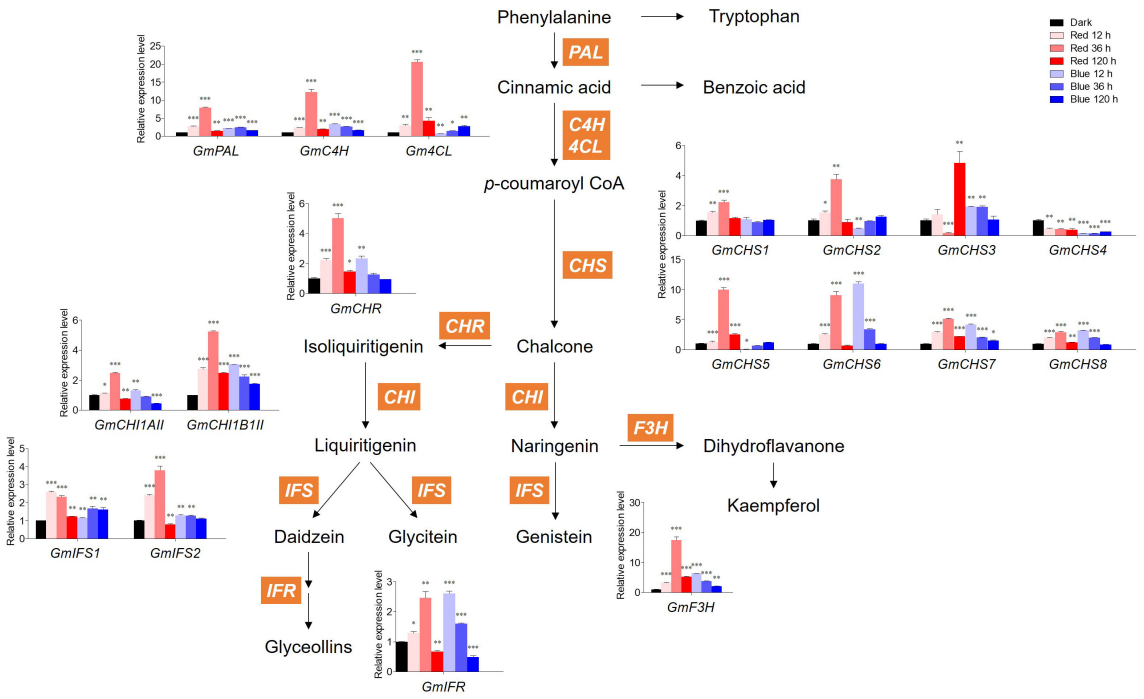


FIGURE 4
Biosynthetic gene expression associated with flavonoids. The transcript levels were normalized with the reference gene *GmELF1a* and expressed as relative values of the dark-grown level. The following abbreviations of genes indicate *PAL*, phenylalanine ammonia lyase; *C4H*, Cinnamic acid 4-hydroxylase; *4CL*, 4-coumarate:CoA ligase; *CHS*, chalcone synthase; *CHR*, chalcone reductase; *CHI1AII*, chalcone isomerase 1A Type II; *CHI1B1II*, chalcone isomerase 1B1 Type II; *IFS*, isoflavone synthase; *IFR*, isoflavone reductase; *F3H*, flavanone 3-hydroxylase. Asterisks indicate statistically significant differences compared with dark in each gene (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$) by Fisher's LSD test.

4 Discussion

4.1 Metabolic profiles in response to photosynthetic light and light irradiation periods

Circular heatmap data clearly showed the relative level of photosynthetic light response for a set of metabolites, including amino acids, organic acids, flavonoids, phenolic acids, free sugars, alcohol sugars, and sugar acids. These metabolic profiling results enabled the distinction between light-responsive and nonresponsive metabolites and revealed that the photosynthetic light response of specialized metabolites is linked to several metabolites. Photosynthetic light, such as red and blue light, markedly affects plant metabolism. Light promotes the metabolism of phenylalanine, a key amino acid precursor for flavonoid synthesis, *via* the upregulation of the phenylalanine ammonia lyase reaction (Seo et al., 2015; Liu et al., 2018). Light also increases flavonoid concentrations by activating the biosynthesis and metabolism of lipids (Maldini et al., 2015).

Metabolic profiling provides insight into the degree to which germinating soybean seedlings respond to short-term or long-term irradiation of two different light wavelengths with respect to their chemical composition, including primary and specific secondary metabolites. Even for compounds belonging to the same category, metabolic changes based on the presence or absence of light, light wavelength, and irradiation period showed significant light response differences. Through untargeted metabolomics based on GC-MS and HPLC analysis, specific metabolites that can be controlled by the wavelength and period of light were identified, allowing for further studies on their light response mechanisms.

4.2 Light-induced synthesis of two flavonoid branches of isoflavone and flavonol

The specific responses of flavonoids to photosynthetic light have been distinguished from those of other metabolites. Most plants have a common major flavonoid pathway, but they frequently derive specific branches to adapt to variable environmental conditions (García-Calderón et al., 2020). Connecting primary and secondary metabolism, the chalcone synthesis pathway provides precursors for the synthesis of multibranched downstream metabolites (Austin and Noel, 2003). Two branches of flavonoids, isoflavone and flavonol, are derived from chalcone. Although many studies have explored the effect of light on isoflavone production in soybeans, the light-controlled coregulation of the isoflavone and flavonol pathways remains underexplored. Our results revealed that the synthesis of kaempferol glycosides was highly dependent on red and blue irradiation periods, whereas isoflavone accumulation was type-dependent in response to each light quality.

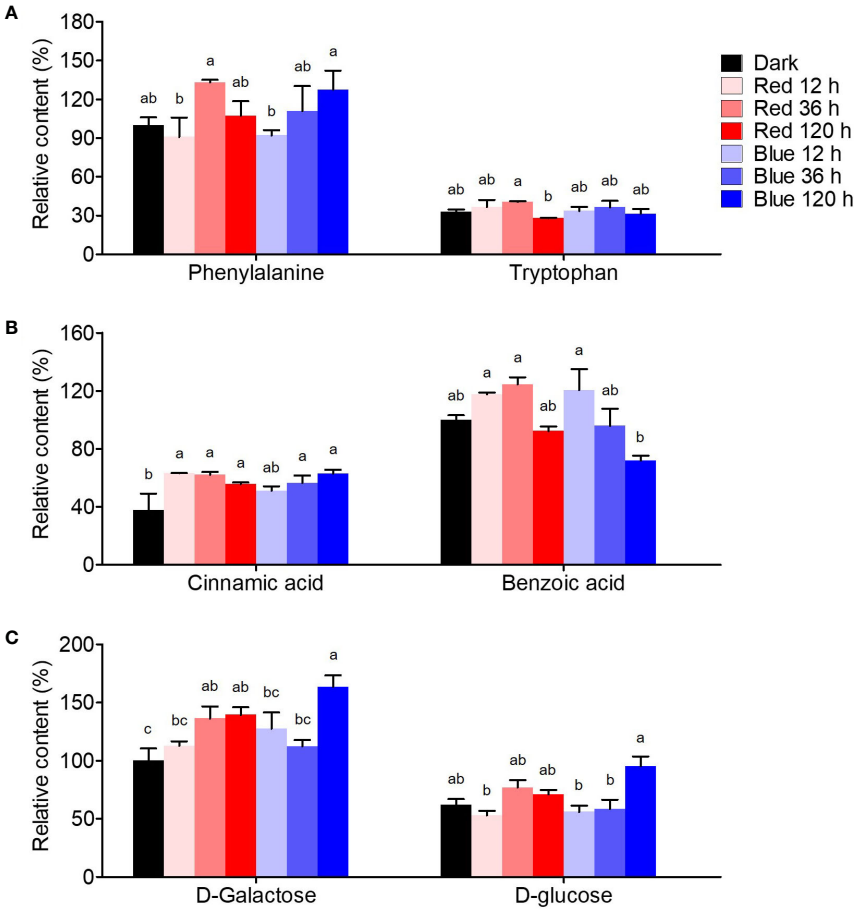


FIGURE 5
Flavonoid synthesis-related metabolites, including amino acid **(A)**, phenolic acid **(B)**, and free sugar **(C)**. The content of amino acid, phenolic acid, and free sugar was expressed as relative content (%) of phenylalanine, benzoic acid, and D-galactose of dark-grown seedlings in each metabolite group, respectively. Different letters (a–c) above the bars in each metabolite indicate significant differences at $p < 0.05$ by Tukey's studentized range (HSD) test.

Light-induced isoflavone accumulation was regulated differently based on the isoflavone type. Type-dependent synthetic pathways exist for isoflavone. The genistein synthetic pathway differs from that of daidzein and glycitein (Figure 4), where genistein is from naringenin while daidzein and glycitein are from isoliquiritigenin (Ralston et al., 2005). In addition, different types of isoflavone are organ-specifically accumulated, presenting different isoflavone compositions in each organ of soybean plants (Graham, 1991; Yang et al., 2020; Lim et al., 2021). Although genistin and malonyl genistin were increased by photosynthetic light, they did not differ significantly from other isoflavone types. Regarding the chemical structure of isoflavones, glycitein has a methoxy group, unlike daidzein and genistein. According to Fu et al. (2016), flavonol methyl derivatives in tobacco negatively correlate with ultraviolet light and red/far-red light. Red and blue light irradiation decreased the glycitein type in the present study, validating the negative correlations between isoflavone methyl derivatives and photosynthetic light. Long-term red light irradiation induced a continuous decrease in the glycitein type, indicating that the negative correlation is more dependent on red light. In addition, red and blue light-responded accumulation of isoflavone in soybean seedlings was limited to daidzein and genistein types, corroborating previous reports (Azad et al., 2018). These differences in synthetic pathways, tissue-specific accumulation, and chemical structures would have resulted in different light-responded accumulations depending on each isoflavone type.

Several studies have shown the enhancing effects of blue light and ultraviolet light on various flavonoids, including isoflavone (Jenkins et al., 1995; Zoratti et al., 2014; Taulavuori et al., 2018). However, little is known about the effect of red light on isoflavones. Specific red-light-responsive isoflavones remain unexplored, but red light has been shown to increase isoflavone glycoside content without affecting isoflavone aglycones (Kirakosyan et al., 2006). In addition, red and far-red light influenced isoflavone accumulation based on the shade tolerance level of soybean plants, with a decrease in the red/far-red ratio increasing leaf isoflavones and reducing hypocotyl isoflavones in high shade-tolerance soybeans (Qin et al., 2017). Although the isoflavone accumulation pattern in response to light quality varied depending on the type, this study found that short-term red light irradiation could be efficient for producing malonyl daidzin and malonyl genistin, the dominant isoflavone in soybeans. Long-term red light irradiation, however, negatively affected the production of glycitein-type isoflavones.

All kaempferol glycoside types increased in response to photosynthetic light. The accumulation of kaempferol glycosides was highly dependent on the light irradiation period, indicating that the accumulation of flavonol is more light-dependently regulated than isoflavone. Although the content of all kaempferol species was increased by red and blue light, Kf2 production was promoted by blue light rather than red light. The regulation of specific kaempferol derivatives by controlling light factors is rarely studied, but blue light promoted the accumulation of kaempferol aglycone compared to red light in Chinese cabbage seedling and lettuce leaf (Kim et al., 2015; Lee et al., 2019). Notably, the attenuation of blue light can result in an increase in certain kaempferol derivatives and a decrease in quercetin derivatives in pea leaves (Sipola et al., 2015). In addition, blue light stabilized the UV-induced high concentration of kaempferol derivatives in kale seedlings, whereas this effect did not apply to kohlrabi seedlings (Neugart et al., 2021). Although both light wavelength and irradiation period have a significant effect on flavonol regulation depending on the crop, our results indicate that light irradiation period rather than the light wavelength is a critical factor for the regulation of kaempferol glycoside accumulation in soybean seedlings.

Kaempferol mainly accumulates in aerial parts that receive direct light in many plants such as soybean, turnip, and fennel, and is not found or in trace amounts in root tissue with or without light (Soliman et al., 2002; Fernandes et al., 2007; Karimi et al., 2011; Lim et al., 2021). Our previous study revealed that ultraviolet light-induced Kf2 accumulation in soybean seedlings occurred primarily in the hypocotyl (Lim et al., 2021). Therefore, blue light-induced Kf2 is inferred by an increase in its content in hypocotyl. Various flavonoids, as well as isoflavones and kaempferol derivatives, are generally regulated by photosynthetic light. It has known that higher light intensity and longer period in red and blue light irradiation induced higher level of anthocyanin in red kale sprouts (Carvalho and Folta, 2014). In addition, flavonoids in buckwheat sprouts, including orientin, isoorientin, vitexin, isovitexin, quercetin 3-*O*-robinobioside, and rutin, were also enhanced by sequential red and blue light irradiation (Nam et al., 2018).

4.3 Flavonoid biosynthetic genes in response to red and blue lights

Overall, with some exceptions, flavonoid biosynthetic genes were maximally expressed at 36 h under red light, peaking at 12 h and gradually decreasing under blue light. The expression of many flavonoid biosynthetic genes in soybean seedlings peaked at 36 h of red light, showing a similar pattern to malonyl-type isoflavone accumulation. The expression patterns of *GmCHS2/5* decreased at 12 h under blue light and then gradually increased, consistent with the accumulation pattern of glycitein-type isoflavones, indicating that *GmCHS2/5* are mainly involved in the synthesis of glycitein types only in response to blue light. These results indicate that the time lag between gene expression and metabolite accumulation is dependent on light wavelength. Different gene-to-metabolite patterns depending on light wavelength have also been observed in strawberry fruits (Zhang et al., 2018). Under blue light, strawberry fruits had lower gene expression levels, inconsistent with their high anthocyanin content, whereas under red light, most genes remained active even though the anthocyanin content was low.

The expression of *GmCHSs* in soybean seedlings was highly variable based on the wavelength and period of light. *GmCHSs* were upregulated in response to red light rather than blue light, but exceptionally, *GmCHS6/7/8* were upregulated in response to both red and blue lights. Differences in flavonoid gene expression in response to red and blue lights are explained by differences in photoreceptors and their signaling pathways. Gene expression is regulated at the transcriptional and posttranscriptional levels by light signals detected by photoreceptors and transduced in photosystems (Martınez-Hernández et al., 2002; Floris et al., 2013). Blue light induces *CHS*, which is predominantly mediated by cry 1 and does not require phytochrome. However, red light upregulates cryptochrome 1-mediated *CHS* expression because phytochrome regulates the cry 1 inductive pathway (Wade et al., 2001). Our qRT-PCR results demonstrated that the mechanism by which the *GmCHS* expression level was differentially regulated depended on light wavelength because red and blue lights have distinct signaling pathways. Previous reports have shown that *CHS6/7/8* sensitively responds to ultraviolet light (Lim et al., 2020). UV-A/blue light-induced *CHS* expression in a signal transduction pathway distinct from the UV-B response (Christie and Jenkins, 1996). However, our previous and current studies demonstrated that *GmCHS6/7/8* were significantly upregulated by UV-A/UV-B/blue light compared to other *CHSs* and were highly sensitive to shorter wavelengths, including UV-A/B/blue light.

UV radiation, as a stress factor, can downregulate the expression of flavonoid biosynthesis genes (Lim et al., 2020; Lim et al., 2021). However, our results show that excessive light exposure, regardless of light wavelength, can inhibit the expression level of flavonoid biosynthetic genes. Light intensity and light wavelength strongly influence the expression of genes involved in flavonoid synthesis (Xu et al., 2014). Flavonoid accumulation patterns were not consistent with most gene expression patterns. Long-term light irradiation inhibited flavonoid biosynthetic gene expression, resulting in a periodic discrepancy between gene expression and flavonoid accumulation. The gene-to-metabolite discrepancy is commonly observed and is due to the time lag from gene transcription to metabolite synthesis (Nakabayashi et al., 2017). Nevertheless, the expressions of most genes involved in flavonoid synthesis were highly upregulated by short-term red light irradiation (12 and 36 h) compared to blue light, exhibiting a strong correlation between red light and the accumulation of malonyl types.

4.4 The relationship between flavonoids, upstream metabolites, and primary metabolites

Various intermediate metabolites are required and are involved in flavonoid synthesis. The red/blue light-induced increase in phenylalanine, a precursor amino acid of flavonoids, resulted in an increase in flavonol content. Tryptophan, which is in a

competitive pathway with phenylalanine, did not change significantly in response to light, indicating that phenylalanine-related metabolism is more light-dependent. Although phenylalanine metabolism contributes to the formation of various metabolites such as proteins, tyrosine derivatives, polyamines, tannins, and lignin as well as flavonoids (Bassard et al., 2010; Barros and Dixon, 2020), it was highly related to flavonoid synthesis in soybean seedlings under red and blue lights. Under blue light, cinnamic acid increased, whereas benzoic acid significantly decreased, indicating that blue light inhibits the synthesis of cinnamic to benzoic acid. The reduction of phenylalanine and cinnamic acid by long-term red light irradiation affected malonyl isoflavone accumulation. Although sugars are involved in various metabolisms, the sufficient provision of free sugars induced by red and blue lights has a significant effect on the increase in flavonoid content. Consequently, the sufficient production of flavonoid precursors induced by red and blue lights and the inhibition of the competitive pathway by blue light resulted in the upregulation of isoflavone and flavonol synthesis in soybean seedlings.

5 Conclusion

Here, we suggest that a sequential light irradiation system with single red or blue light is an effective method to regulate the metabolism of specific isoflavone and flavonol in soybean seedling. Among the eight metabolite categories, flavonols were the most significant components that were increased by red and blue light irradiation. Also, it was observed that some specific isoflavones and phenolic acids were light induced metabolites. The light irradiation period rather than the light wavelength was an important factor in the regulation of kaempferol glycoside accumulation. In particular, kaempferol-3-*O*-(2, 6-dirhamnosyl)-galactoside showed higher accumulation under blue light than red light. Higher accumulation of kaempferol derivatives was induced by a longer light irradiation period; however, kaempferol-3-*O*-diglucoside production unexpectedly stopped increasing after 36 hours of both red and blue light. On the contrary, the pattern of the light-responded accumulation of isoflavones considerably varied depending on the type of isoflavone. Red light was effective in increasing total isoflavone levels by inducing high accumulation of malonyl daidzin and malonyl genistin, the main isoflavones of soybean seedling. However, longer irradiation of red light (120 h) caused a reduction of those main isoflavones, indicating that isoflavone is significantly affected not only by the light wavelength but also by the irradiation period. Daidzein increased only in response to red light, while glycitin types decreased in response to the prolonged red light irradiation, suggesting that isoflavone structural specificity results in the accumulation of different isoflavone profiles in response to light. In conclusion, our findings imply that a single red and blue light can be used to selectively regulate isoflavone and flavonol in soybean seedlings and that the light wavelength and irradiation period are critical in deciding which flavonoids are being targeted.

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

SE designed and supervised the project. YL designed the experiment and drafted the manuscript. YL and S-JK performed laboratory experiments and analyzed data. SE participated in the material preparation. SE and S-JK revised the manuscript. All authors contributed to the article and approved the final version of the manuscript.

Funding

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (NRF-2021R1A6A3A01086968, NRF-RS-2022-00156231) and by KAERI R&D program (523320-23).

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.2023.1128001/full#supplementary-material

SUPPLEMENTARY TABLE 1
Gene specific primers sequences for qRT-PCR.

SUPPLEMENTARY FIGURE 1
Light spectra of the red (A) and blue (B) LEDs used in this experiment.

SUPPLEMENTARY FIGURE 2
Total ion chromatograms by GC-MS of soybean seedling grown under different light treatments.

References

Anthony, M. S., Clarkson, T. B., and Williams, J. K. (1998). Effects of soy isoflavones on atherosclerosis: Potential mechanisms. *Am. J. Clin. Nutr.* 68, 1390S–1393S. doi: 10.1093/ajcn/68.6.1390S

Arjmandi, B. H., and Smith, B. J. (2002). Soy isoflavones' osteoprotective role in postmenopausal women: Mechanism of action. *J. Nutr. Biochem.* 13, 130–137. doi: 10.1016/S0955-2863(02)00172-9

Austin, M. B., and Noel, J. P. (2003). The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* 20, 79–110. doi: 10.1039/B100917F

Azad, M. O. K., Kim, W. W., Park, C. H., and Cho, D. H. (2018). Effect of artificial LED light and far infrared irradiation on phenolic compound, isoflavones and antioxidant capacity in soybean (*Glycine max* l.) sprout. *Foods* 7, 174. doi: 10.3390/foods7100174

Barros, J., and Dixon, R. A. (2020). Plant phenylalanine/tyrosine ammonia-lyases. *Trends Plant Sci.* 25, 66–79. doi: 10.1016/j.tplants.2019.09.011

Bassard, J. E., Ullmann, P., Bernier, F., and Werck-Reichhart, D. (2010). Phenolamides: bridging polyamines to the phenolic metabolism. *Phytochemistry* 71, 1808–1824. doi: 10.1016/j.phytochem.2010.08.003

Bian, Z. H., Yang, Q. C., and Liu, W. K. (2015). Effects of light quality on the accumulation of phytochemicals in vegetables produced in controlled environments: A review. *J. Sci. Food Agric.* 95, 869–877. doi: 10.1002/jsfa.6789

Carvalho, S. D., and Folta, K. M. (2014). Sequential light programs shape kale (*Brassica napus*) sprout appearance and alter metabolic and nutrient content. *Hort Res.* 1, 8. doi: 10.1038/hortres.2014.8

Christie, J. M., and Jenkins, G. I. (1996). Distinct UV-b and UV-a/blue light signal transduction pathways induce chalcone synthase gene expression in arabidopsis cells. *Plant Cell* 8, 1555–1567. doi: 10.1105/tpc.8.9.1555

De Melo, G. O., Malvar, D. D. C., Vanderlinde, F. A., Rocha, F. F., Pires, P. A., Costa, E. A., et al. (2009). Antinociceptive and anti-inflammatory kaempferol glycosides from sedum dendroideum. *J. Ethnopharmacol.* 124, 228–232. doi: 10.1016/j.jep.2009.04.024

Di Ferdinando, M., Brunetti, C., Fini, A., and Tattini, M. (2012). Flavonoids as antioxidants in plants under abiotic stresses. *Abiotic Stress responses Plants*, 159–179. doi: 10.1007/978-1-4614-0634-1_9

Eom, S. H., Jin, S. J., Jeong, H. Y., Song, Y., Lim, Y. J., Kim, J. I., et al. (2018). Kudzu leaf extract suppresses the production of inducible nitric oxide synthase, cyclooxygenase-2, tumor necrosis factor-alpha, and interleukin-6 *via* inhibition of JNK, TBK1 and STAT1 in inflammatory macrophages. *Int. J. Mol. Sci.* 19, 1536. doi: 10.3390/ijms19051536

Fernandes, F., Valentão, P., Sousa, C., Pereira, J. A., Seabra, R. M., and Andrade, P. B. (2007). Chemical and antioxidative assessment of dietary turnip (*Brassica rapa* var. rapa l.). *Food Chem.* 105, 1003–1010. doi: 10.1016/j.foodchem.2007.04.063

Floris, M., Bassi, R., Robaglia, C., Alboresi, A., and Lanet, E. (2013). Post-transcriptional control of light-harvesting genes expression under light stress. *Plant Mol. Biol.* 82, 147–154. doi: 10.1007/s11103-013-0046-z

Fu, B., Ji, X., Zhao, M., He, F., Wang, X., Wang, Y., et al. (2016). The influence of light quality on the accumulation of flavonoids in tobacco (*Nicotiana tabacum* l.) leaves. *J. Photochem. Photobiol. B Biol.* 162, 544–549. doi: 10.1016/j.jphotobiol.2016.07.016

García-Calderón, M., Pérez-Delgado, C. M., Palove-Balang, P., Betti, M., and Márquez, A. J. (2020). Flavonoids and isoflavonoids biosynthesis in the model legume *Lotus japonicus*; connections to nitrogen metabolism and photorespiration. *Plants* 9, 774. doi: 10.3390/plants9060774

Graham, T. L. (1991). Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates. *Plant Physiol.* 95, 594–603. doi: 10.1104/pp.95.2.594

Gu, E. J., Kim, D. W., Jang, G. J., Song, S. H., Lee, J. I., Lee, S. B., et al. (2017). Mass-based metabolomic analysis of soybean sprouts during germination. *Food Chem.* 217, 311–319. doi: 10.1016/j.foodchem.2016.08.113

Ho, H. M., Chen, R. Y., Leung, L. K., Chan, F. L., Huang, Y., and Chen, Z. Y. (2002). Difference in flavonoid and isoflavone profile between soybean and soy leaf. *Biomed. Pharmacother.* 56, 289–295. doi: 10.1016/S0753-3322(02)00191-9

Jenkins, G. I., Christie, J. M., Fuglevand, G., Long, J. C., and Jackson, J. A. (1995). Plant responses to UV and blue light: Biochemical and genetic approaches. *Plant Sci.* 112, 117–138. doi: 10.1016/0168-9452(95)04260-1

Karimi, E., Jaafar, H. Z., and Ahmad, S. (2011). Phytochemical analysis and antimicrobial activities of methanolic extracts of leaf, stem and root from different varieties of *Labisa pumila* benth. *Molecules* 16, 4438–4450. doi: 10.3390/molecules16064438

Kim, B. G., Kim, J. H., Kim, J., Lee, C., and Ahn, J. (2008). Accumulation of flavonols in response to ultraviolet-b irradiation in soybean is related to induction of flavanone 3-beta-hydroxylase and flavonol synthase. *Mol. Cells* 25, 247. doi: 10.14348/.1970.0.0

Kim, Y. J., Kim, Y. B., Li, X., Choi, S. R., Park, S., Park, J. S., et al. (2015). Accumulation of phenylpropanoids by white, blue, and red light irradiation and their organ-specific distribution in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *J. Agric. Food Chem.* 63, 6772–6778. doi: 10.1021/acs.jafc.5b02086

Kirakosyan, A., Kaufman, P., Nelson, R. L., Kasperbauer, M. J., Duke, J. A., Seymour, E., et al. (2006). Isoflavone levels in five soybean (*Glycine max*) genotypes are altered by phytochrome-mediated light treatments. *J. Agric. Food Chem.* 54, 54–58. doi: 10.1021/jf052458w

Křížová, L., Dadáková, K., Kašparovská, J., and Kašparovský, T. (2019). Isoflavones. *Molecules* 24, 1076. doi: 10.3390/molecules24061076

Kusano, M., Fukushima, A., Redestig, H., and Saito, K. (2011). Metabolomic approaches toward understanding nitrogen metabolism in plants. *J. Exp. Bot.* 62, 1439–1453. doi: 10.1093/jxb/erq417

Lee, M., Xu, J., Wang, W., and Rajashekar, C. B. (2019). The effect of supplemental blue, red and far-red light on the growth and the nutritional quality of red and green leaf lettuce. *Am. J. Plant Sci.* 10, 2219–2235. doi: 10.4236/ajps.2019.1012157

Li, T., Wang, Y. H., Liu, J. X., Feng, K., Xu, Z. S., and Xiong, A. S. (2019). Advances in genomic, transcriptomic, proteomic, and metabolomic approaches to study biotic stress in fruit crops. *Crit. Rev. Biotechnol.* 39 (5), 680–692. doi: 10.1080/07388551.2019.1608153

Li, Y., Xin, G., Liu, C., Shi, Q., Yang, F., and Wei, M. (2020). Effects of red and blue light on leaf anatomy, CO 2 assimilation and the photosynthetic electron transport capacity of sweet pepper (Capsicum annuum l.) seedlings. *BMC Plant Biol.* 20, 1–16. doi: 10.1186/s12870-020-02523-z

Lim, Y. J., Jeong, H. Y., Gil, C. S., Kwon, S. J., Na, J. K., Lee, C., et al. (2020). Isoflavone accumulation and the metabolic gene expression in response to persistent UV-b irradiation in soybean sprouts. *Food Chem.* 303, 125376. doi: 10.1016/j.foodchem.2019.125376

Lim, Y. J., Lyu, J. I., Kwon, S. J., and Eom, S. H. (2021). Effects of UV-a radiation on organ-specific accumulation and gene expression of isoflavones and flavonols in soybean sprout. *Food Chem.* 339, 128080. doi: 10.1016/j.foodchem.2020.128080

Liu, Y., Fang, S., Yang, W., Shang, X., and Fu, X. (2018). Light quality affects flavonoid production and related gene expression in *Cyclocarya paliurus*. *J. Photochem. Photobiol. B Biol.* 179, 66–73. doi: 10.1016/j.jphotobiol.2018.01.002

Ma, Y., Wang, P., Gu, Z., Sun, M., and Yang, R. (2022). Effects of germination on physio-biochemical metabolism and phenolic acids of soybean seeds. *J. Food Compos. Anal.* 112, 104717. doi: 10.1016/j.jfca.2022.104717

Ma, M., Wang, P., Yang, R., and Gu, Z. (2018). Effects of UV-b radiation on the isoflavone accumulation and physiological-biochemical change of germinated soybean induced by UV-b. *Food Chem.* 250, 259–267. doi: 10.1016/j.foodchem.2018.01.051

Ma, M., Zhang, H., Xie, Y., Yang, M., Tang, J., Wang, P., et al. (2020). Response of nutritional and functional composition, anti-nutritional factors and antioxidant activity in germinated soybean under UV-b radiation. *LWT* 118, 108709. doi: 10.1016/j.lwt.2019.108709

Mackinney, G. (1941). Absorption of light by chlorophyll solutions. *Biol. Chem.* 140, 315–322. doi: 10.1016/S0021-9258(18)51320-X

Maldini, M., Natella, F., Baima, S., Morelli, G., Scaccini, C., Langridge, J., et al. (2015). Untargeted metabolomics reveals predominant alterations in lipid metabolism following light exposure in broccoli sprouts. *Int. J. Mol. Sci.* 16, 13678–13691. doi: 10.3390/ijms160613678

Martınez-Hernández, A., López-Ochoa, L., Arguello-Astorga, G., and Herrera-Estrella, L. (2002). Functional properties and regulatory complexity of a minimal RBCS light-responsive unit activated by phytochrome, cryptochrome, and plastid signals. *Plant Physiol.* 128, 1223–1233. doi: 10.1104/pp.010678

Naik, J., Misra, P., Trivedi, P. K., and Pandey, A. (2022). Molecular components associated with the regulation of flavonoid biosynthesis. *Plant Sci.* 317, 111196. doi: 10.1016/j.plantsci.2022.111196

Na Jom, K., Frank, T., and Engel, K. H. (2011). A metabolite profiling approach to follow the sprouting process of mung beans (*Vigna radiata*). *Metabolomics* 7, 102–117. doi: 10.1007/s11306-010-0236-5

Nakabayashi, R., Mori, T., Nishizawa, T., and Saito, K. (2017). Temporal lag between gene expression and metabolite accumulation in flavonol biosynthesis of arabidopsis roots. *Phytochem. Lett.* 22, 44–48. doi: 10.1016/j.phytol.2017.09.001

Nam, T. G., Lim, Y. J., and Eom, S. H. (2018). Flavonoid accumulation in common buckwheat (*Fagopyrum esculentum*) sprout tissues in response to light. *Hortic. Environ. Biotechnol.* 59, 19–27. doi: 10.1007/s13580-018-0003-5

Neugart, S., Majer, P., Schreiner, M., and Hideg, É. (2021). Blue light treatment but not green light treatment after pre-exposure to UV-b stabilizes flavonoid glycoside changes and corresponding biological effects in three different brassicaceae sprouts. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.611247

Park, Y., Kwon, S. H., Jang, Y. L., Lee, D. H., Yang, S. O., Eo, H. J., et al. (2022). Nutritional composition and phytochemical screening in different parts of *Hibiscus syriacus* l. *Food Sci. Nutr.* 10, 3034–3042. doi: 10.1002/fsn3.2899

Qin, W. T., Feng, Y. R., Lei, Z., Yang, C. Q., Wu, H. J., Nasir, I., et al. (2017). Effect of shading signal on isoflavone biosynthesis of soybean seedling. *Nat. Prod. Res. Dev.* 29, 1470. doi: 10.16333/j.1001-6880.2017.9.003

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

Seo, J. M., Arasu, M. V., Kim, Y. B., Park, S. U., and Kim, S. J. (2015). Phenylalanine and LED lights enhance phenolic compound production in tartary buckwheat sprouts. *Food Chem.* 177, 204–213. doi: 10.1016/j.foodchem.2014.12.094

Shah, A., and Smith, D. L. (2020). Flavonoids in agriculture: Chemistry and roles in, biotic and abiotic stress responses, and microbial associations. *Agronomy* 10, 1209. doi: 10.3390/agronomy10081209

Shu, X. L., Frank, T., Shu, Q. Y., and Engel, K. H. (2008). Metabolite profiling of germinating rice seeds. *J. Agric. Food Chem.* 56, 11612–11620. doi: 10.1021/jf802671p

Siipola, S. M., Kotilainen, T., Sipari, N., Morales, L. O., Lindfors, A. V., Robson, T. M., et al. (2015). Epidermal UV-a absorbance and whole-leaf flavonoid composition in pea respond more to solar blue light than to solar UV radiation. *Plant Cell Environ.* 38, 941–952. doi: 10.1111/pce.12403

Sohn, S. I., Pandian, S., Oh, Y. J., Kang, H. J., Cho, W. S., and Cho, Y. S. (2021). Metabolic engineering of isoflavones: an updated overview. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.670103

Soliman, F. M., Shehata, A. H., Khaleel, A. E., and Ezzat, S. M. (2002). An acylated kaempferol glycoside from flowers of *Foeniculum vulgare* and *F. dulce*. *Molecules* 7, 245–251. doi: 10.3390/70200245

Sugiyama, A., Yamazaki, Y., Hamamoto, S., Takase, H., and Yazaki, K. (2017). Synthesis and secretion of isoflavones by field-grown soybean. *Plant Cell Physiol.* 58, 1594–1600. doi: 10.1093/pcp/pcx084

Taulavuori, K., Pyysalo, A., Taulavuori, E., and Julkunen-Tiitto, R. (2018). Responses of phenolic acid and flavonoid synthesis to blue and blue-violet light depends on plant species. *Environ. Exp. Bot.* 150, 183–187. doi: 10.1016/j.envexpbot.2018.03.016

Thoma, F., Somborn-Schulz, A., Schlehuber, D., Keuter, V., and Deerberg, G. (2020). Effects of light on secondary metabolites in selected leafy greens: A review. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00497

Vacek, J., Klejdus, B., Lojková, L., and Kubán, V. (2008). Current trends in isolation, separation, determination and identification of isoflavones: A review. *J. Sep. Sci.* 31, 2054–2067. doi: 10.1002/jssc.200700569

Wade, H. K., Bibikova, T. N., Valentine, W. J., and Jenkins, G. I. (2001). Interactions within a network of phytochrome, cryptochrome and UV-b phototransduction pathways regulate chalcone synthase gene expression in arabidopsis leaf tissue. *Plant J.* 25, 675–685. doi: 10.1046/j.1365-313x.2001.01001.x

Wang, J., Fang, X., Ge, L., Cao, F., Zhao, L., Wang, Z., et al. (2018). Antitumor, antioxidant and anti-inflammatory activities of kaempferol and its corresponding glycosides and the enzymatic preparation of kaempferol. *PloS One* 13, e0197563. doi: 10.1371/journal.pone.0197563

Wang, S. Y., Zhang, Y. J., Zhu, G. Y., Shi, X. C., Chen, X., Herrera-Balandrano, D. D., et al. (2022). Occurrence of isoflavones in soybean sprouts and strategies to enhance their content: A review. *J. Food Sci.* 87, 1961–1982. doi: 10.1111/1750-3841.16131

Wei, P., Liu, M., Chen, Y., and Chen, D. C. (2012). Systematic review of soy isoflavone supplements on osteoporosis in women. *Asian Pac. J. Trop. Med.* 5, 243–248. doi: 10.1016/S1995-7645(12)60033-9

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 For.* 45, 765–776. doi: 10.1007/s11056-014-9435-7

Yang, W. T., Cho, K. M., and Lee, J. H. (2020). Comparative analysis of isoflavone aglycones using microwave-assisted acid hydrolysis from soybean organs at different growth times and screening for their digestive enzyme inhibition and antioxidant properties. *Food Chem.* 305, 125462. doi: 10.1016/j.foodchem.2019.125462

Zhang, X., Bian, Z., Yuan, X., Chen, X., and Lu, C. (2020). A review on the effects of light-emitting diode (LED) light on the nutrients of sprouts and microgreens. *Trends Food Sci. Technol.* 99, 203–216. doi: 10.1016/j.tifs.2020.02.031

Zhang, Y., Jiang, L., Li, Y., Chen, Q., Ye, Y., Zhang, Y., et al. (2018). Effect of red and blue light on anthocyanin accumulation and differential gene expression in strawberry (*Fragaria× ananassa*). *Molecules* 23, 820. doi: 10.3390/molecules23040820

Zoratti, L., Karppinen, K., Luengo Escobar, A., Häggman, H., and Jaakola, L. (2014). Light-controlled flavonoid biosynthesis in fruits. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00534

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 03 March 2023
DOI 10.3389/fpls.2023.1139274

OPEN ACCESS

EDITED BY
Feng Xu,
Yangtze University, China

REVIEWED BY
Peter Lindemann,
Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany
Juan Hua,
Shenyang Agricultural University, China

*CORRESPONDENCE
Petra M. Bleeker
p.m.bleeker@uva.nl

SPECIALTY SECTION
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

RECEIVED 06 January 2023
ACCEPTED 01 February 2023
PUBLISHED 03 March 2023

CITATION
Kortbeek RWJ, Galland MD, Muras A, Therezan R, Maia S, Haring MA, Schuurink RC and Bleeker PM (2023) Genetic and physiological requirements for high-level sesquiterpene-production in tomato glandular trichomes. *Front. Plant Sci.* 14:1139274. doi: 10.3389/fpls.2023.1139274

Genetic and physiological requirements for high-level sesquiterpene-production in tomato glandular trichomes

Ruy W. J. Kortbeek, Marc D. Galland, Aleksandra Muras, Rodrigo Therezan, Sofia Maia, Michel A. Haring, Robert C. Schuurink and Petra M. Bleeker*

Green Life Science Research Cluster, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, Netherlands

Type-VI glandular trichomes of wild tomato *Solanum habrochaites* PI127826 produce high levels of the sesquiterpene 7-epizingiberene and its derivatives, making the plant repellent and toxic to several pest insects and pathogens. How wild tomato trichomes achieve such high terpene production is still largely unknown. Here we show that a cross (F1) with a cultivated tomato produced only minute levels of 7-epizingiberene. In the F2-progeny, selected for the presence of the 7-epizingiberene biosynthesis genes, only three percent produced comparable amounts the wild parent, indicating this trait is recessive and multigenic. Moreover, trichome density alone did not explain the total levels of terpene levels found on the leaves. We selected F2 plants with the "high-production active-trichome phenotype" of PI127826, having trichomes producing about 150 times higher levels of terpenes than F2 individuals that displayed a "low-production lazy-trichome phenotype". Terpene quantities in trichomes of these F2 plants correlated with the volume of the storage cavity and shape of the gland. We found that trichome morphology is not a predetermined characteristic, but cavity volume rather depended on gland-cell metabolic activity. Inhibitor assays showed that the plastidial-precursor pathway (MEP) is fundamental for high-level production of both cytosolic as well as plastid-derived terpenes in tomato trichomes. Additionally, gene expression profiles of isolated secretory cells showed that key enzymes in the MEP pathway were higher expressed in active trichomes. We conclude that the MEP pathway is the primary precursor-supply route in wild tomato type-VI trichomes and that the high-production phenotype of the wild tomato trichome is indeed a multigenic trait.

KEYWORDS

tomato, glandular trichome, sesquiterpenes, MEP pathway, metabolic activity, *Solanum habrochaites*, specialised metabolism

Introduction

Trichomes are hair-like structures present on the aerial parts of approximately 30% of all land plants and they constitute an important part of the plant's defence system against herbivorous insects. Non-glandular trichomes form a physical barrier e.g. by impeding movement over the plant's surface, perturbating feeding behaviour or disruption of the insect digestive system (Kariyat et al., 2017; Andama et al., 2020). Glandular trichomes can produce, store and/or secrete a large number of volatile-and non-volatile specialised metabolites. These metabolites act as defence compounds by being repellent, causing toxicity, interfere with insect development, entrap them, or attract predatory insects (Glas et al., 2012; Kortbeek et al., 2019; Schuurink and Tissier, 2020).

The tomato genus displays seven different trichome types; non-glandular type II, III, V and glandular types I, IV, VI and VII (Luckwill, 1943). The presence and distribution of these depend on the tomato species, accession and tissue, and can be influenced by environmental conditions or stresses (Vendemiatti et al., 2017; Chen et al., 2018; Escobar-Bravo et al., 2018). Some accessions of the wild tomato *Solanum habrochaites*, including PI127826 and LA1777, are known to produce high amounts of specific sesquiterpenoids in their type-VI trichomes making them resistant against multiple insects and pathogens (Fridman et al., 2005; Sallaud et al., 2009; Schilmiller et al., 2009; Bleeker et al., 2011a; Akhtar et al., 2013; Nakashima et al., 2016). This trichome type consists of four secretory cells on top of a multicellular stalk and they can be found on the green parts of most tomato species, including cultivated tomato (*S. lycopersicum*; Glas et al., 2012).

In both *S. lycopersicum* and *S. habrochaites*, sesquiterpenes are synthesised in the cytosol from acetyl-CoA *via* the mevalonate (MVA) pathway (Hemmerlin et al., 2012) with some enzymes localised to the peroxisomes. Here, acetyl-CoA is first metabolised to the common C_5 isoprenoid precursors isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP). Next, prenyltransferase farnesyl diphosphate synthase (FPS) catalyses the head-to-tail condensation of IPP/DMAPP to form the C_{15} precursor farnesyl diphosphate (FPP). In cultivated tomato, FPP is mainly utilised by the sesquiterpene synthase *Sl*TPS12 to form α-humulene and β-caryophyllene while *S. habrochaites* additionally produces a mix of germacrenes *via Sh*TPS9 (Schilmiller et al., 2010; Bleeker et al., 2011b). Both species also generate IPP/DMAPP from pyruvate and glyceraldehyde 3-phosphate (G3P) *via* the plastidial 2-C-methyl-D-erythritol 4-phosphate pathway (MEP). In trichomes of cultivated tomato, IPP/DMAPP are condensated by a *cis*-prenyltransferase, neryl-diphosphate synthase 1 (NDPS1), to generate the C_{10} monoterpene-precursor neryl-diphosphate (NPP; Schilmiller et al., 2009). NPP is utilised predominantly by phellandrene synthase 1 (PHS1, also referred to as TPS20) generating a mix of monoterpenes (Schilmiller et al., 2009). In the plastids of *S. habrochaites* accessions the *cis*-allelic variant of NDPS1, *Z*-isoprenyl pyrophosphate synthase (zFPS), condensates IPP/DMAPP to the C_{15} sesquiterpene precursor *Z,Z*-farnesyl pyrophosphate (zFPP; Sallaud et al., 2009; Kang et al., 2014) which is utilised by *S. habrochaites*-specific sesquiterpene terpene synthases. In PI127826 the 7-epizingiberene synthase (ZIS) generates 7-epizingiberene (7epiZ) and *R*-curcumene (Bleeker et al., 2012) while in LA1777 the santale/bergamotene synthase (SBS) forms a mix of santalenes and bergamotenes (Sallaud et al., 2009). In some accessions, these sesquiterpenes are further oxidised by which they gain toxicity to insects and/or pathogens (Frelichowski and Juvik, 2001; Zabel et al., 2021).

In type-VI trichomes, terpenes are stored inside an intercellular storage-cavity, located central to the four glandular cells (Tissier et al., 2017). Wild tomato genotypes like *S. habrochaites* often have larger storage cavities compared to trichomes of cultivated tomatoes, and this size difference seems to be positively correlated to quantity of specialised metabolites inside the trichome (Bergau et al., 2015). Furthermore, type-VI trichomes can differ in morphology. The shape of the glandular heads on cultivated tomato are often described as "lobed" or "clover shaped", referring to the four secretory cells which are distinctly visible. In contrast, the glandular heads of some *S. habrochaites* accessions, e.g. LA1777 or PI126449, are referred to as "globular" or "spherical", in which the individual secretory cells are not distinguishable (Ben-Israel et al., 2009; Bergau et al., 2015).

Despite the differences in biochemistry and morphology between *S. lycopersicum* and *S. habrochaites* trichomes, it is possible to transfer terpenoid production from a wild species to a cultivar as both species have active MVA and MEP pathways providing the universal precursor IPP/DMAPP. Indeed, we previously showed that that transgenic cultivated tomato can produce sesquiterpene 7epiZ after introducing two genes only; i.e. *zFPS* and *ZIS* from PI127826 (Bleeker et al., 2012). Similarly, we showed that santalenes/bergamotenes can be produced by cultivated tomato trichomes with the introgression of the *zFPS* and *SBS* locus (*Sst2*) from LA1777 to a *S. lycopersicum* cultivar (cv), MicroTom (Therezan et al., 2021). However, in all these cases terpene quantities remained relatively low compared to the wild species. The transgenic *zFPS/ZIS* tomato plants, for example, produced ~200 times less 7epiZ compared to PI127826 (Bleeker et al., 2012). This implies that besides the prenyl-transferases and terpene synthases, additional factors are required to achieve high-level production of these terpenes. The previously reported 7epiZ-oxidase (*Sh*ZO) of *S. habrochaites* PI127826 is responsible for the oxidation of 7epiZ to 9-hydroxy-zingiberene (9HZ) and 9-hydroxy-10,11-epoxyzingiberene (9H10epoZ), metabolites that are toxic to whiteflies and several microbes in a dose-dependent manner (Zabel et al., 2021). When produced in sufficient quantities, these metabolites could protect the pest-sensitive cultivated tomato, acting as natural defence compounds.

In this paper, we aimed to study the genetic basis of the high-terpene production and storage capacity of wild tomato type-VI trichomes. We generated an interspecific F2 population of a cultivated tomato and *S. habrochaites* PI127826 and showed that it requires multiple recessive loci for an individual trichome to produce 7epiZ in high concentrations, a trait independent of having a high density of glandular trichomes. We selected individuals with contrasting terpene-production phenotypes while being

homozygous for the *S. habrochaites* alleles required for the production of 7epiZ and derivatives (i.e. *zFPS*, *ZIS* and *ShZO*). This material enabled us to conclude the capacity to store high terpene requires multiple recessive loci. The large storage volume, and consequent morphology, of *S. habrochaites* type VI-trichomes is the result of a high-level MEP pathway activity.

Results

The genetic background of a cultivated tomato impedes 7-epizingiberene biosynthesis

To study the production of 7epiZ and its derivatives 9HZ and 9H10epoZ in the genetic background of a cultivated tomato, the leaf surface of an interspecific F1 cross between PI127826 (male) and a cultivated tomato (female) were analysed for trichome densities and terpene abundance. Leaf washes of PI127826 revealed a total of 176 ng mg^{-1} FW terpenes with 7epiZ as the most abundant terpene (70 ng mg^{-1}), along with its derivatives 9HZ (12 ng mg^{-1}) and 9H10epoZ (10 ng mg^{-1}; Figure 1A; Supplemental Figure S1). As expected, the quantity of terpenes found in leaf washes of the cultivar were much lower with a total of 1.6 ng mg^{-1} FW, of which monoterpene β-phellandrene was the most abundant (0.5 ng mg^{-1} FW). Surprisingly, the F1 plants also exhibited much lower terpene levels compared to the wild parent. Total terpene levels in the F1 were 12.3 ng mg^{-1} FW, of which 1.1 ng mg^{-1} was 7epiZ, 0.2 ng mg^{-1} 9HZ, 0.8 ng mg^{-1} 9H10epoZ and 2.5 ng mg^{-1} β-phellandrene (Figure 1A; Supplemental Figure S1). Interestingly, when PI127826 was crossed to *S. habrochaites* accession LA1777, resulting in an *S. habrochaites* F1 (F1-hab), 7epiZ levels of the hybrid exceeded to those of PI127826 (Figure 1A). Total terpene content in F1-hab leaf extracts also exceeded those of both parents, mostly through the additional production of bergamotenes and santalenes (Supplemental Figure S1).

The quantity of trichome-derived mono-and sesquiterpenes on leaf surface is a result of the total number of the terpene-producing type-VI trichomes multiplied by the production per individual trichome. To explain the differences in terpene levels found, we counted type-VI trichomes on both the adaxial and abaxial leaf surface (Figure 1C). Additionally, we recorded acylsugar-producing glandular type-I and IV (taken together as type-I/IV) and non-glandular (NG) trichomes (Supplemental Figure S2). Type-VI trichome density of PI127826 appeared to be ~2 times higher compared to the cultivar, whereas the density on the interspecific

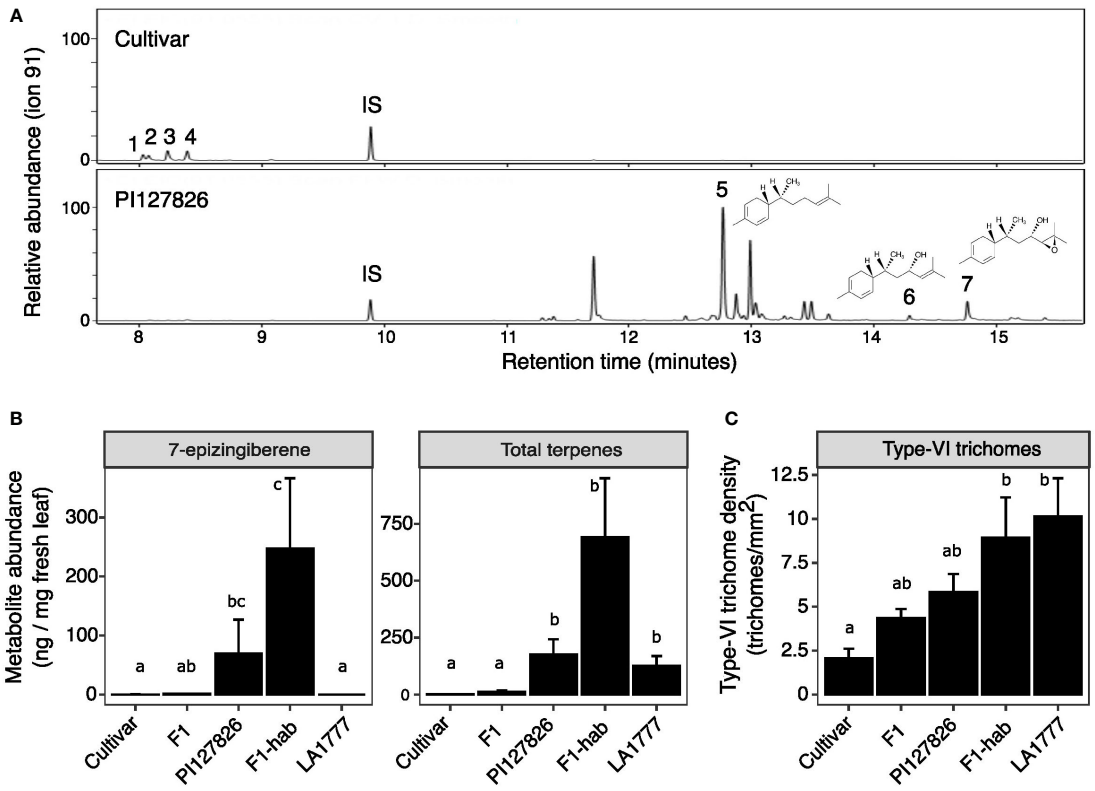


FIGURE 1
Metabolite levels and trichome density on leaves of tomato hybrids and their parents. **(A)** GC-MS chromatograms of the *S. lycopersicum* cultivar and *S. habrochaites* PI127826. Peaks: (1) 2-carene; (2) α-phellandrene; (3) terpinolene; (4) β-phellandrene/D-limonene; (5) 7-epizingiberene; (6) 9-hydroxy-zingiberene; (7) 9-hydroxy-10,11-epoxyzingiberene; (IS) internal standard. Molecular structures are given for 5, 6 and 7. **(B)** 7-epizingiberene and total mono-and sesquiterpene abundance on the leaves of the cultivar x PI127826 F1 hybrid (F1), the *S. habrochaites* hybrid PI127826 x LA1777 (F1-hab) and their parents. Bars represent mean level of metabolites per mg fresh leaf ± SE (n = 3). **(C)** Type-VI trichomes density on the leaf. Bars represent mean ± SE (n = 3) summed densities over the abaxial and adaxial surface. Statistics were performed after log transformation; letters indicate significant groups ($p < 0.05$) according to a Tukey HSD *post-hoc* test after ANOVA.

F1 exhibited an intermediate phenotype (Figure 1B). This intermediate phenotype of the F1 was also observed for type-I/IV and NG trichome density. Type-I/IV trichomes were typically highly abundant in PI127826 but absent on leaves of the cultivar. On the contrary, NG trichome density was low on PI127826 and high on the cultivar (Supplemental Figure S2). In addition, F1-hab also displayed an intermediate type-VI trichome-density phenotype between both *S. habrochaites* parents PI127826 and LA1777 (Figure 1C). The difference in type-VI trichome density did not fully explain the difference in terpene phenotype. Terpenes levels in the F1 were much lower than expected from the inherited genetic dose of the wild parent.

High-level terpene production comprises a multigenic trait

To investigate the genetics behind the quantitative traits involving terpene production by type-VI trichomes, we created an F2 population of the cultivar and PI127826 by selfing the F1. Next, a subpopulation was selected by screening the F2 progeny for homozygous alleles of *ShzFPS*, *ShZIS* (Solyc08g005640; Solyc08g005680) and *ShZO* (Solyc01g008670) to fix the synthesis of 7epiZ and its derivatives. This subpopulation (n = 392) was screened for mono-and sesquiterpene abundance on the leaf and 7epiZ, being the most abundant terpene, was taken as a proxy. The distribution of 7epiZ levels measured in the population were not normally distributed but instead appeared highly skewed towards low quantities (Figure 2A). In fact, only 84 F2 plants had notable levels of 7epiZ, of which 12 displayed parental PI127826 levels or higher. Presuming a Mendelian inheritance, the 12:380 (1:32) segregation of the "high terpene level" phenotype indicates that multiple unlinked loci from the wild parent required. Testing different genetic models, the combination of two recessive and two dominant loci from PI127826 best fitted the observed segregation pattern (Table 1). When taking the production of 7epiZ derivatives into account as well, this model again fitted our observations best (Supplemental Table S1).

To assess the contribution of trichome density to the 7epiZ levels, we counted the type-VI trichomes on leaf discs taken from the F2 plants and the F2 genotypes were grouped into 10 classes according to the summed number of trichomes on both sides (Supplemental Figure S3A). Next, we compared the trichome densities to 7epiZ levels using the 84 F2-genotypes with notable 7epiZ levels. Surprisingly, mean 7epiZ levels did not significantly differ between trichome-density classes (ANOVA; $p > 0.05$; Supplemental Figure S3B). Moreover, the 12 F2 genotypes with parental PI127826 7epiZ levels were represented in five of the 10 trichome-density classes, ranging from the class with a low number of trichomes per leaf disc to the class with the highest number of trichomes (Supplemental Figure S3B). We selected a subset of F2-plants originating from each trichome-density class (Supplemental Figure S3C) and with high and low levels of 7epiZ to investigate the relationship between trichome density and the abundance of terpenes in closer detail. Hereto we quantified and summed the mono- and sesquiterpenes on the leaves and determined the trichome density per mm^2 of the leaf. This confirmed that that, in this subset of plants, there is no significant linear correlation between the type-VI trichome-density and the quantity of volatile terpenes per mg of fresh leaf (Figure 2B; $r^2 = 0.015$, $p = 0.25$).

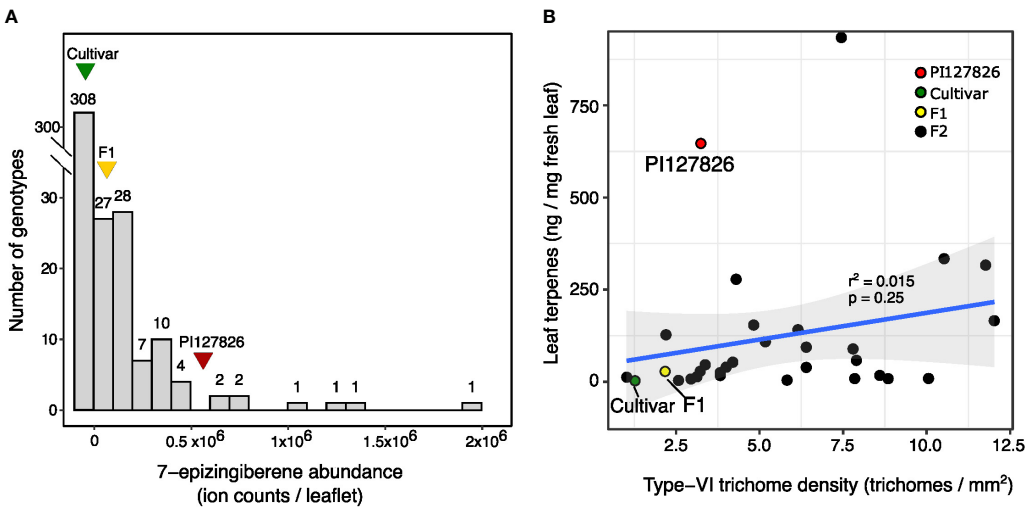


FIGURE 2
Distribution of 7-epizingberene levels and trichome density of the F2 population. **(A)** Histogram shows the distribution of 7-epizingiberene levels on leaflets of 393 F2-genotypes selected for the presence of *zFPS*, *ZIS*, and *ShZO*. F2 plants are assigned to bins of ion counts of the 7-epizingiberene peak as determined by GC-MS. The width of each bin represents one time the standard deviation of 7-epizingiberene in PI127826 (n =4), with the first bin displaying the number of F2 genotypes with zero ion counts. The number on top of each bar displays the number of F2 genotypes assigned to the particular bin. Triangles indicate the bin with the average 7-epizingiberene levels of the respective parental genotypes (n = 4). **(B)** Scatterplot showing the type-VI trichome density of the sub-population (averaged abaxial/adaxial) versus the total terpene levels (i.e. summed mono-and sesquiterpenes) on the leaves of F2 plants from the sub-population. The blue line indicates the linear correlation plus 95% confidence interval (grey). The correlation coefficient (r^2) and p-value are displayed on top of the line. Each dot represents measurements of an individual plant.

TABLE 1 Comparison of observed 7-epizingiberene segregation to theoretical models.

Gene model	Segregation	χ^2	p-value
a - b	1: 15	6.80	< 0.01
a - b - C	1: 20	2.50	0.11
a - b - c	1: 63	2.54	0.11
a - b - C - D	1: 27	1.15	0.28
a - b - c - D	1: 84	6.46	< 0.05
a -b -c - d	1: 255	47.35	< 0.01

Gene models consist of recessive (small letters) and dominant (capital letters) loci. The theoretical segregation pattern of the phenotype based on the corresponding gene model is given under Segregation. Models were compared to the observed low:high 7-epizingiberene segregation of 1:32 using the Chi-square test for goodness-of-fit (χ^2).

The results indicate that there is a large variation in volatile production by type-VI trichomes between F2-genotypes which explains why high 7epiZ levels can be found on leaflets with low trichome densities and vice versa. It moreover suggests that trichome density and terpene production by type-VI trichomes segregate independently and should be regarded as genetically separate traits.

Gland-cell metabolic activity: Active and lazy trichomes

To investigate the accumulated quantity of terpenes by type-VI trichomes, the glandular heads ("glands") were isolated from the subset of F2-plants and analysed by GC-MS. Type-VI glands of the cultivar had an averaged total terpene content of 0.51 ng/gland, mainly consisting of plastidial monoterpenes and to a lesser extend cytosolic sesquiterpenes. Terpene levels of PI127826 type-VI glands were ~35 times higher with 17.7 ng/gland, while levels in F1-glands were ~1.8 times higher (0.9 ng/gland) compared to the cultivar. A broad range of terpene levels was found in glands of the selected F2 progeny ranging from 0.05 ng/gland up to 21.5 ng/gland (Figure 3A).

From the available material, we selected six F2-individuals based on their extreme terpene levels in the type-VI glands. Three F2-genotypes exhibited relatively high terpene quantities per gland, hence we refer to them as having "active" trichomes. Another three F2-genotypes were selected for their low levels of terpenes per gland, referred to as having a "lazy" trichome phenotype (Figure 3B, upper panel). Light microscopy allowed visualisation of the internal storage cavity of their type-VI trichomes making the different phenotypes clearly visible (Figure 3C) and made it possible to calculate the volume of the storage cavities. The storage-cavity volume on the leaves of lazy F2 genotypes ranged from 0.7 picolitre (F2-455) to 4 picolitre (F2-151). Storage-cavity volumes of the active F2 plants ranged from 14 picolitre (F2-28) up to 22 picolitre (F2-127). The cavity volumes of the glands (Figure 3B, lower panel) largely mirrored their terpene content, where it appears that larger cavity volumes corresponded to larger quantities of terpenes.

It is furthermore worthwhile to note that we did not observe large variations in gland volume of type-VI trichomes within a genotype; a single F2 individual had either all active or all lazy trichomes, but never a mixture. Additionally, the cavity volume seems to be correlated to the shape of the glands. Active genotypes with large cavities had smooth and spherical-shaped glands while on lazy genotypes the contours of the individual cells became visible and therefore the glands appeared to be more lobed (Supplemental Figure S4).

Inhibition of terpene metabolic flux

To investigate whether cavity volume is a fixed trait that facilitates terpene accumulation, or rather is a consequence of terpenoid accumulation, we inhibited the flux through the plastidial precursor pathway. Cuttings of PI127826 were treated with 10 µM fosmidomycin, inhibiting 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) activity (Zeidler et al., 1998). After 14 days of treatment, the type-VI trichome head-cells from newly developed leaves were analysed for terpene content and storage-cavity volume. In control plants, terpene levels in PI127826 glands measured 12.3 ng/gland, which was reduced with 98% to 0.21 ng/gland in fosmidomycin-treated cuttings (Figure 4A). Inhibition of the MEP-pathway did not only affect the accumulation of terpenes produced in the plastid, but also sesquiterpenes produced by cytosol-localised terpene synthases (Supplemental Figure. S5). In addition, the treatment reduced the cavity volume of the glands with 95% from 25.0 picolitre in control plants to 1.2 picolitre in treated cuttings (Figures 4B, C). The change in cavity volume was accompanied by a difference in physical appearance of the glands. Instead of having smooth and spherical-shaped glands, newly formed glands of fosmidomycin-treated cuttings appeared more lobed (Figure 4C) as were the glands of the above mentioned lazy F2-genotypes. Interestingly, type-VI trichomes on leaves that were already present on the cuttings prior to the treatment were much less affected in cavity size and shape; fosmidomycin reduced the cavities on these trichomes with only 16% (Supplemental Figure S6).

Following up, we treated both active and lazy F2-genotypes with selective inhibitors to examine if the MEP and the MVA pathway contribution differently to terpene accumulation depending on the genotype. Cuttings of PI127826 and F2-73 ("active"), the cultivar and F2-411 ("lazy") were treated with 10 µM fosmidomycin to

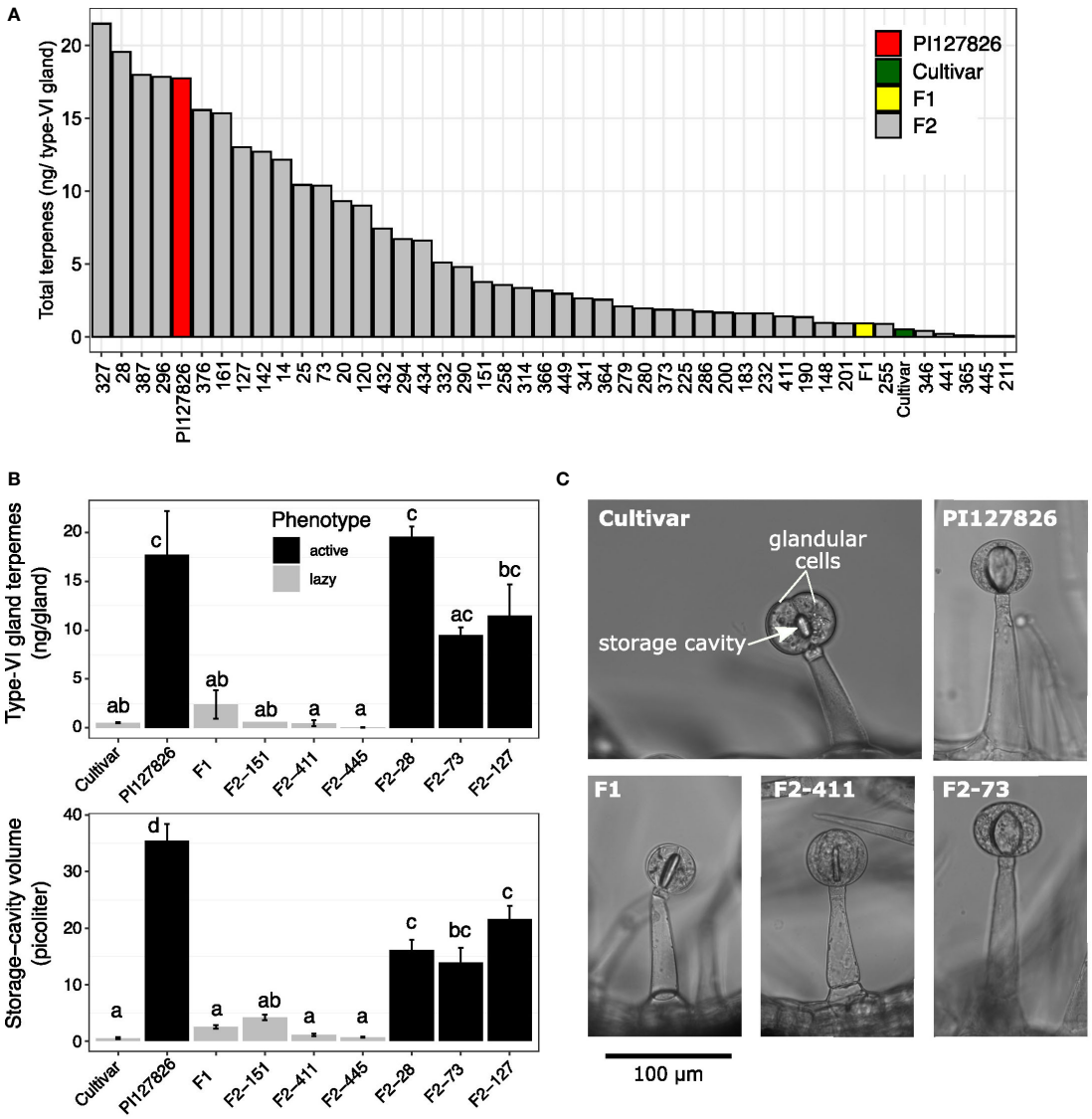


FIGURE 3
Volatile levels in relation to type-VI trichome densities of the F2 progeny. **(A)** Total amount of mono-and sesquiterpenes per type-VI glandular head of the selected F2-individuals (n = 1) and their parents. **(B)** Representative photographs showing the middle section of type-VI trichomes on the parental genotypes and an active (F2-73) and lazy (F2-411) F2 individuals. Visible are two glandular cells surrounding the central-storage cavity. **(C)** Quantification of the storage-cavity volume (upper panel) of type-VI trichomes (n = 10-20) and total mono-and sesquiterpenes per gland (lower panel) of the selected active (black bars) and lazy (grey bars) F2-genotypes and the parents (n = 3). Letters indicate significant groups ($p < 0.05$) according to Tukey HSD *post-hoc* test after ANOVA.

block the MEP pathway, or 10 µM mevastatin to block the MVA pathway by inhibition of HMG-CoA Reductase (HMGR). After 14 days of treatment, type-VI trichome-heads were isolated from newly formed leaves and their terpene content was analysed.

As observed above, fosmidomycin treatment greatly reduced both plastid- and cytosol-derived terpenes in the glands of active genotypes (Figure 5A). In this experiment, the levels of plastid-derived 7epiZ and its derivatives were for 98% reduced in PI127826 glands. Gland cells of F2-73 also displayed a reduction of 98%; from 5.5 ng/gland in control conditions to 0.12 ng/gland after treatment. It was furthermore confirmed that blocking of the MEP pathway reduces cytosolic terpenes (mainly β-caryophyllene, α-humulene and germacrenes) as well. Those terpenes were for 99% depleted in PI127826 (5.2 ng/gland in control conditions versus 0.065 ng/gland upon treatment) and were completely absent in glands of F2-73 (from 1.4 ng/gland control conditions versus undetectable levels upon treatent). In comparison to fosmidomycin, the effect of mevastatin on terpenoid production in "active" glands was less pronounced. Though both plastidial and cytosolic terpenes appeared less abundant upon inhibition of the MVA pathway, these differences were not significant (Figure 5A). Again, and especially apparent in the active genotypes, changes in terpene levels coincided with a reduction in storage-cavity volume, particularly apparent in active F2-73 (Figures 5B, C). This F2 exhibited a 60% reduction in volume under mevastatin-treatment compared to control (Figure 5B) indicating that also other, MVA-

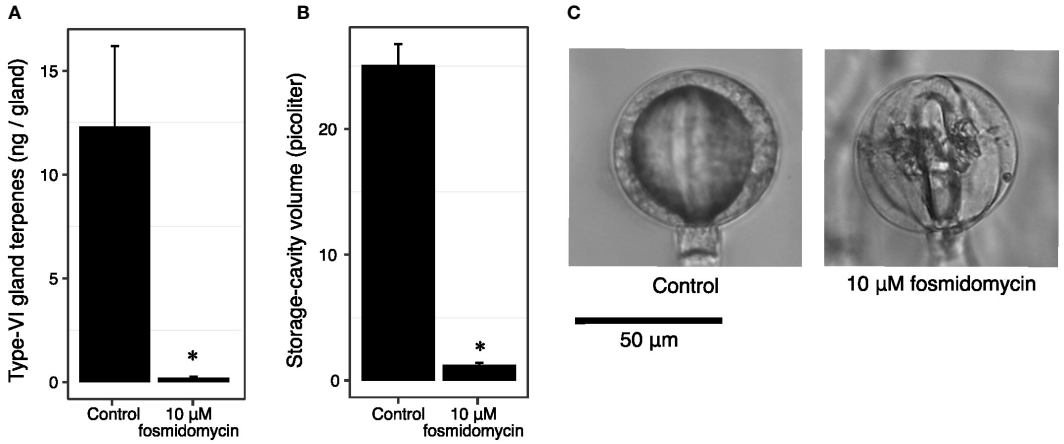


FIGURE 4
Fosmidomycin treatment of PI127826 cuttings reduces terpene content and storage cavity volume of type-VI trichomes. Cuttings of PI127826 (n = 3) were grown for 14 days in hydroponic solution supplied with or without 10 µM fosmidomycin. **(A)** Quantification of total mono-and sesquiterpene terpene levels in type-VI glands. **(B)** Volume of the storage cavity. **(C)** Representative photographs of a type-VI head displaying the central storage cavity after 14 days under control conditions or with fosmidomycin treatment. Significant differences are annotated with an asterisk: T-test $p < 0.01$.

pathway dependent, products that were not analysed here, are present in the storage cavity.

Terpene levels of "lazy" genotypes F2-411 and the cultivar, were overall less affected by inhibitor treatment. In general, their trichomes contained much lower terpene levels compared to the active genotypes. Fosmidomycin reduced the levels of plastidial terpenes with 36% in the cultivar (from 0.13 to 0.083 ng/gland) and with 63% in F2-411 (0.16 to 0.060 ng/gland; Figure 5A). These reductions were however not statistically significant and the cavity volumes did not significantly alter under inhibition of the MEP pathway (Figure 5B). As for the "active" genotypes, no significant differences were observed in the abundance of either plastidial-or cytosolically synthesised terpenes of "lazy" genotypes treated with MVA-inhibitor mevastatin (Figure 5A).

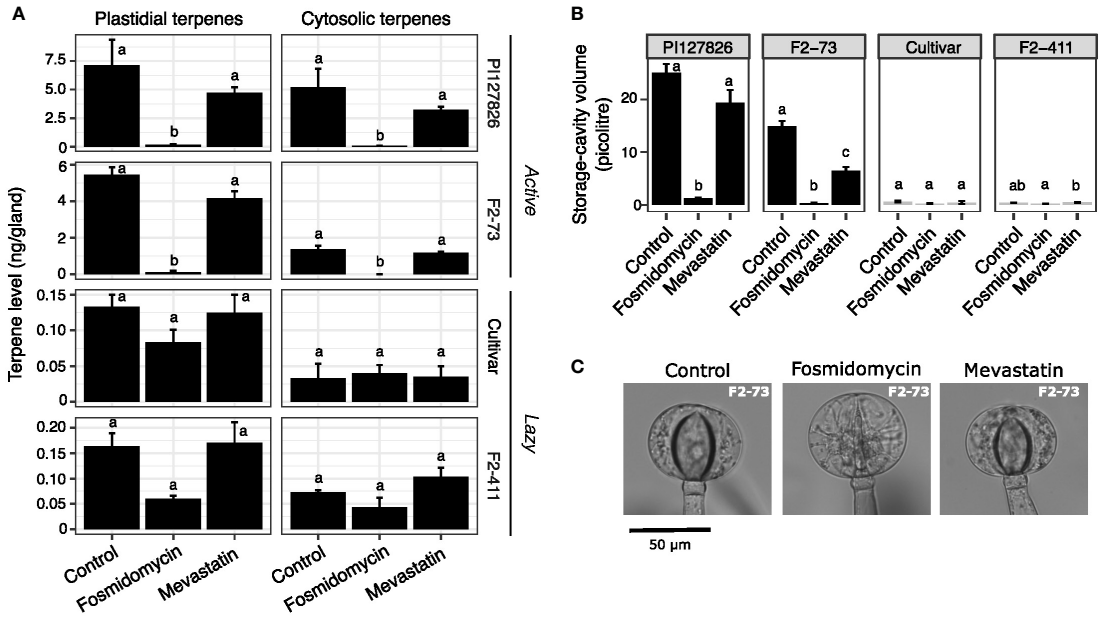


FIGURE 5
The effect of MEP/MVA pathway inhibition on plastidial-and cytosolically produced terpenes and the storage cavities. Cuttings (n = 3) of active (PI127826 and F2-73) and lazy (cultivar and F2-411) genotypes were grown for 14 days in hydroponic solution supplied with 10 µM fosmidomycin or 10 µM mevastatin. **(A)** Terpene levels in the type-VI trichome glands after 14 days of treatment. Terpene levels are summed according to the sub-cellular localisation (i.e. plastidial of cytosolic) of the corresponding terpene synthase (Zhou et al., 2020). The summed terpene quantities were Log_2-transformed prior to ANOVA. Letters indicate significant groups ($p < 0.05$) after a Tukey HSD *post-hoc* test. **(B)** Volume of the storage cavity after 14 days of treatment. Letters indicate significant groups ($p < 0.05$) according to Tukey HSD *post-hoc* test after ANOVA of Log_2-transformed cavity volumes **(C)** Representative microscope images type-VI trichome glands on the leaves of F2-73 after 14 days of treatment.

10.3389/fpls.2023.1139274

MEP pathway upregulated in gland secretory cells

To investigate the transcript levels of the MVA- and MEP-pathway genes, we isolated the glands of type-VI trichomes from the leaves of the lazy cultivar and active F2 genotype F2-28 (Figure 3B) and compared gene expression by mRNA-sequencing of enzymes in the MVA and MEP precursor pathways (Figure 6A). The gland-isolation procedure involved a sieving step to enrich for intact type-VI glands resulting in only minimal contamination of

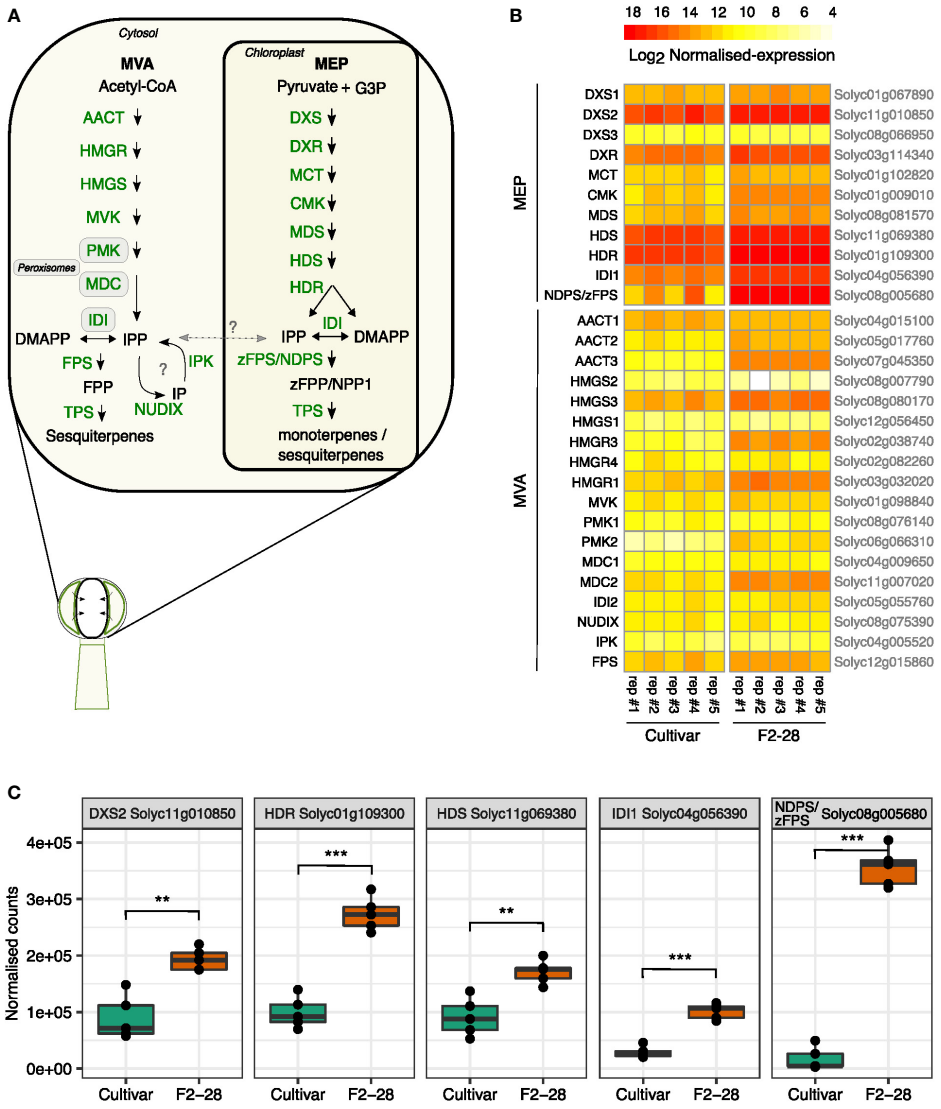


FIGURE 6
Gene expression patterns of the MEP and MVA pathway in secretory cells of F2-28 and the cultivar. **(A)** A schematic overview of the MEP and MVA pathway for mono-and sesquiterpene synthesis in the secretory cells of type-VI trichomes of tomato. Each arrow arrows indicates one enzymatic step. Metabolites are indicated in black, enzymes in green. Question marks indicate hypothetical steps that are not validated in tomato trichomes. **(B)** Heatmap showing the normalised counts (Log_2 transformed) of transcripts obtained by mRNA-sequencing of secretory cells of type-VI trichomes isolated from the leaves of the cultivar (n = 5) and F2-28 (n = 5) cuttings. Plotted are transcripts mapping to the different homologous enzymes of the MEP and MVA pathway as encoded by genes in the *S. lycopersicum* genome (Heinz). Gene annotations are given on the left side of the heatmap, gene annotations are given on the right. **(C)** Boxplots of the normalised counts of the highest expressed genes in the MEP pathway of the cultivar and F2-28. Each dot represents one biological replicate. Asterisks show significant differences in expression (T-test) with ** $p < 0.01$ and *** $p < 0.001$. Enzyme abbreviations: 1-deoxy-d-xylulose 5-phosphate synthase (DXS); 1-deoxy-d-xylulose 5-phosphate reductoisomerase (DXR); 2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase (MCT); 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol kinase (CMK); 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase (MDS); (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (HDS); (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase; (HDR); isopentenyl diphosphate isomerase (IDI); *Z*-*Z*-farnesyl diphosphate synthase (zFPS); neryl diphosphate synthase (NDPS); terpene synthase (TPS); acetyl-CoA C-acetyltransferase (AACT); 3-hydroxy-3-methylglutaryl-CoA synthase (HMGS); 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR); mevalonate kinase (MVK); phospho-mevalonate kinase (PMK); mevalonate-5-diphosphate decarboxylase (MDC); isopentenyl phosphate kinases (IPK); nudix hydrolase (NUDIX); farnesyl diphosphate synthase (FPS). Metabolite abbreviations: glyceraldehyde 3-phosphate (G3P); isopentenyl phosphate (IP); isoprene diphosphate (IPP); dimethylallyl diphosphate (DMAPP); neryl diphosphate (NPP); farnesyl diphosphate (FPP); *Z*-*Z*-farnesyl diphosphate (zFPP).

other cells, or other trichome types (Supplemental Figure S8). We thus obtained mRNA profiles highly specific for the type-VI secretory cells.

In both genotypes, genes of the MEP pathway were relatively higher expressed than those of the MVA pathway (Figure 6B; Supplemental Dataset 1). Differential-expression analysis comparing the two genotypes showed that in the high-producing glands of F2-28, all genes of the MEP-pathway were significantly higher expressed compared to the cultivar (Supplemental Figure S9). The biggest difference in expression was found for genes in the last part of the MEP pathway; *zFPS* compared to *NDPS1* (20-fold), *IDI1* (4-fold), *HDR* (3-fold), *HDS* (2-fold) and the also first committed step of the pathway *DXS2* (2-fold; Figure 6C). Interestingly, most genes encoding enzymes of the MVA pathway were significantly higher expressed in F2-28 as well; *AACT3* (23-fold), *HMGS3* (3-fold), *HMGR3* (19-fold), *MVK* (1.5-fold), *PMK* (30-fold), *DMC1* (1.5-fold), *DMC2* (4.5-fold; Supplemental Figure S9; Supplemental Dataset 1). However, compared to the MEP pathway, expression levels of the MVA pathway genes were generally much lower. Moreover, in contrast to the MEP pathway, the final enzymatic steps of MVA pathway, *IDI2* and *FPS*, were equally expressed in the active F2-28 and the lazy cultivar (Supplemental Figure S9). Also, IPK and NUDIX, exhibiting relatively low expression levels in tomato secretory cells anyway, did not differ (Supplemental Dataset 1). Together, the results show that in the secretory cells MEP pathway genes are expressed much higher than the MVA pathway. In addition, transcript levels of the final steps of the plastidial-precursor pathway are particularly high in the active F2 trichomes compared to the cultivar.

Finally, the transcript levels of confirmed regulators of terpene biosynthesis in tomato trichomes were compared (Supplemental Figure S10). The expression of transcription factor *SCL3*, recently identified to be a regulator of *DXS2* (Yang et al., 2021) was 2-fold higher in F2-28 compared to the cultivar, in line with the higher transcript levels of *DXS2*. There was no differential expression in other transcription factors of potential interest. *EOT1*, *MYC1* and *MYB75* (Spyropoulou et al., 2014; Xu et al., 2018; Gong et al., 2021) did not significantly differ between the two genotypes studied here.

Discussion

High-level 7-epizingiberene production is a multigenic trait

We discovered that introgression of a metabolic trait into a tomato cultivar is challenging as the capacity to produce high quantities of terpenes seems to be constrained in a cultivated tomato (i.e. *S. lycopersicum*). Despite the fact that cultivated tomato has the suitable micro-organ to produce terpenes, i.e. a glandular-trichome type-VI, we showed that in the offspring of an interspecific cross with a high-producing wild tomato, terpene levels were much lower than expected based on the gene dosage of biosynthetic enzymes. The synthesis of sesquiterpenes, including 7epiZ and its derivatives, and monoterpenes is (co-)dominant, hence the F1 hybrid is able produce 7epiZ, 9HZ, 9H10epoZ as well as monoterpenes. However, terpene levels in F1 plants were only 7% of that of PI127826, and not significantly higher than those in a cultivar (Figure 1A). In contrast, terpene levels in the F1 made from two *S. habrochaites* accessions (F1-hab; LA1777 x PI127826) were similar to PI127826. This indicates that the germplasms of PI127826 and LA1777 share the (genetic) factors necessary to accumulate high levels of terpenes in type-VI trichomes or, alternatively, lack inhibiting factors that may be present in *S. lycopersicum*.

The pre-selection of the F2 plants for homozygous alleles of the full 7epiZ pathway (i.e. *zFPS*/*ZIS*/*ShZO*) allowed us to study rate limiting factors for high production in the plastid without interference of monoterpene synthesis *via NDPS1*/*PHS1*. The segregation of 7epiZ levels in this F2 population, which were highly skewed to low levels as found in the cultivar and the F1 (Figure 2A), indicated a multigenic recessive nature of high terpene production by type-VI trichomes. Our results are in line with previous findings describing trichome-produced specialised metabolite levels in segregating material of different interspecific crosses (Zamir et al., 1984; Frelichowski and Juvik, 2001; Zhang et al., 2008; Ben-Israel et al., 2009; Ji et al., 2022). Interspecific F1-hybrids made with various other *S. habrochaites* accessions (i.e. LA1777, LA1363, LA0407 and PI126449) or *S. pennellii* (LA0716) also displayed a disproportionate low production of specialised metabolites. Moreover, the frequency distributions of metabolite levels in segregating populations were also highly skewed towards the low levels found in the F1s (Zamir et al., 1984; Frelichowski and Juvik, 2001; Zhang et al., 2008; Ben-Israel et al., 2009; Ji et al., 2022). Similar frequency distributions were observed when F1-hybrids were back crossed to the wild parent pointing out the involvement of recessive alleles (Zamir et al., 1984; Frelichowski and Juvik, 2001).

Frelichowski and Juvik (2005) calculated three recessive loci to regulate santalene and bergamotene levels in LA1777. Likewise, biosynthesis of the methyl ketone 2-tridecanone in *S. habrochaites* LA0704, was predicted to require three genes in addition to the two known methyl ketone synthases (Fery and Kennedy, 1987). The genetic models that fit the frequency distribution of high 7epiZ levels observed here also comprise either three recessive loci, or two recessive loci plus one or two dominant ones (Table 1). These loci are irrespective of the *zFPS*/*ZIS* locus on chromosome 8 (Solyc08g005680 and Solyc08g005670) and the *ShZO* locus on chromosome 1 (Solyc01g008670), and do not seem to be linked. Noticeably, the examples above entail type-VI trichome specialised metabolites that originate from different precursor pathways. The SCAs in LA1777 and 7epiZ (derivatives) in PI127826 are MEP-pathway derived, while the terpenoids in LA1363 originate from the cytosolic MVA-pathway, and methyl-ketones (e.g. 2-tridecanone and 2-undecanone) are derived from (plastidial) fatty-acid biosynthesis. The factors that determine "high-level metabolite production" must therefore be involved in facilitating the flux, or enhance the storage capacity, independent of a particular pathway.

Trichome activity and density are two independent traits

The total quantity of terpenes on a leaf-surface area is determined by a combination of type-VI trichome-density, the metabolic activity of the individual trichomes and their capacity to store and retain the metabolites. While it is obvious that the presence of high numbers of glandular trichomes contributes to the levels a plant can produce, trichome density alone could not explain the total terpene levels produced by the studied F2 plants. We found individuals with high numbers of trichomes producing relatively low levels of terpenes, and vice versa (Figure 2B; Supplemental Figure S3B). Hence, density can contribute, but does not explain high 7epiZ levels at the leaf level by itself. Likewise, the work of Ben-Israel et al. (2009) on the production of methyl-ketones in type-VI trichomes, showed that trichome density had a minor effect on total levels found on the leaf. Interestingly, trichome shape and allelic variants of the biosynthetic enzymes correlated better to the methyl-ketone concentrations (Ben-Israel et al., 2009). In conclusion, the "productivity" phenotype of type-VI trichomes segregates independently from trichome density (Figure 2) and should therefore be regarded as a separate trait that, in combination with density, will determine the total level of metabolites produced on the leaf surface.

Type-VI trichome density, typically higher in PI127826, appease to be co-dominant as the F1 hybrid exhibits a density phenotype intermediate to both parents (Figure 1B). A similar pattern was observed for type-I/IV and non-glandular trichomes (Supplemental Figure S1), suggesting that regulatory control over trichome development (e.g. initiation and type-differentiation) in general is regulated by co-dominant factors. Several studies indeed discovered trichome initiation and development is under control of transcription-factor protein-complexes that require all elements of the complex to be present (Guevara-García et al., 2005; Kang et al., 2016; Gao et al., 2017; Xu et al., 2018).

Type-VI storage cavity volume and shape are determined by activity of the secretory cells

The difference in morphology of type-VI trichome glandular heads between *S. habrochaites* and *S. lycopersicum* accessions was proposed earlier to be related to the metabolite quantities inside the trichome (Bergau et al., 2015). The formation of "globular" shaped trichomes may be due to the "inflated-balloon" effect, where expansion of the internal cavity pushes the four secretory cells, encapsulated by a cuticular sac, outwards creating a spherical shape (Ben-Israel et al., 2009). The expansion of the cavity is likely facilitated by loosening of the gland-cell's cell wall at the side facing the storage cavity, a process that seems to be more pronounced in *S. habrochaites* type-VI trichomes compared to *S. lycopersicum* (Bergau et al., 2015). Loosening of the cell walls, e.g. by pectinases or expansins, may therefore be a prerequisite for cavity expansion. Still, the accumulation of terpenoids remains a determining factor. The inhibitor assays done here show that the cavity volume in newly developed trichomes is dependent on the ability to produce terpenoids (Figure 5B). This "filling" of the cavity must take place predominantly during the very early stages of trichome development, as matured trichomes exposed to fosmidomycin treatment were much less affected (Supplemental Figure S6). This indicates that the volume of the storage cavity is not a predetermined trait but rather the consequence of metabolite accumulation early in trichome development. The fact that we observed only minor variation in cavity volume of trichomes on an individual plant, whether it is an "active" or "lazy" genotype (Figure 3B, lower panel), suggests that the metabolic activity of all secretory cells is co-regulated.

A high metabolic flux through the MEP-pathway for the productivity of both plastidial as cytosolically localised terpene synthases

Compartmentalised production of IPP seems an efficient way to accommodate increased terpenoid production (Aharoni et al., 2003; Wu et al., 2006; Dong et al., 2016). With genes of the MEP pathway highly expressed, trichomes accumulated high amounts of terpenes *via* both cytosolic- and plastid-localised synthases (Figure 5A; Zhou and Pichersky, 2020). Evidence for precursor-crosstalk (i.e IPP/DMAPP) between plastids and the cytosol has accumulated over the past years (Hemmerlin et al., 2003; Liao et al., 2016; Henry et al., 2018; Gutensohn et al., 2021). The substantial reduction in terpene content observed after fosmidomycin treatment of active genotypes (Figure 5A) suggests that type-VI trichomes have a high metabolic flux through the MEP pathway driving the synthesis of terpenes in the plastid, and also in the cytosol. This indicates that also cytosolic terpene synthases in *S. habrochaites* and active F2-genotypes largely rely on precursors derived from MEP pathway. The contribution of the MVA pathway to the total pool of IPP in the secretory cells indeed seems limited as no significant effect was observed on either mono- or sesquiterpenes levels under treatment with mevastatin (Figure 5A). Combining those results with our gene-expression data (Figure 6), we hypothesised that the metabolic flux of isoprenoid precursors IPP/DMAPP in trichomes of *S. habrochaites* predominantly runs through a highly active MEP pathway. Such directional flux was described previously for snapdragon flowers (Dudareva et al., 2005) where sesquiterpene biosynthesis in the cytosol appeared to rely largely on the IPP generated by the MEP pathway. A bidirectional precursor-flux was shown previously for tomato fruit (Gutensohn et al., 2021). However, the MVA pathway contributes to only little amounts of terpenes in tomato trichomes (Figure 5A).

Although we did not investigate this further, we cannot exclude that genetic factors from *S. lycopersicum* repress the synthesis of terpenes. These could include (post-)transcriptional repressors of (regulators of) terpene biosynthesis, metabolite transport and storage capacity, the availability of carbon and energy or the branching of available carbon to other pathways

(e.g. flavonoids). We did not see large differences in gene expression of (confirmed) transcriptional regulators of terpene metabolism, except for a 2-fold higher transcript level of *MIXTA-like* and *SCL3* which is in line with enhanced activation of the MEP pathway in the active F2-genotype. Through quantitative trait locus (QTL) mapping, Bennewitz et al. (2018) indicated multiple loci of *S. habrochaites* LA1777 correlated to the particular "globular" shape of its type-VI trichomes. They found a QTL on chromosome 1 that shows strong correlation to the shape, with additional loci on chromosome 7 and chromosome 11. As they used a backcross population, these QTLs indicate dominant loci. Here we showed the connection between terpene accumulation, storage-cavity volume and trichome morphology; the loci proposed by Bennewitz et al. (2018) may thus contain genes involved in terpene biosynthesis additional to gland shape. The enhanced metabolic flux can be the result of multiple upstream or downstream processes including transcriptional activation and the capacity to transport and store end-products in the storage cavity.

Methods

Plant material

The F1 was generated by transferring pollen from *S. habrochaites* PI127826 onto the stamen of a *S. lycopersicum* cultivar obtained from Enza Zaden Enkhuizen after which F1 seed was harvested. Next, the F1 was selfed to generate the *S. lycopersicum* x *S. habrochaites* PI127826 F2 population. F2 individuals were preselected for the presence of Tomato Mosaic Virus resistance of the cultivar and the 7-epizingiberene synthase (*ZIS;* Solyc08g005640), *zFPS* (Solyc08g005640) and a P450 (Solyc01g008650) from PI127826, using SNP markers. F1-hab was generated by crossing *S. habrochaites* PI127826 (male) to *S. habrochaites* LA1777 (female). The F2s, F1s and prenatal plants were maintained as cuttings and grown in enclosed climate-controlled greenhouse compartments (22–25°C; 16/8h light/dark). Biological replicates originate by growing multiple cuttings from the sample mother plant. All experiments were performed on ~4 weeks old cuttings except when indicated differently. Unless indicated otherwise, the leaf material and type-VI trichomes originated from 4th fully developed leaf from the shoot apex, of which the lateral leaflets after the leaf's terminal leaflet were taken.

Trichome phenotyping

Trichome density

Two leaf discs (r = 2 mm) were taken per leaf surface (i.e. abaxial/adaxial) and the total number of non-glandular trichomes, type-VI trichomes and type I-and VI trichomes on each leaf disc was counted by eye under a stereo microscope. Per trichome type, the number of trichomes was summed over the abaxial-and adaxial surface to obtain the total number of trichome per leaf-disc area. Mean densities over the two leaf discs were taken to determine the density per biological replicate (n = 3 cuttings).

Type-VI trichome-density classes

To assign the F2 genotypes to trichome-density classes, as displayed in Supplemental Figure S3A, two leaf discs (r = 2 mm) were taken from each genotype. One leaf disc was used to estimate the number of type-VI trichomes on the abaxial surface, the second leaf disc for the abaxial surface. Next, the adaxial and adaxial surface were summed giving the total number of trichomes per leaf disc. For the parental genotypes, the mean summed number of trichomes was taken after outlier analysis (n= 4-5 cuttings).

Microscopy and storage-cavity measurements

Strips of the peripheral ends were cut from leaflets and mounted on a microscope slide, immersed in a drop of water and covered by a cover slip. Trichome images were taken with a EVOS^{TM} digital inverted microscope (www.thermofisher.com) under bright light using a 20X objective. The height (medial plane) and width (transverse plane) of the cavities were measured using ImageJ (https://imagej.nih.gov) and the cavity volume was calculated according to the formula for prolate spheroids: volume = $(4/3)*a^2*b$, were a is the radius of the cavity's width and b the radius of its height.

Volatile analysis

Metabolite extraction

Leaflets were first weighted and subsequently submerged in 1mL *n*-hexane spiked with 0.5 ng/µL benzyl acetate as internal standard. After briefly rocking for 5 seconds the hexane was transferred to a 1.5 mL Eppendorf Safe-Lock tube containing ~10 mg Na_2SO_4(s) and vortexed for 5 seconds. The tubes were spun down for 5 minutes at 12.000 rpm where after the hexane extract was transferred to a 2 mL GC-MS glass vial and stored under N_2(g) at -20°C until GC-MS analysis. For type-VI trichome gland-cell analysis, per sample 150 trichomes were collected from the adaxial side of a leaflet using the tip of a stretched-out glass Pasteur-pipette (final radius ~80-100 µM) while under a stereomicroscope. Volatiles were extracted in Eppendorf Safe-Lock tubes containing 300 µL *n*-hexane spiked with 0.5 ng/µL benzyl acetate. During this procedure, the tip of the Pasteur pipette picks up the type-VI head-cells by capillary force after brief contact. The pipette was washed in the *n*-hexane with every 30 head-cells collected to prevent clogging of the tip. After collection, ~10 mg Na_2SO_4(s) was added to the tube followed by 5 seconds of vortexing. The tubes were spun down for 5 minutes at 12.000 rpm where after ~150 µL of the hexane extract was transferred to a 2 mL GC-MS glass vial containing an autosampler insert and stored under N_2(g) at -20°C until GC-MS analysis.

Data acquisition

Gas chromatograph-mass spectrometry was performed using an Agilent 7890A gas chromatograph coupled to an Agilent 7200

mass spectrometer. Depending on the sample, 1 or 2 µL of sample was injected, heated to 275°C and separated on a HP-5MS column (30 m × 250 µm x 0.25 µm; Agilent) using helium as carrier gas (flow rate: 1 mL min^{-1}). The column was heated to 40°C for 3 min, after which the temperature increased 15°C/min up to 250°C which was held for 5 min. Ionisation of the sample was done at 70 eV with an emission of 35 µA at 230°C and data was collected with rate of 5 scans/second within a mass range of 40 – 300 mu.

Data analysis

Chromatographic peaks were integrated on the spectral base peak using Masshunter Quantitative software (Agilent). Peak areas were corrected for the internal standard, sample dilution, injected sample volume and, depending on the experiment, normalised for fresh leaf weight or number of collected trichomes. Normalised peak areas were quantified using authentic standards if possible. β-caryophyllene was used for quantification of 7-epizingiberene and its derivatives, α/β-bergamotene and α-santalene. The 95% confidence interval of PI127826 levels (n = 4) was determined by the mean peak area ± 1.96*(standard deviation/√n).

Inhibition assays

Freshly made cuttings were grown on Terra Aquatica (www.terraaquatica.com/) hydroponic solution containing TriPart-Grow/TriPart-Micro/Tripart-Bloom/water: 1/1/1/9997 (v/v/v/v) in a climate-controlled greenhouse-compartment (22–25°C, 16/8h light/dark). After 1 week, roots emerged and cuttings were transferred to fresh hydroponic solutions supplemented with 10 µM fosmidomycin (Sigma-Aldrich, Catalog number: F8682) or 10 µM mevastatin (Sigma-Aldrich, Catalog number: M2537) or without inhibitors. After 14 days, the first lateral leaflets after the terminal leaflet of the 2nd or 3rd leaf from the apex were taken for microscopy and volatile analysis of the trichome head-cells. To ensure effective inhibitor treatment, cuttings were checked for development-related appearances that are typical after treatment with these inhibitors. Fosmidomycin-treated cuttings displayed bleaching of the newly formed leaves, while mevastatin-treated cuttings showed a stunted-growth phenotype (Supplemental Figure S7).

Gland isolation and RNA extraction

Approximately 15-20 leaflets were collected in 50 mL tubes and were fully submerged in 30 mL ice-cold 70% EtOH. Tubes were hand-shaken for 20 times followed by 10 seconds of vortexing, whereafter the leaflets were briefly checked under a stereo microscopy for remaining type-VI trichomes. When many trichomes were still attached to the leaflets the shaking/vortexing step was repeated. The leaflets were next removed, and the extract was sieved through a 70 µm nylon cell strainer (Corning® product # 431751) and collected into a fresh 50 mL tube. A small fraction on the collected samples (~50 µL) was inspected under a microscope for the enrichment of type-VI glands (Supplemental Figure S8) and to assure minimal contamination. The tubes were centrifuged at 3.000g for 5 minutes at 4°C. The pellet was resuspended in 1 mL 70% EtOH and, for each RNA-extraction, two resuspended pellets were combined into a fresh 2 mL Eppendorf® tube. The Eppendorf tubes were centrifuged at 13.000g for 5 minutes at 4°C whereafter the supernatant was removed, and the pellet was frozen in N_2 (l). Steel beads were added to the tube and shaken for 60 seconds using a paint shaker, then frozen in N_2 (l) and stored at -80°C until further processing. The grinded tissue was used as input material for RNA extraction using the NucleoSpin® RNA XS kit (Macherey-Nagel product # 740902) following manufacturer's instructions and was eluted in 25 µL water. To further purify the RNA samples, 5 µL 10M of LiCl was added and samples were left overnight at -20°C. Precipitated RNA was spun down at 13.000 rpm for 20 minutes at 4°C and subsequently washed twice with 500µL 70% EtOH and resuspended in 10 µL water.

Library preparation and mRNA sequencing

Prior to library preparation, poly-A enrichment was done using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (New England BioLabs). RNA-sequencing libraries were generated using the NEBNext® Ultra II Directional RNA Library Prep Kit for Illumina with the NEBNext® Multiplex Oligos for Illumina® (Unique Dual Index Primer Pairs) according to the manufacturers' protocols for larger insert size libraries (New England BioLabs). The size distribution of the libraries with indexed adapters was assessed using a 2200 TapeStation System with Agilent D5000 ScreenTapes (Agilent Technologies) and libraries were quantified using a QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific) using the NEBNext® Library Quant Kit for Illumina (New England BioLabs) according to the manufacturer's instructions. Libraries were clustered and sequenced (2x 150 bp) on a NovaSeq 6000 Sequencing System (Illumina) using a NovaSeq 6000 S4 Reagent Kit v1.5 (300 cycles, Illumnia). The mRNA-Seq reads were analysed using a custom Snakemake (Mölder et al., 2021) pipeline coined "snakemake-rnaseq" that is based on fastp, Kallisto and Subread (Liao et al., 2014; Bray et al., 2016; Chen et al., 2018). Read counts were normalised using DESeq2 in R (Love et al., 2014).

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://github.com/BleekerLab/genetic_requirements/releases/tag/v0.5.0.

Author contributions

PB, RK and MG conceived the study. RK and PB wrote the manuscript. RK, AM, RT and SM performed the experiments. MG performed the bioinformatic analyses. RS and MH provided input

for the experiments and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by an NWO VIDI-grant from the Dutch Organisation for Scientific Research (NWO; grant number 12988) obtained by PB, covering all project cost and the salary of RK, AM and MG. RT was funded by TKI T&U "Topsectoren" grand number TU-2017-012 obtained by PB.

Acknowledgments

The authors would like to thank the research and breeding departments of Enza Zaden Research and Development B.V. for their sizeable contribution to this manuscript, especially for providing the plant material, for creating, maintaining and genotyping the segregating population, and for help with metabolite analysis. Michelle van der Gragt and Sanne Ypenburg are thanked for their help with counting the trichomes on the F2 population.

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.2023.1139274/full#supplementary-material

References

Aharoni, A., Giri, A. P., Deuerlein, S., Griepink, F., de Kogel, W., Verstappen, F. W., et al. (2003). Terpenoid metabolism in wild-type and transgenic arabidopsis plants. *Plant Cell* 15 (12), 2866–2884. doi: 10.1105/tpc.016253

Akhtar, T. A., Matsuba, Y., Schauvinhold, I., Yu, G., Lees, H. A., Klein, S. E., et al. (2013). The tomato cis–prenyltransferase gene family. *Plant J.* 73 (4), 640–652. doi: 10.1111/tpj.12063

Andama, J. B., Mujiono, K., Hojo, Y., Shinya, T., and Galis, I. (2020). Nonglandular silicified trichomes are essential for rice defense against chewing herbivores. *Plant Cell Environ.* 43 (9), 2019–2032. doi: 10.1111/pce.13775

Ben-Israel, I., Yu, G., Austin, M. B., Bhuiyan, N., Auldridge, M., Nguyen, T., et al. (2009). Multiple biochemical and morphological factors underlie the production of methylketones in tomato trichomes. *Plant Physiol.* 151 (4), 1952–1964. doi: 10.1104/pp.109.146415

Bennewitz, S., Bergau, N., and Tissier, A. (2018). QTL mapping of the shape of type VI glandular trichomes in tomato. *Front. Plant Sci.* 9, 1–12. doi: 10.3389/fpls.2018.01421

Bergau, N., Bennewitz, S., Syrowatka, F., Hause, G., and Tissier, A. (2015). The development of type VI glandular trichomes in the cultivated tomato solanum lycopersicum and a related wild species s. habrochaites. *BMC Plant Biol.* 15 (1), 1–15. doi: 10.1186/s12870-015-0678-z

Bleeker, P. M., Diergaarde, P. J., Ament, K., Schütz, S., Johne, B., Dijkink, J., et al. (2011a). Tomato-produced 7-epizingiberene and r-curcumene act as repellents to whiteflies. *Phytochemistry* 72 (1), 68–73. doi: 10.1016/j.phytochem.2010.10.014

Bleeker, P. M., Mirabella, R., Diergaarde, P. J., VanDoorn, A., Tissier, A., Kant, M. R., et al. (2012). Improved herbivore resistance in cultivated tomato with the sesquiterpene biosynthetic pathway from a wild relative. *Proc. Natl. Acad. Sci.* 109 (49), 20124–20129. doi: 10.1073/pnas.1208756109

Bleeker, P. M., Spyropoulou, E. A., Diergaarde, P. J., Volpin, H., De Both, M. T., Zerbe, P., et al. (2011b). RNA-Seq discovery, functional characterization, and comparison of sesquiterpene synthases from solanum lycopersicum and solanum habrochaites trichomes. *Plant Mol. Biol.* 77 (4), 323–336. doi: 10.1007/s11103-011-9813-x

Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527. doi: 10.1038/nbt.3519

Chen, G., Klinkhamer, P. G., Escobar-Bravo, R., and Leiss, K. A. (2018). Type VI glandular trichome density and their derived volatiles are differently induced by jasmonic acid in developing and fully developed tomato leaves: Implications for thrips resistance. *Plant Sci.* 276, 87–98. doi: 10.1016/j.plantsci.2018.08.007

Dong, L., Jongedijk, E., Bouwmeester, H., and Van Der Krol, A. (2016). Monoterpene biosynthesis potential of plant subcellular compartments. *New Phytol.* 209 (2), 679–690. doi: 10.1111/nph.13629

Dudareva, N., Andersson, S., Orlova, I., Gatto, N., Reichelt, M., Rhodes, D., et al. (2005). The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. *Proc. Natl. Acad. Sci.* 102 (3), 933–938. doi: 10.1073/pnas.0407360102

Escobar-Bravo, R., Ruijgrok, J., Kim, H. K., Grosser, K., Van Dam, N. M., Klinkhamer, P. G., et al. (2018). Light intensity-mediated induction of trichome-associated allelochemicals increases resistance against thrips in tomato. *Plant Cell Physiol.* 59 (12), 2462–2475. doi: 10.1093/pcp/pcy166

Fery, R. L., and Kennedy, G. G. (1987). Genetic analysis of 2-tridecanone concentration, leaf trichome characteristics, and tobacco hornworm resistance in tomato. *J. Am. Soc. Hortic. Sci.* 112 (5), 886–891. doi: 10.21273/JASHS.112.5.886

Frelichowski, J. E.Jr., and Juvik, J. A. (2001). Sesquiterpene carboxylic acids from a wild tomato species affect larval feeding behavior and survival of helicoverpa zea and spodoptera exigua (Lepidoptera: Noctuidae). *J. Economic Entomology* 94 (5), 1249–1259. doi: 10.1603/0022-0493-94.5.1249

Frelichowski, J. E.Jr., and Juvik, J. A. (2005). Inheritance of sesquiterpene carboxylic acid synthesis in crosses of *Lycopersicon hirsutum* with insect-susceptible tomatoes. *Plant Breed.* 124, 277–281. doi: 10.1111/j.1439-0523.2005.01076.x

Fridman, E., Wang, J., Iijima, Y., Froehlich, J. E., Gang, D. R., Ohlrogge, J., et al. (2005). Metabolic, genomic, and biochemical analyses of glandular trichomes from the wild tomato species lycopersicon hirsutum identify a key enzyme in the biosynthesis of methylketones. *Plant Cell* 17 (4), 1252–1267. doi: 10.1105/tpc.104.029736

Gao, S., Gao, Y., Xiong, C., Yu, G., Chang, J., Yang, Q., et al. (2017). The tomato b-type cyclin gene, SlCycB2, plays key roles in reproductive organ development, trichome initiation, terpenoids biosynthesis and prodenia litura defense. *Plant Sci.* 262, 103–114. doi: 10.1016/j.plantsci.2017.05.006

Glas, J. J., Schimmel, B. C., Alba, J. M., Escobar-Bravo, R., Schuurink, R. C., and Kant, M. R. (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *Int. J. Mol. Sci.* 13 (12), 17077–17103. doi: 10.3390/ijms131217077

Gong, Z., Luo, Y., Zhang, W., Jian, W., Zhang, L., Gao, X., et al. (2021). A SlMYB75-centred transcriptional cascade regulates trichome formation and sesquiterpene accumulation in tomato. *J. Exp. Bot.* 72 (10), 3806–3820. doi: 10.1093/jxb/erab086

Guevara-García, A., San Román, C., Arroyo, A., Cortés, M. E., de la Luz Gutiérrez-Nava, M., and León, P. (2005). Characterization of the arabidopsis clb6 mutant

illustrates the importance of posttranscriptional regulation of the methyl-d-erythritol 4-phosphate pathway. *Plant Cell* 17 (2), 628–643. doi: 10.1105/tpc.104.028860

Gutensohn, M., Henry, L. K., Gentry, S. A., Lynch, J. H., Nguyen, T. T., Pichersky, E., et al. (2021). Overcoming bottlenecks for metabolic engineering of sesquiterpene production in tomato fruits. *Front. Plant Sci.* 1107. doi: 10.3389/fpls.2021.691754

Hemmerlin, A., Harwood, J. L., and Bach, T. J. (2012). A raison d'être for two distinct pathways in the early steps of plant isoprenoid biosynthesis? *Prog. Lipid Res.* 51 (2), 95–148. doi: 10.1016/j.plipres.2011.12.001

Hemmerlin, A., Hoeffler, J., Meyer, O., Tritsch, D., Kagan, I. A., Grosdemange-Billiard, C., et al. (2003). Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells. *J. Biol. Chem.* 278 (29), 26666–26676. doi: 10.1074/jbc.M302526200

Henry, L. K., Thomas, S. T., Widhalm, J. R., Lynch, J. H., Davis, T. C., Kessler, S. A., et al. (2018). Contribution of isopentenyl phosphate to plant terpenoid metabolism. *Nat. Plants* 4 (9), 721–729. doi: 10.1038/s41477-018-0220-z

Ji, W., Mandal, S., Rezenom, Y. H., and McKnight, T. D. (2022). Specialized metabolism by trichome-enriched rubisco and fatty acid synthase components. *Plant Physiol.* 191 (2), 1199–213. doi: 10.1093/plphys/kiac487

Kang, J., Campos, M. L., Zemelis-Durfee, S., Al-Haddad, J. M., Jones, A. D., Telewski, F. W., et al. (2016). Molecular cloning of the tomato hairless gene implicates actin dynamics in trichome-mediated defense and mechanical properties of stem tissue. *J. Exp. Bot.* 67 (18), 5313–5324. doi: 10.1093/jxb/erw292

Kang, J., Gonzales-Vigil, E., Matsuba, Y., Pichersky, E., and Barry, C. S. (2014). Determination of residues responsible for substrate and product specificity of *Solanum habrochaites* short-chain cis-prenyltransferases. *Plant Physiol.* 164 (1), 80–91. doi: 10.1104/pp.113.230466

Kariyat, R. R., Smith, J. D., Stephenson, A. G., De Moraes, C. M., and Mescher, M. C. (2017). Non-glandular trichomes of solanum carolinense deter feeding by manduca sexta caterpillars and cause damage to the gut peritrophic matrix. *Proc. R. Soc. B: Biol. Sci.* 284 (1849), 20162323. doi: 10.1098/rspb.2016.2323

Kortbeek, R. W., van der Gragt, M., and Bleeker, P. M. (2019). Endogenous plant metabolites against insects. *Eur. J. Plant Pathol.* 154 (1), 67–90. doi: 10.1007/s10658-018-1540-6

Liao, P., Hemmerlin, A., Bach, T. J., and Chye, M. (2016). The potential of the mevalonate pathway for enhanced isoprenoid production. *Biotechnol. Adv.* 34 (5), 697–713. doi: 10.1016/j.biotechadv.2016.03.005

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30 (7), 923–930. doi: 10.1093/bioinformatics/btt656

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 (12), 1–21. doi: 10.1186/s13059-014-0550-8

Luckwill, L. C. (1943). The genus lycopersicon: An historical, biological, and taxonomic survey of the wild and cultivated tomatoes. *Aberd. Univ. Stud.* 120, 1–44.

Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., et al. (2021). Sustainable data analysis with snakemake. *F1000Research* 10 (33), 1–28. doi: 10.12688/f1000research.29032.2

Nakashima, T., Wada, H., Morita, S., Erra-Balsells, R., Hiraoka, K., and Nonami, H. (2016). Single-cell metabolite profiling of stalk and glandular cells of intact trichomes with internal electrode capillary pressure probe electrospray ionization mass spectrometry. *Analytical Chem.* 88 (6), 3049–3057. doi: 10.1021/acs.analchem.5b03366

Sallaud, C., Rontein, D., Onillon, S., Jabès, F., Duffé, P., Giacalone, C., et al. (2009). A novel pathway for sesquiterpene biosynthesis from z, z-farnesyl pyrophosphate in the wild tomato solanum habrochaites. *Plant Cell* 21 (1), 301–317. doi: 10.1105/tpc.107.057885

Schilmiller, A. L., Miner, D. P., Larson, M., McDowell, E., Gang, D. R., Wilkerson, C., et al. (2010). Studies of a biochemical factory: Tomato trichome deep expressed sequence tag sequencing and proteomics. *Plant Physiol.* 153 (3), 1212–1223. doi: 10.1104/pp.110.157214

Schilmiller, A. L., Schauvinhold, I., Larson, M., Xu, R., Charbonneau, A. L., Schmidt, A., et al. (2009). Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proc. Natl. Acad. Sci.* 106 (26), 10865–10870. doi: 10.1073/pnas.0904113106

Schuurink, R., and Tissier, A. (2020). Glandular trichomes: micro-organs with model status? *New Phytol.* 225 (6), 2251–2266. doi: 10.1111/nph.16283

Spyropoulou, E. A., Haring, M. A., and Schuurink, R. C. (2014). Expression of terpenoids 1, a glandular trichome-specific transcription factor from tomato that activates the terpene synthase 5 promoter. *Plant Mol. Biol.* 84 (3), 345–357. doi: 10.1007/s11103-013-0142-0

Therezan, R., Kortbeek, R., Vendemiatti, E., Legarrea, S., de Alencar, S. M., Schuurink, R. C., et al. (2021). Introgression of the sesquiterpene biosynthesis from solanum habrochaites to cultivated tomato offers insights into trichome morphology and arthropod resistance. *Planta* 254 (1), 1–16. doi: 10.1007/s00425-021-03651-y

Tissier, A., Morgan, J. A., and Dudareva, N. (2017). Plant volatiles: Going 'in' but not 'out' of trichome cavities. *Trends Plant Sci.* 22 (11), 930–938. doi: 10.1016/j.tplants.2017.09.001

Vendemiatti, E., Zsögön, A., Silva, G. F. F. E., de Jesus, F. A., Cutri, L., Figueiredo, C. R. F., et al. (2017). Loss of type-IV glandular trichomes is a heterochronic trait in tomato and can be reverted by promoting juvenility. *Plant Sci.* 259, 35–47. doi: 10.1016/j.plantsci.2017.03.006

Wu, S., Schalk, M., Clark, A., Miles, R. B., Coates, R., and Chappell, J. (2006). Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nat. Biotechnol.* 24 (11), 1441–1447. doi: 10.1038/nbt1251

Xu, J., van Herwijnen, Z. O., Dräger, D. B., Sui, C., Haring, M. A., and Schuurink, R. C. (2018). SlMYC1 regulates type VI glandular trichome formation and terpene biosynthesis in tomato glandular cells. *Plant Cell* 30 (12), 2988–3005. doi: 10.1105/tpc.18.00571

Yang, C., Marillonnet, S., and Tissier, A. (2021). The scarecrow-like transcription factor SlSCL3 regulates volatile terpene biosynthesis and glandular trichome size in tomato (*Solanum lycopersicum*). *Plant J.* 107, 1102–1118. doi: 10.1111/tpj.15371

Zabel, S., Brandt, W., Porzel, A., Athmer, B., Bennewitz, S., Schäfer, P., et al. (2021). A single cytochrome P450 oxidase from solanum habrochaites sequentially oxidizes 7-epi-zingiberene to derivatives toxic to whiteflies and various microorganisms. *Plant J.* 105 (5), 1309–1325. doi: 10.1111/tpj.15113

Zamir, D., Selilaben-David, T., Rudich, J., and Juvik, J. A. (1984). Frequency distributions and linkage relationships of 2-tridecadone in interspecific segregating generations of tomato. *Euphytica* 33 (2), 481–488. doi: 10.1007/BF00021148

Zeidler, J. G., Schwender, J., Muller, C., Wiesner, J., Weidemeyer, C., Beck, E., et al. (1998). Inhibition of the non-mevalonate 1-deoxy-D-xylulose-5-phosphate path-way of plant isoprenoid biosynthesis by fosmidomycin. *Z. Naturforsch.* 53C, 980–986. doi: 10.1515/znc-1998-11-1208

Zhang, X., Thacker, R. R., and Snyder, J. C. (2008). Occurrence of 2, 3-dihydrofarnesoic acid, a spidermite repellent, in trichome secretions of lycopersicon esculentum× l. hirsutum hybrids. *Euphytica* 162 (1), 1–9. doi: 10.1007/s10681-007-9489-0

Zhou, F., and Pichersky, E. (2020). The complete functional characterisation of the terpene synthase family in tomato. *New Phytol.* 226 (5), 1341–1360. doi: 10.1111/nph.16431

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 31 May 2023
DOI 10.3389/fpls.2023.1181311

OPEN ACCESS

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

REVIEWED BY
Dong Wook Lee,
Chonnam National University,
Republic of Korea
Umar Gani,
University of British Columbia, Canada

*CORRESPONDENCE
Sun-Hwa Ha
sunhwa@khu.ac.kr

†These authors have contributed equally to this work

RECEIVED 07 March 2023
ACCEPTED 10 May 2023
PUBLISHED 31 May 2023

CITATION
Yu JS, You MK, Lee YJ and Ha S-H (2023) Stepwise protein targeting into plastoglobules are facilitated by three hydrophobic regions of rice phytoene synthase 2. *Front. Plant Sci.* 14:1181311. doi: 10.3389/fpls.2023.1181311

Stepwise protein targeting into plastoglobules are facilitated by three hydrophobic regions of rice phytoene synthase 2

Ji Su Yu†, Min Kyoung You†, Yeo Jin Lee and Sun-Hwa Ha*

Department of Genetics and Biotechnology, Graduate School of Green-Bio Science, College of Life Sciences, Kyung Hee University, Yongin, Republic of Korea

Plastoglobules (PGs) are plastidial lipid droplets enclosed by a polar monolayer born from the thylakoid membrane when plants require active lipid metabolism, including carotenogenesis, under the environmental stress and during plastid transition. Despite the fact that many proteins are reported to target PGs, their translocation mechanism has remained largely unexplored. To elucidate this process, we studied the influence of three hydrophobic regions (HR)—HR1 (1–45th aa), HR2 (46–80th aa), and HR3 (229–247th aa)—of rice phytoene synthase 2 (OsPSY2, 398 aa), which has previously shown to target PGs. As results, HR1 includes the crucial sequence (31–45th aa) for chloroplast import and the stromal cleavage occurs at a specific alanine site (64th aa) within HR2, verifying that a N-terminal 64-aa-region works as the transit peptide (Tp). HR2 has a weak PG-targeting signal by showing synchronous and asynchronous localization patterns in both PGs and stroma of chloroplasts. HR3 exhibited a strong PG-targeting role with the required positional specificity to prevent potential issues such as non-accumulation, aggregation, and folding errors in proteins. Herein, we characterized a Tp and two transmembrane domains in three HRs of OsPSY2 and propose a spontaneous pathway for its PG-translocation with a shape embedded in the PG-monolayer. Given this subplastidial localization, we suggest six sophisticated tactics for plant biotechnology applications, including metabolic engineering and molecular farming.

KEYWORDS

carotenoid, cleavage site, hydrophobic helix, transmembrane domain, transit peptide

Introduction

Chloroplasts are organelles where photosynthesis occurs and that are vital for plant survival. They also participate in metabolic processes for carbohydrates, amino acids, lipids, vitamins, and isoprenoids (Menke, 1962; Chen et al., 2018). Chloroplasts consist of subcellular aqueous regions, such as the stroma and thylakoid lumen, and various membranous structures, including the outer and inner envelope, internal thylakoid

membrane, and plastoglobules (PGs). PGs are lipid monolayer droplets associated with the thylakoid membranes (Bailey and Whyborn, 1963; Greenwood et al., 1963; Rottet et al., 2015). The proteins found in various chloroplast structures, excluding the outer envelope, possess a transit peptide (Tp) that is susceptible to cleavage in the precursor N-terminal region. The Tp facilitates protein localization from the cytoplasm to the stroma *via* consecutive traversal of two envelopes through one translocon at the outer envelope membrane of chloroplasts and another translocon at the inner envelope membrane of chloroplast complexes and then is cleaved by a stromal processing peptidase. The cleaved mature proteins have different fates: some remain in the stroma and some are embedded in the inner envelope or engage in secondary transportation-mediated translocation toward the thylakoid or thylakoid-derived structures such as PGs in the help of additional signals (Jarvis and Robinson, 2004; Jarvis, 2008; Paila et al., 2015; Xu et al., 2021).

To precisely sort proteins to thylakoidal structures, several independent routes have been identified, including the chloroplast secretion (cpSec) and chloroplast twin-arginine translocation (cpTat) pathways for thylakoid lumen proteins, the chloroplast signal recognition particle (cpSRP)-dependent pathway, the chloroplast spontaneous pathway, and the chloroplast Guided Entry of tail-anchored proteins (cpGET) pathway for thylakoid membrane proteins (Jarvis and Robinson, 2004; Zhu et al., 2022). In general, thylakoid lumen proteins distinctly use the cpSec (when unfolded) and cpTat (when folded) pathways. Moreover, they also have different types of luminal targeting peptides that are cleaved by specific thylakoidal processing peptidases (TPP) (Palmer and Stansfeld, 2020; Xu et al., 2021). Some thylakoid membrane proteins, such as light-harvesting chlorophyll *a/b*-binding proteins (LHCBs), use the cpSRP pathway, which requires a stromal protein complex and a thylakoid membrane-bound integrase. Many proteins, including photosynthetic subunits such as PsaK and PsbW are spontaneously targeted without requiring additional proteins or energy input (Mant et al., 2001; Woolhead et al., 2001; Jarvis and Robinson, 2004; Xu et al., 2021). Recent research has suggested that *Arabidopsis* Get3b can recognize tail-anchored cargo proteins, which have a single transmembrane (TM) domain in their C-terminus followed by a tail of 30 amino acids, and help to deliver them into the thylakoid membrane *via* the cpGET pathway (Anderson et al., 2021; Zhu et al., 2022). In addition, thylakoid membrane proteins generally have different numbers of TM domains in the mature protein region to span the membrane. For example, there are three TM domains in LHCB1, a protein linked to the cpSRP pathway, two TM domains in PsaK, which participates in the spontaneous pathway, and one TM domain in PsbW, which also features a cleavable signaling peptide that is cleaved by TPP inside the thylakoid lumen (Jarvis and Robinson, 2004; Xu et al., 2021; Zhu et al., 2022). This suggests that the specific translocation pathway to be used by a thylakoid membrane-bound protein may depend on the number of TM domain which typically consists of 17-25 hydrophobic amino acids in an α-helix structure (Engelman et al., 1986; Pasquier et al., 1999; White and von Heijne, 2005).

PGs have been recognized as a hub for metabolic activities related to isoprenoid-derived lipids, including phylloquinone, plastoquinone, tocopherols, carotenoids, and apocarotenoids, as well as neutral lipids such as galactolipids, phospholipids, sterol esters, triacylglycerols, and free fatty acids through metabolomic analyses using a diverse array of analytical tools to examine various plastid types (Steinmuller and Tevini, 1985; Vidi et al., 2006; Gaude et al., 2007; van Wijk and Kessler, 2017). Protein profiling has also confirmed the presence of specific metabolic enzymes—including esterases, lipases, and thioesterases—that are involved in the metabolism of neutral lipids. Other enzymes, including ubiE-methyltransferase-related enzymes and ABC1 kinase, which are involved in quinone metabolism, tocopherol cyclase, and carotenoid cleavage dioxygenase 4 have also been identified in PGs in *Arabidopsis* chloroplasts (Vidi et al., 2006; Ytterberg et al., 2006; Rubio et al., 2008; Joyard et al., 2009; Rottet et al., 2016; van Wijk and Kessler, 2017). In addition, one study of saffron found that expression patterns calibrated to the highest levels of β-carotene and β-ionone release were induced during stigma development and that carotenoid cleavage dioxygenase 4 was localized to PGs (Rubio et al., 2008). Another study of *Capsicum* chromoplast PGs confirmed the presence of ζ-carotene desaturase (ZDS), lycopene β-cyclase (*β*-LCY), and β-carotene hydroxylase (CrtR-*β*) during carotenoid biosynthesis (Ytterberg et al., 2006). These results strongly imply that PGs might be the crucial site where lipid metabolites such as isoprenoids (including carotenoids) undergo biosynthesis, degradation, accumulation, and metabolic communication with thylakoids during plant development and environmental adaptation (Brehelin and Kessler, 2008; van Wijk and Kessler, 2017).

Phytoene synthase (PSY), the first and rate-limiting enzyme involved in the carotenoid biosynthesis pathway, has been found to exhibit diverse chloroplast localization, including in the stroma, thylakoid membrane, and PGs, in maize, rice, saffron, and cabbage (Shumskaya et al., 2012; You et al., 2017; Ahrazem et al., 2019; Cao et al., 2021). For example, PG localization of three rice PSYs (OsPSYs) has been reported using a maize protoplast system (Shumskaya et al., 2012). Moreover, our group has previously shown that the OsPSY2 (LOC_Os12g43130; NP_001391692) localizes to PGs in both etiolated and greening rice protoplasts (You et al., 2017). In this study, we predicted that OsPSY2 contains three putative TM helix domains with high hydrophobicity *via* bioinformatics approaches, including querying of four public databases (i.e., the ChloroP 1.1 Server, TMpred, TopPred, and HMMTOP) and the N-terminal 80-amino-acid portion of OsPSY2 (PTp) harbors two putative TM helices with dual localization patterns of fluorescent proteins as cargo, showing speckle-like spots on PGs with a slightly blurred background signal dispersed over the stroma of chloroplasts as a PG-preferential Tp sequence (You et al., 2017). Since the publication of this paper, we hypothesized that the first and second putative TM helix structures in the PTp may be associated with the Tp signal and an additional signal that is partly crucial for PG-targeting of OsPSY2, respectively.

To date, despite there being many cases of PG-targeted proteins, the PG-targeting mechanism has remained largely uncharacterized (Ytterberg et al., 2006; Rubio et al., 2008; Shumskaya et al., 2012; Rottet et al., 2016; You et al., 2017). We determined that OsPSY2 is one of the

most promising candidates for identify PG-targeting signals, so we therefore dissected its chloroplast targeting region and membrane spanning domains by generating various combinatory constructs using two fluorescent proteins—i.e., synthetic green fluorescent protein (sGFP) and red fluorescent mCherry protein—as cargo to be delivered. *Via* microscopic fluorescent observation and molecular biological analysis, we elucidate the PG-localization mechanism of OsPSY2 in chloroplasts, and propose strategies for achieving enhanced precision in protein targeting to specific sub-organelle compartments. This precise targeting approach holds great potential for various applications in the field of plant biotechnology.

Materials and methods

Vector construction

We previously described the six fluorescent gene fragments such as sGFP (sG), mCherry (mC), a transit peptide of rice RuBisCO small subunit 1 (RTp)-sG, PTp-sG, OsPSY2:mC, and OsPSY2(ΔPTp):mC, which were cloned into *pB2GW7* vector for plant transformation (Dubin et al., 2008; You et al., 2015; You et al., 2017). Using the same system, we first combined RTp and OsPSY2 (ΔPTp):mC to generate a RTp-OsPSY2(ΔPTp):mC fragment with overhung recombination sites (attB1/attB2). Next, we PCR-amplified hydrophobic regions (HRs) to obtain new five fragments that contained only one of the following configurations of three hydrophobic helixes (HHs) of OsPSY2. These included HR1, short HR1 (sHR1), HR2, extended HR2 (eHR2), and HR3. Eleven sGFP or mCherry chimeric gene fragments that each included one of these HR fragments and either RTp or PTp are as follows: two for HH1 (no RTp or PTp), one for HH2 (no RTp or PTp), three for HH2 (with RTp), two for HH3 (with RTp), and three for HH3 (with PTp). HR2 was additionally split inside the HH2 structure into front (HR2f) and back (HR2b) halves. Each of these fragments also had overhung attB1/attB2 sites. All PCR reactions were conducted using Phusion® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) and KOD FX High Success-rate DNA Polymerase (Toyobo, Osaka, Japan). We introduced all twelve gene fragments into a *pB2GW7* vector *via pDONR221* by consecutive recombination using Gateway® BP Clonase® II Enzyme Mix and Gateway® LR Clonase® II Enzyme Mix. All procedures followed the manufacturer's instructions (Invitrogen, Waltham, MA, USA).

Rice protoplast isolation and PEG-mediated transfection

Rice (*Oryza sativa* L. Japonica cv. "Ilmi") seeds were harvested at full maturity (60 days after flowering), were dehusked using a TR-200 Electromotion rice husker (Kett, Tokyo), and were then sterilized by serial treatment of 70% ethanol (once for 1 min), 2% sodium hypochlorite (twice for 20 min), and distilled water (five times). Seeds were then left to germinate on Murashige and Skoog agar medium (Duchefa, Haarlem, Netherlands). Once germinated, they were grown at 28°C for nine days in the dark then for 20 hours under light conditions in a plant growth chamber. Using these seedlings, protoplast isolation was conducted as per our previous description (Bart et al., 2006; You et al., 2015) with minor modifications. In brief, the leaf sheaths of seedlings were cut to 0.5 mm pieces with a sharp razor blade then promptly immersed in a cell wall-degrading enzyme solution (i.e., containing Cellulase R-10 and Macerozyme R-10; Yakult Honsha, Japan) with additional ingredients (i.e., 4 mM $CaCl_2$ and 50 μg/mL ampicillin). Plates with rice strips soaked in solution were placed in a desiccation chamber for vacuum infiltration for 10 minutes then incubated with gentle shaking at 50 rpm in the dark at room temperature for 4 hours. This enzyme reaction was stopped by adding three volumes of W5 solution (i.e., 6 mM glucose, 154 mM NaCl, 125 mM $CaCl_2$, 10.7 mM KCl, and 2 mM MES adjusted to pH 5.7). After removal of cell debris by filtering twice through 70 and 40 μm cell strainers (SPL, Pocheon, South Korea), rice protoplasts were acquired from the pellet and suspended in MMG buffer (i.e., 0.6 M mannitol, 15 mM $MgCl_2$, and 5 mM MES adjusted to pH 5.7) at a concentration of 10^7 cells/mL.

For transfection, 10^6 cells were counted using a Marienfeld hemocytometer counting system (Marienfeld-Superior, Berlin, Germany). These aliquots were used for each transfection of plasmid DNA (10 μg) accompanied by an equal volume of 40% PEG-3350 (Sigma, St. Louis, MO, USA) contained in a 0.4 M mannitol and 100 mM $Ca(NO_3)_2$ solution and incubated for 15 minutes. The PEG-mixed protoplasts were then washed twice with two volumes of W5 solution and 1 mL of W5 solution, then resuspended in 1 mL of incubation solution, and incubated at 28°C in the dark overnight. During the transfection experiments, all plastic products were coated with 5% fetal calf serum for 10 seconds and all buffers were filtered using a 0.45μm Sartorius Minisart® syringe filter (Sartorius, Gottingen, Germany).

Confocal microscope analysis

To observe fluorescence signals indicating subcellular and suborganellar localization in the rice protoplasts, we used Carl Zeiss LSM700 and Carl Zeiss LSM800 inverted confocal microscopes (Carl Zeiss, Oberkochen, Germany). Fluorescence was detected at two excitation and emission wavelengths, i.e., 488 and 509–518 nm for sGFP, 587 and 585–610 nm for mCherry, and 639–680 and 660–700 for chlorophyll autofluorescence, respectively. Fluorescence images were captured using ZEN 2009 Light Edition and ZEN 2.3 Blue Edition software (Carl Zeiss).

Molecular analysis

The same transfected protoplasts used for fluorescence observations (i.e., approximately 10^6 cells) were then subjected to protein and RNA extraction. Following microscopy, cells were pelleted, immediately frozen in liquid N_2, and stored at −80°C until extraction. To extract total protein, rice protoplasts were

suspended in 50 µL of 2× Laemmli Sample Buffer (GenDEPOT, Barker, TX, USA), placed in a hot water bath for 10 minutes, then centrifuged at 4°C. Next, 10 µL of the supernatant was loaded onto a 12% acrylamide gel for protein separation *via* SDS-PAGE. Next, the gel was transferred to an Immun-Blot® PVDF Membrane (Bio-Rad) using a Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad). Concurrently, a duplicate gel was stained using an EZ-Silver Staining Kit for Protein (Biosesang, Seongnam, Korea) to quantify protein content and estimate the quality of protein extracts. Next, a Western blot was performed using an anti-rabbit polyclonal antibody against GFP (Abcam, Cambridge, UK), an antibody against mCherry (Novus Biologicals, Denver, USA), and an antibody against actin (Abcam, Cambridge, UK). An Anti-Rabbit IgG HRP Conjugate (Promega, Madison, WI, USA) was used as a secondary antibody to react with the primary antibody. Two primary antibodies against GFP and actin were used at dilutions of 1:10,000, and the mCherry antibody was diluted to 1:5000. After incubation with West-Q Femto Clean ECL Solution (GenDEPOT), immunoreactive bands were detected and imaged using an Alliance Mini HD camera (UVITEC, Cambridge, UK).

Total RNA was then extracted from the same rice protoplast samples using RNeasy® Plant mini kits (Qiagen, Hilden, Germany). First strand cDNA was then synthesized using AccuPower® RT Premix (Bioneer, Daejeon, Korea) with all procedures performed according to the manufacturer's instructions. Quantitative real-time PCR (qRT-PCR) was then performed with the following gene-specific Fwd/Rev primer sets: 5′-GACAAGCAGAAGAACGGCATCA-3′/5′-GGCGGCGGTCACGAACT-3′ for sGFP and 5′-GAAGTAAGGAAGGAGGAGGA-3′/5′-AAGGTGTTCAGTTCCAAGG-3′ for rice ubiquitin (AK061988), which was used as a reference to normalize RNA content. Three technical replicates were performed for each condition. The reaction conditions used for qRT-PCR followed a method that was previously described (You et al., 2017).

Results

Generation of recombinant proteins evaluating subcellular localization roles of three hydrophobic regions in OsPSY2

In a previous study, we reported that the PTp sequence facilitates the simultaneous translocation of foreign proteins into PGs and stroma within chloroplasts. This suggests that there is a need for additional signaling for the intact PG localization of OsPSY2 (You et al., 2017). Based on the predicted TM helix sequences with high hydrophobicity (Supplementary Figure S1), we designated three HH structures as HH1 (1–21st aa), HH2 (48–66th aa), and HH3 (233–252nd aa). We also designated the partitioned a PTp (1–80th aa), and two sets of two regions that covered either HH1 or HH2 as follows: 'HR1 (1–45th aa) and HR2 (46–80th aa)' and 'sHR1 (1–30th aa) and eHR2 (31–80th aa)'. Finally, we designated HH3 as 'HR3 (229–247th aa)'. These residue locations corresponded to a commonly proposed region found in multiple databases (Figure 1A). Next, to accurately interpret the results, we used six vectors to express fluorescent proteins such as sGFP or mCherry under the 35S promoter as subcellular and suborganellar markers; these vectors included: sG and mC for the cytosol, RTp-sG for the stroma, PTp-sG for simultaneous but preferring PGs than stroma, OsPSY2:mC for PGs, and OsPSY2 (ΔPTp):mC for the cytosol (You et al., 2015; You et al., 2017) (Figure 1B). Thus, in this study, we produced a total of fourteen recombinant proteins that should show sGFP or/and mCherry fluorescence; this data helped us to assess whether HH is likely to be a subplastidial targeting signal. We therefore assessed the following possible combinations: two for HH1, four for HH2, three for HH3, and another three for HH3 together with both HH1 and HH2 as the PTp (Figure 1B) and additional two for HH2 cut in half front to back (Supplementary Figure S2A).

Chloroplast import property in N-terminal sequences including hydrophobic region 1

Localization of OsPSY2 in PGs within chloroplasts and the fact that its cleaved Tp length is shorter than 80 aa (i.e., as predicted by ChloroP) has previously been examined in maize and rice protoplasts (Shumskaya et al., 2012; You et al., 2017) and pea and rice chloroplasts (Welsch et al., 2008; You et al., 2017), respectively. To determine how OsPSY2 enters the stroma before PGs, two N-terminal regions including HH1 (i.e., HR1 and sHR1) were fused with sGFP and individually transfected into green rice protoplasts from 10-day-old seedlings. Their green fluorescent signals showed different pattern in the stroma and cytosol as matched with the RTp-sG and sG stromal and cytosolic localization markers, respectively (Figure 2A). This data therefore suggests that HR1 was responsible for the chloroplast targeting of OsPSY2 and also that the sHR1 lost its function as Tp sequence. This in turn suggests that the 31–45th aa position in the HR1 may be essential for translocation into chloroplasts.

To confirm whether this HR1 was involved in cleavage after translocation into chloroplasts, we performed a Western blot analysis to determine the size of sGFP (Figure 2B). The original GFP size of 27 kDa was detected in RTp-sG and sG, whereas HR1-sG generated a much larger GFP band (~32 kDa), suggesting that this is its uncleaved size. We conclude that the HR1 has no cleavable site. To bioinformatically predict the cleavage position of OsPSY2, we then analyzed its protein sequence using the TargetP algorithm (https://services.healthtech.dtu.dk/service.php?TargetP) (Almagro Armenteros et al., 2019). This analysis generated three proposed alanine (A) positions that show high cleavage probability: i.e., A43 in HR1 and A64 and A70 in the HR2 (Figure 2C and Supplementary Table S1). These data suggest that OsPSY2 is more likely to be truncated in HR2 and not in the HR1, since A43 is experimentally excluded (Figure 2B).

The role of hydrophobic region 2 as a partial PG-targeting signal and for cleavage in chloroplasts

To acertain the influence of HH2 on the translocation of OsPSY2, we cotransfected green rice protoplasts with PTp-sG or

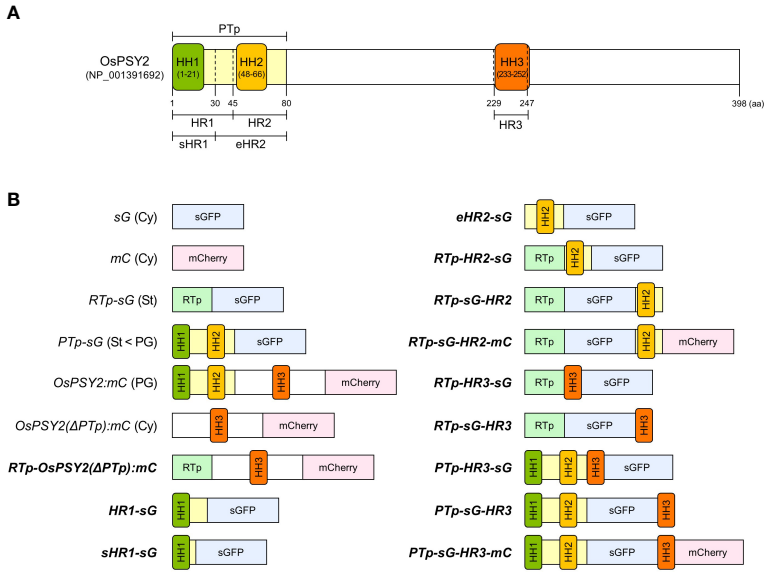


FIGURE 1
Schematic representation of the OsPSY2 protein with positional information for each region used in this study as well as various fluorescent protein combinations used for rice protoplast transfection. **(A)** OsPSY2 is denoted by the predicted amino acid positions of a transit peptide as predicted by the ChloroP 1.1 Server and hydrophobic regions (HRs) individually containing three hydrophobic helices (HHs) that were identified using the TMpred, TopPred, and HMMTOP databases. **(B)** All chimerically recombined gene fragments including sGFP or/and mCherry are shown in italic for this study. The newly manipulated ones are in boldface (hereafter in all figures). All recombined DNA fragments were cloned into a plant expression vector (i. e., *pB2GW7*) for protein expression in rice protoplasts. OsPSY2, *Oryza sativa* phytoene synthase 2; PTp, the prediction-based N-terminal cleavable transit peptide of OsPSY2; OsPSY2(ΔPTp), OsPSY2 with PTp removed; sHR1, short HR1; eHR2, extended HR2; sG, sGFP; mC, mCherry; RTp, the transit peptide of RuBisCO small subunit 1; Cy, cytosol; St, stroma; PG, plastoglobule.

one of three new recombinant fluorescent proteins harboring regions containing HH2 with either mC or OsPSY2:mC; we also transfected protoplasts with RTp-sG-HR2-mC alone (Figure 3A). The eHR2 resulted in cytosolic localization, confirming that it alone was not functional as Tp despite containing HH2 and a 30-aa N-terminal region including HH1 (i.e., sHR1) might necessitate first entry into chloroplasts again. To understand the unique features of putative HR2 in PTp, we used RTp instead of HR1 and placed HR2

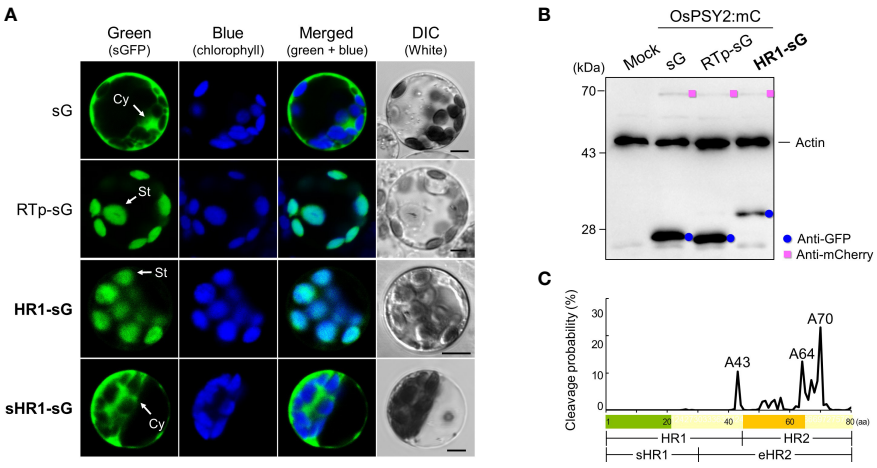


FIGURE 2
Effect of two N-terminal regions of the OsPSY2 protein including the hydrophobic helix (HH) 1 on subcellular localization in rice protoplasts. **(A)** Individual and merged images of sGFP, chlorophyll, and visible protoplasts using confocal microscopy. sG and RTp-sG were used as a cytosolic (Cy) and stromal (St) localization markers, respectively. Scale bars indicate 5 μm. **(B)** Western blot analysis to compare the amount and size of sGFP. Rice protoplasts were cotransfected with each experimental vector and an OsPSY2:mC vector as a transfection control. This blot was performed with three antibodies simultaneously: Anti-GFP (blue circle), anti-mCherry (pink rectangle), and anti-rice actin (marked at 45 kDa, to be used as an internal reference of the total protein amount). **(C)** Prediction of cleavage probability in the N-terminal 80-amino-acid region of OsPSY2 as implemented by the TargetP 2.0 program (https://services.healthtech.dtu.dk/service.php?TargetP). Estimated cleavage scores of the 80 amino acids of PTp are shown in Supplementary Table S1.

10.3389/fpls.2023.1181311

both before and after sG, thus generating both RTp-HR2-sG and RTp-sG-HR2. The green fluorescence signal of RTp-HR2-sG spread throughout the stroma in discrete specks. This punctuate pattern exactly overlapped with the red PG signals of OsPSY2:mC when cotransfected. The overall dual targeting pattern of RTp-HR2-sG (i.e., including both the stroma and PGs) was similar to PTp-sG, providing evidence that HR2 may be involved in PG-targeting. In contrast, the green fluorescence signal of RTp-sG-HR2 was displayed only in the stroma, indicating that HR2 placement is dependent on RTp. To verify this finding, we generated an additional collocation of mC in the C-terminus to form RTp-sG-HR2-mC. We then observed two asynchronous patterns of either stroma or PGs in individual cells. Interestingly, we found that when RTp (used as a substitute of HR1) and HR2 were close and not far, as in RTp-HR2-sG and PTp-sG, they exerted the dual-target pattern. However, when they were spaced further apart, as in RTp-sG-HR2 and RTp-sG-HR2-mC, they displayed a single-target pattern. This suggests that RTp/HR1 and HR2 may be in competition and also that the position of HR2 in cargo proteins works crucially as a PG-targeting signal: i.e., at the N-terminus as in RTp-HR2-sG and PTp-sG, in the middle as in RTp-sG-HR2-mC, but never at the C-terminus as in RTp-sG-HR2.

To determine whether and where cleavage occurs for OsPSY2, we conducted Western blots using rice protoplasts and compared the sizes of fluorescent proteins produced (Figure 3B). These results showed that the uncleaved GFP band of eHR2-sG was coincident with a failure to target chloroplasts. Moreover, it also showed the same cleaved size GFP bands from RTp-HR2-sG and PTp-sG were greater than the 28 kDa protein marker as well as the mature GFP produced by RTp-sG. This proves that the cleavage occurred in HR2 in both RTp-HR2-sG and PTp-sG and that the mature GFP size was equivalent to the theoretical size following cleavage at the A64 position. This means that the estimated size is either 29 or 28 kDa if cleaved at either A64 or A70 in HR2. Moreover, the fact that RTp-sG-HR2 displayed a bigger GFP band than PTp-sG (which was equivalent to RTp-HR2-sG) confirmed that RTp-HR2-sG had a second cleavage in HR2 following the first cleavage of RTp. Thus, cleavage occurred once in the HR2 of PTp-sG, twice in RTp and the HR2 of RTp-HR2-sG, and once in RTp of RTp-sG-HR2 and RTp-sG-HR2-mC (Figure 3C).

To further verify which region is involved in PG-targeting and whether cleavage actually occurs at the A64 position, a HR2 was divided into two of a 21-aa in the front (HR2f, 46-66th aa) and a 17-aa on the back (HR2b, 64-80th aa) and recombined them with RTp and sGFP in the N- and C-terminus, respectively (Supplementary Figures S2A, B). The cleavage positions of RTp-linked protein sequences were commonly predicted the probability of A64 in HR2 with cysteine (C) 48 in RTp region (Supplementary Figure S2C). The subplastidial localization showed different fluorescent patterns between RTp-HR2f-sG and RTp-HR2b-sG: the former in only stroma and the latter in PGs and stroma with similarity to PTp-sG, confirming that PG-targeting ability of a PTp might be due to the 16 aa-region on the back of HR2 after cleavage at the A64 (Supplementary Figure S2D). Western blot also confirmed the occurrence of actual cleavage within HR2f showing smaller GFP band in RTp-HR2f-sG (27 kDa) than RTp-HR2-sG (29 kDa) as predictably cleaved sizes at the A64 (Supplementary Figure S2E).

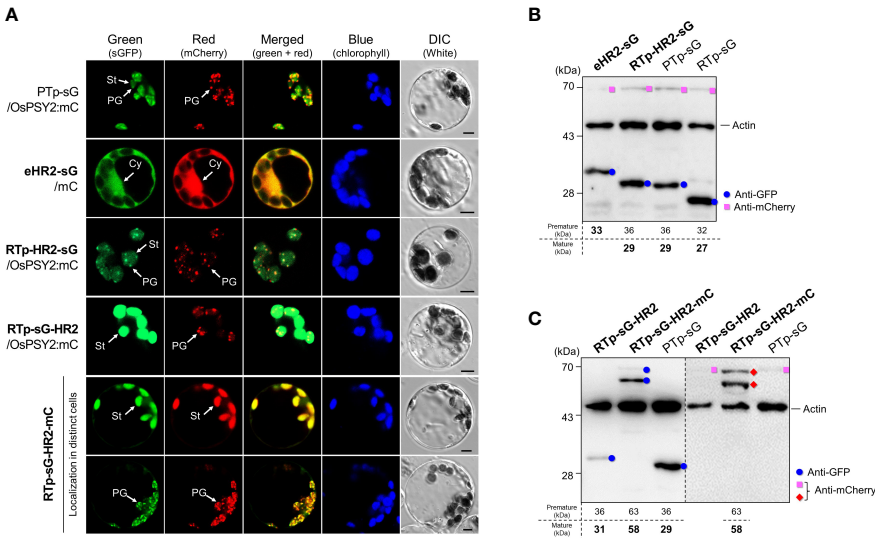


FIGURE 3
Effect of the hydrophobic region (HR) 2 on subcellular and suborganellar localization in rice protoplasts. **(A)** Individual and merged images of sGFP, mCherry, chlorophyll, and visible protoplasts using confocal microscopy. PTp-sG, OsPSY2:mC, and mC were used as markers for dual localization into plastoglobules (PGs) and stroma (St) as well as single localization into a PG and the cytosol (Cy), respectively. Scale bars indicate 5 μm. **(B)** Western blot analysis to confirm the cleavage in HR2 when placed right after a Tp signal such as RTp or HR1 within PTp. **(C)** Western blot analysis to confirm that no cleavage occurred in HR2 when placed at the C-terminus of sGFP. In both **(B, C)**, rice protoplasts were cotransfected with an experimental vector and an OsPSY2:mC vector used as a transfection control. In one exceptional case we used RTp-sG-HR2-mC with intrinsic mC. The blot shown in **(B)** was performed with three antibodies simultaneously: Anti-GFP (blue circle), anti-mCherry (red diamond for the experimental vector or pink rectangle for the control vector), and anti-rice actin (marked at 45 kDa and used as an internal reference indicating total protein amount). The blot shown in **(C)** was performed with two antibodies simultaneously: Anti-GFP and anti-rice actin in the left panel, and anti-mCherry and anti-rice actin in the right panel, respectively. The expected protein size is noted under the blot.

Taken together, our results indicate that HR2 containing HH2 requires a front-facing domain containing another HH structure, such as RTp or HR1 of PTp, for plastidial cleavage and chloroplast import. With respect to PG localization, HR2 can play a part unless it is positioned at the C-terminus of a plastidial cargo protein.

Role of hydrophobic region 3 as a novel PG-targeting signal requiring a precise location

In a previous study, we hypothesized that a signal other than PTp may be required for complete PG targeting, since there was a discrepancy in subplastidial localization of PTp, which was simultaneously localized to PGs and the stroma, and OsPSY2, which localized only to PGs (You et al., 2017). To confirm the presence of this signal, we prepared a new RTp-OsPSY2(ΔPTp):mC and compared its localization pattern to OsPSY2:mC and OsPSY2 (ΔPTp):mC (Figure 4A). We found that the red fluorescence signal of RTp-OsPSY2(ΔPTp):mC showed the same PG-localization pattern of OsPSY2:mC, confirming our assumption that there is another PG-targeting signal in OsPSY2. In addition, the RTp-cleaved mature protein was also confirmed as having the same size as the cytosolic-targeted OsPSY2(ΔPTp):mC (Figure 4B).

At this point, we speculated whether HH3 works as a novel PG-signal located in the mature region of OsPSY2. Therefore, we contrived two recombinant proteins to collocate HR3 before and after sG, obtaining RTp-HR3-sG and RTp-sG-HR3. In rice protoplasts, RTp-sG-HR3 displayed a PG-targeting pattern, which was consistent with that of the cotransfected OsPSY2:mC (i.e., showing specks dotted within chloroplasts). This pattern was totally different than RTp-sG, which showed stromal localization, as well as RTp-HR3-sG, which had no fluorescent signal (Figure 4C). To further analyze these proteins on the molecular level, we examined protein production *via* Western blot (Figure 4D). Compared to the cleaved 27 kDa GFP band of RTp-sG, RTp-sG-HR3 produced a cleaved 29 kDa GFP band but RTp-HR3-sG formed only a weak premature GFP band. Additional quantitative real-time-PCR was performed to examine whether there were differences in transcript levels (Figure 4E). Interestingly, all three showed similar levels of normal GFP transcription despite different patterns of fluorescence and protein expression. These findings suggest that HR3 causes no protein production if located at the N-terminus of a cargo protein, where

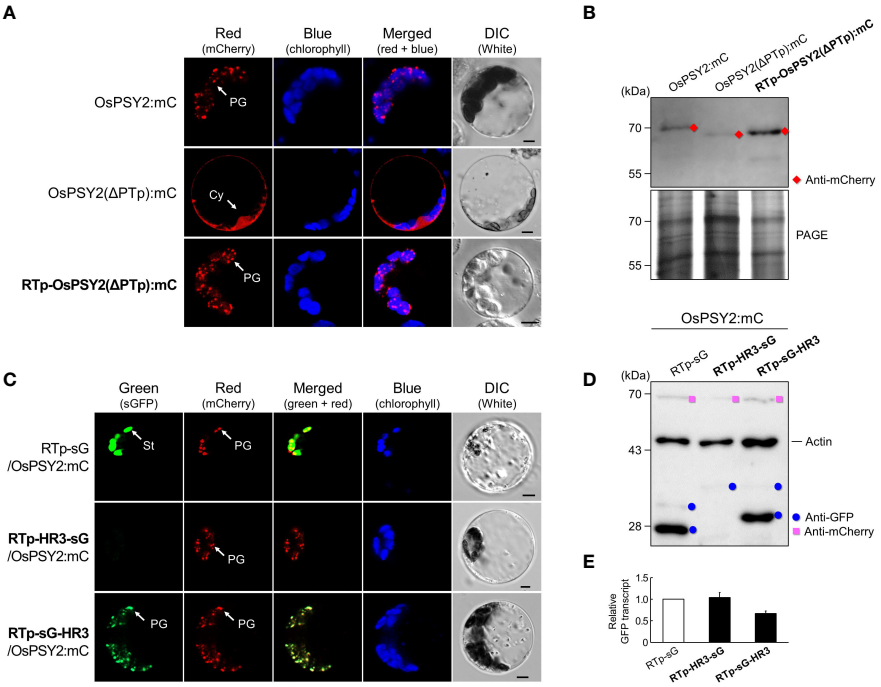


FIGURE 4
Effect of the hydrophobic region (HR) 3 on PG localization in rice protoplasts. **(A)** Individual and merged images of mCherry, chlorophyll, and visible protoplasts using confocal microscopy. OsPSY2:mC and OsPSY2(ΔPTp):mC were used as localization markers of plastoglobules (PGs) and the cytosol (Cy), respectively. Scale bars indicate 5 μm. **(B)** Western blot analysis to confirm the presence of a PG-targeting signal in an area other than PTp in OsPSY2. Rice protoplasts were individually transfected with three experimental vectors. This blot was conducted with anti-mCherry (red diamond) and a polyacrylamide gel electrophoresis (PAGE) image is shown to indicate total protein amounts. **(C)** Individual and merged images of sGFP, mCherry, chlorophyll, and visible protoplasts using confocal microscopy. RTp-sG and OsPSY2:mC were used as localization markers for the stroma (St) and plastoglobules (PGs), respectively. Scale bars indicate 5 μm. **(D)** Western blot analysis to reveal the precise location of HR3 that exhibits PG-targeting capacity. Rice protoplasts were cotransfected with an experimental vector and an OsPSY2:mC vector, used as a transfection control. The blot was performed with three antibodies simultaneously: Anti-GFP (blue circle), anti-mCherry (pink rectangle), and anti-rice actin (marked at 45 kDa and used as an internal reference of total protein amount). **(E)** Quantitative real-time-PCR to indicate relative transcript levels depending on different protein levels for the three vectors. The rice ubiquitin gene (AK061988) was used as a reference to normalize RNA amounts. qRT-PCR was performed using three technical replicates with three biological samples per vector.

it is placed side-by-side with RTp, but also that HR3 alone can exert a near-perfect PG-targeting role if located at the C-terminus of a chloroplastidial protein when translocated by a simple Tp such as RTp.

The mode of hydrophobic region 3 action for PG localization under the influence of hydrophobic region 1 and 2

In addition to the PG-targeting role of HR3 in collaboration with RTp, we also examined the characteristics of HR3 when towed by PTp. To do so, we transfected rice protoplasts with three recombinant proteins that included HR3 in different places around fluorescent cargo proteins: i.e., the N-terminus of sG, the C-terminus of sG, and in between sG and mC (Figure 5A). In the PTp-HR3-sG case, we identified two different types of green fluorescence, either aggregation or (rarely) a PG-targeting pattern. This matched the pattern of cotransfected OsPSY2-mC. At the same time, PTp-sG-HR3 did not show any fluorescent signal. Taken together, these data suggest that the close arrangement of three HHs in a row easily forms fluorescent aggregates and also that the 'trapped' pattern in which sGFP is between two HHs disturbs fluorescent expression. Next, we tried to reproduce the configuration of OsPSY2 *via* mC-fusion at C-terminus of PTp-sG-HR3, generating PTp-sG-HR3-mC. Observation of its red fluorescent signal revealed PG localization. This presented evidence that HR3 has a PG-targeting ability and requires a precise location to be functional.

To further understand the protein dynamics at play, we performed subsequent Western blots and observed that the fluorescent proteins corresponding to sG and mC (e.g., PTp-sG-HR3-mC) were normally produced at the mature size (Figure 5B). This finding revealed that the barely observed fluorescent green signals were caused by posttranslational effects; e.g., clumping that caused relatively lower amounts of GFP production for PTp-HR3-sG and folding errors despite normal GFP production for PTp-sG-HR3 and PTp-sG-HR3-mC. In addition, the mC contained in PTp-sG-HR3-mC showed successful PG-targeted red fluorescence with normal protein expression. It implies that the HR3 position does not affect PG targeting itself but that there must be distance between HR2 and HR3 and that HR3 must not be positioned at the C-terminus of a cargo protein when translocated by PTp containing HR1 and HR2.

Discussion

Since the reports of "osmiophilic globules" in faba bean, spinach, and beet chloroplasts (Bailey and Whyborn, 1963; Greenwood et al., 1963), PGs have been identified as active subcompartments. That is, rather than passive storage droplets for lipid metabolism including carotenogenesis, they engage in two-way communication with the thylakoid membrane (Austin et al., 2006; Brehelin and Kessler, 2008; Zita et al., 2022). In addition that the presence of carotenoid biosynthetic enzymes such as ZDS, β-LCY, and CrtR-β with high levels of ketoxanthophylls such as capsanthin and capsorubin has been found in the chromoplast PGs of red bell pepper (Ytterberg et al., 2006; van Wijk and Kessler, 2017), the enriched accumulation of carotenogenic enzymes such as PSY (i.e., by 7.5-fold), phytoene desaturase (7.2-fold), carotenoid isomerase (12-fold) was reported alongside large increases of lycopene and β-carotene in chromoplasts relative to chloroplasts in the PG proteome and metabolome of tomato (Zita et al., 2022). These results strongly support the notion that the PG-localized carotenoid biosynthetic enzymes, which includes PSY, might directly contribute to carotenoid biosynthesis. Previously, Shumskaya et al. (2012) and our group provide evidence of the PG localization of OsPSY2 (You et al., 2017) and we suggested that the prediction-based Tp region of OsPSY2 drove the dual

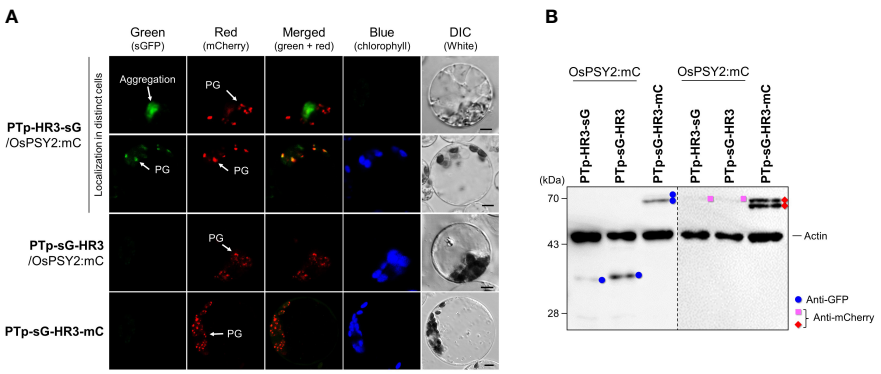


FIGURE 5
Differential effects of the hydrophobic region (HR) 3 on the PG-localization depending on PTp in rice protoplasts. **(A)** Individual and merged images of sGFP, mCherry, chlorophyll, and visible protoplasts using confocal microscopy. OsPSY2:mC was used as a plastoglobule (PG) localization marker. Scale bars indicate 5 µm. **(B)** Western blot analysis to determine the factors influencing aggregation and loss of fluorescence. Rice protoplasts were cotransfected with an experimental vector and an OsPSY2:mC vector used as a transfection control. In one exceptional case PTp-sG-HR3-mC was used, which has intrinsic mC. The blot was performed with two antibodies simultaneously: Anti-GFP (blue circle) and anti-rice actin (marked at 45 kDa) on the left panel, and anti-mCherry (red diamond for experimental vector or pink rectangle for the control vector) and anti-rice actin on the right panel, respectively.

localization of this protein to PGs and the stroma. This finding left room for further study to elucidate PG-targeting signals in areas other than PTp (Figure 1A).

In this study, we generated a variety of recombinant proteins with stepwise combinations of three HRs of OsPSY2 (Figure 1B). Using these recombinant proteins, we investigated protein fluorescence and expression in rice protoplast systems. First, we identified the critical region for chloroplast targeting as a 15-aa-peptide in the region between residues 31–45 located between HH1 and HH2 (Figures 2A, B). We also identified cleavage site A64 positioned at the end of HH2 (Figures 2C and 3B). Next, we determined that the factual Tp of OsPSY2 is a 64-aa-long protein including HH1 and most of HH2; this is compared to the 80-aa-long sequence of PTp identified *via* program-driven prediction (Figure 6A). After translocation into a chloroplast, OsPSY2 might undergo a secondary translocation into PGs. To understand this mechanism, HR2 and HR3 were individually relocated at either the N- or C-terminus of sGFP and in a middle location between sGFP and mCherry at a position after stroma cleavage (Figure 1B). We attested the weak PG-targeting capacity of HR2 to positional effects caused by different influences on HR1 and RTp that result in three patterns of dual localization into PG (by HR2) and the stroma (by HR1 or RTp). These patterns were biased toward PGs for PTp-sG, toward the stroma for RTp-HR2-sG and was balanced for RTp-sG-HR2-mC (Figure 3A). Here we determined that for HR2, most of HH2 is cut out by stromal processing at A64; the remaining 16-aa-region hangs at the N-terminus of the mature protein. The possibility that this region may be more influential than HH2 itself as a PG-targeting signal was also verified (Supplementary Figure S2). In addition, the capacity of HR3 to drag cargo proteins into PGs is evidently greater than HR2; this conclusion is supported by our observation of clear PG-targeting patterns RTp-sG-HR3 (visualized using GFP) and PTp-sG-HR3-mC (visualized using mCherry), respectively (Figure 4C and 5A). Taken together, we identified all the features in OsPSY2 to act as a transit peptide, including the critical region for translocation and the cleavage site and features related to weak and strong PG-targeting signals (Figure 6A).

Interestingly, for HR3 we observed diverse aberrant patterns of GFP fluorescence for RTp-HR3-sG, PTp-HR3-sG, PTp-sG-HR3, and PTp-sG-HR3-mC. The first case showed insufficient protein translation (not transcription), but the other three showed different posttranslational errors. This causes aggregates to form in the second case and folding errors to occur in the third and fourth cases (Figures 4D and 5B). Thus, no protein detected for RTp-HR3-sG might be caused by the elimination of unfolded, misfolded, or unimported proteins for protein homeostasis as part of general cell clean-up routines (Sun et al., 2021). This hypothesis is supported by our observation of a very thin, premature size of the GFP band (Figure 4D). This revealed the strong demand of HR3 for an exact

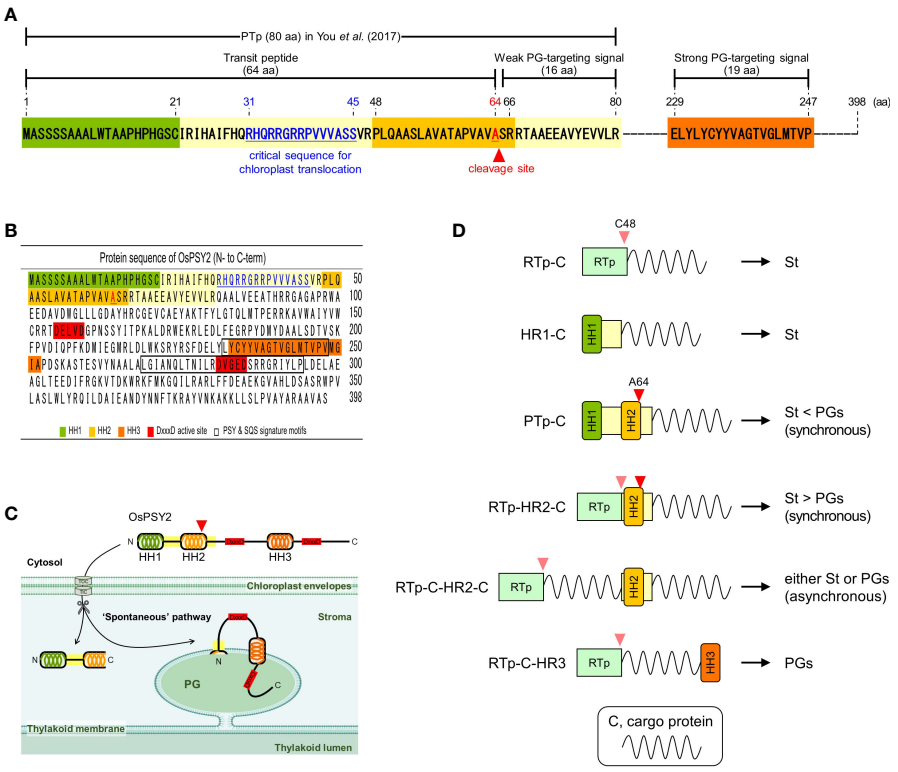


FIGURE 6
Characteristics of OsPSY2 destined for plastoglobuli localization and potential strategies for subplastidial localization. **(A)** Summary of features identified as functioning as a transit peptide and PG-targeting signal in this study. **(B)** Protein sequence information of OsPSY2, including important regions (marked in colors) including the hydrophobic helix (HH) sequences. **(C)** Proposed model indicating how OsPSY2 exists in PGs within plastids. **(D)** Proposal of six strategies for delivery of proteins of interest into the stroma and PGs with different preferences. All abbreviations here refer to Figures 1A, B.

location to preserve proper protein function. Thus, it requires sufficient distance from any other HH structure and must not be located at the C-terminal end of a cargo protein that also contains an N-terminal HH.

In general, PGs are born in highly curved margins of thylakoid membranes by blistering of the monolayer (Austin et al., 2006). It implies that PG-targeting mechanisms have a high probability of following a similar translocation path as do thylakoid membrane proteins: i.e., *via* one of the cpSRP, spontaneous, and/or cpGET pathways (Xu et al., 2021; Zhu et al., 2022). In that respect, the spontaneous pathway, which was previously reported for PsaK (Jarvis and Robinson, 2004), is the most plausible as the PG-translocation path of OsPSY2 since it has both HR2 and HR3 as well as HR1. By considering the overall protein structure, including the regions in which OsPSY2 shows activity (Figure 6B), we propose a model of how OsPSY2 is embedded in the PG-monolayer by two membrane-penetrating parts. This results in exposure of the first on the stromal side and the second inside the PG-monolayer among two DxxxD active sites (Figure 6C).

So far, the best known Tp is a 50-aa-long-RTp which has a predicted HH (1–20th aa) and is known to be cleaved at C48 by stromal processing (Eseverri et al., 2020). Here, we used our results to strategize for achieving subplastidial locations between PGs and the stroma (Figure 6D). The six strategies identified here may be useful for the delivery of cargo proteins into micro-organelle compartments, where they can work best for key plant biotechnology applications such as metabolic engineering, protein engineering, and molecular farming.

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

JY performed all experiments, including bioinformatics, vector cloning, subcellular localization, and molecular analysis, and wrote the draft. MY contributed to the experimental design and data interpretation and wrote the draft with JY. YL contributed to rice protoplast preparation and confocal microscopic analysis. S-HH envisioned and coordinated the project, supervised the manuscript, and is responsible for all contacts and correspondence. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the New Plant Breed Technology Program (PJ016548 to S-HH) funded by the Rural Development Administration of South Korea and by a grant from the National Research Foundation (Number 2021R1A2C2012227 to S-HH) funded by the Korean government (MSIP).

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.2023.1181311/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Prediction of hydrophobic helix (HH) structures in OsPSY2. **(A)** Hydrophobic degree to propose the potent hydrophobic alpha helix structures in protein sequences of OsPSY2 that was analyzed using the TMpred server. Three putative hydrophobic amino acid regions are plotted as solid lines above the peaks. **(B)** Hydrophobic residue distribution of three HH structures of OsPSY2. The number on both sides indicates the amino acid residue position of the OsPSY2. The hydrophobic residues that are generally insoluble and uncharged in aqueous solution (i. e., A, F, I, L, M, P, V, W, and Y), are marked in red letters.

SUPPLEMENTARY FIGURE 2
Verification of PG-targeting signal and cleavage occurrence at the A64 position inside the hydrophobic region (HR) 2 in rice protoplasts. **(A)** Schematic representation showing three recombinant genes of rice RuBisCO small subunit 1 transit peptide (RTp) and each of three different HR2 sequences to express green fluorescent proteins. All DNA fragments in italic were cloned into a plant expression vector (i. e., *pB2GW7*). **(B)** Protein sequence information in a full-length of HR2, a front half-length of HR2 (HR2f), and a back half-length of HR2 (HR2b) together with the simplified protein sequence of RTp. The known and predicted cleavable sites in **(C)** were bold and underlined with triangle marks. **(C)** Prediction results of cleavage probability of three RTp-fused protein regions with different HR2 parts, respectively, by the TargetP 2.0 program. **(D)** Individual and merged images of sGFP, mCherry, chlorophyll, and differential interference contrast (DIC) picture of rice protoplasts using confocal microscopy. OsPSY2:mC was used as a marker for PG localization. **(E)** Western blot analysis to confirm whether cleavage is occurring at A64 position inside HR2. Rice protoplasts were cotransfected with each experimental vector and an OsPSY2:mC vector as a transfection control. This blot was performed with three antibodies simultaneously: anti-GFP (blue circle), anti-mCherry (pink rectangle), and anti-rice actin (marked at 45 kDa, to be used as an internal reference of the total protein amount).

References

Ahrazem, O., Diretto, G., Argandona Picazo, J., Fiore, A., Rubio-Moraga, A., Rial, C., et al. (2019). The specialized roles in carotenogenesis and apocarotenogenesis of the phytoene synthase gene family in saffron. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00249

Almagro Armenteros, J. J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., et al. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* 2 (5), e201900429. doi: 10.26508/lsa.201900429

Anderson, S. A., Satyanarayan, M. B., Wessendorf, R. L., Lu, Y., and Fernandez, D. E. (2021). A homolog of GuidedEntry of tail-anchored proteins3 functions in membrane-specific protein targeting in chloroplasts of arabidopsis. *Plant Cell* 33 (8), 2812–2833. doi: 10.1093/plcell/koab145

Austin, J. R.2nd, Frost, E., Vidi, P. A., Kessler, F., and Staehelin, L. A. (2006). Plastoglobules are lipoprotein subcompartments of the chloroplast that are permanently coupled to thylakoid membranes and contain biosynthetic enzymes. *Plant Cell* 18 (7), 1693–1703. doi: 10.1105/tpc.105.039859

Bailey, J. L., and Whyborn, A. (1963). The osmiophilic globules of chloroplasts II. globules of the spinach-beet chloroplast. *Biochim. Biophys. Acta* 78 (1), 163–174. doi: 10.1016/0006-3002(63)91621-4

Bart, R., Chern, M., Park, C. J., Bartley, L., and Ronald, P. C. (2006). A novel system for gene silencing using siRNAs in rice leaf and stem-derived protoplasts. *Plant Methods* 2, 1–9. doi: 10.1186/1746-4811-2-13

Brehelin, C., and Kessler, F. (2008). The plastoglobule: a bag full of lipid biochemistry tricks. *Photochem. Photobiol.* 84 (6), 1388–1394. doi: 10.1111/j.1751-1097.2008.00459.x

Cao, W., Wang, P., Yang, L., Fang, Z., Zhang, Y., Zhuang, M., et al. (2021). Carotenoid biosynthetic genes in cabbage: genome-wide identification, evolution, and expression analysis. *Genes (Basel)* 12 (12), 2027. doi: 10.3390/genes12122027

Chen, Y., Zhou, B., Li, J., Tang, H., Tang, J., and Yang, Z. (2018). Formation and change of chloroplast-located plant metabolites in response to light conditions. *Int. J. Mol. Sci.* 19 (3), 654. doi: 10.3390/ijms19030654

Dubin, M. J., Bowler, C., and Benvenuto, G. (2008). A modified gateway cloning strategy for overexpressing tagged proteins in plants. *Plant Methods* 4, 1–11. doi: 10.1186/1746-4811-4-3

Engelman, D. M., Steitz, T. A., and Goldman, A. (1986). Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* 15, 321–353. doi: 10.1146/annurev.bb.15.060186.001541

Eseverri, A., Baysal, C., Medina, V., Capell, T., Christou, P., Rubio, L. M., et al. (2020). Transit peptides from photosynthesis-related proteins mediate import of a marker protein into different plastid types and within different species. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.560701

Gaude, N., Brehelin, C., Tischendorf, G., Kessler, F., and Dormann, P. (2007). Nitrogen deficiency in arabidopsis affects galactolipid composition and gene expression and results in accumulation of fatty acid phytyl esters. *Plant J.* 49 (4), 729–739. doi: 10.1111/j.1365-313X.2006.02992.x

Greenwood, A., Leech, R. M., and Williams, J. (1963). The osmiophilic globules of chloroplasts: i. osmiophilic globules as a normal component of chloroplasts and their isolation and composition in vicia faba l. *Biochim. Biophys. Acta* 78 (1), 148–162. doi: 10.1016/0006-3002(63)91620-2

Jarvis, P. (2008). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* 179 (2), 257–285. doi: 10.1111/j.1469-8137.2008.02452.x

Jarvis, P., and Robinson, C. (2004). Mechanisms of protein import and routing in chloroplasts. *Curr. Biol.* 14 (24), R1064–R1077. doi: 10.1016/j.cub.2004.11.049

Joyard, J., Ferro, M., Masselon, C., Seigneurin-Berny, D., Salvi, D., Garin, J., et al. (2009). Chloroplast proteomics and the compartmentation of plastidial isoprenoid biosynthetic pathways. *Mol. Plant* 2 (6), 1154–1180. doi: 10.1093/mp/ssp088

Mant, A., Woolhead, C. A., Moore, M., Henry, R., and Robinson, C. (2001). Insertion of PsaK into the thylakoid membrane in a "Horseshoe" conformation occurs in the absence of signal recognition particle, nucleoside triphosphates, or functional albino3. *J. Biol. Chem.* 276 (39), 36200–36206. doi: 10.1074/jbc.M102914200

Menke, W. (1962). Structure and chemistry of plastids. *Annu. Rev. Plant Physiol.* 13 (1), 27–44. doi: 10.1146/annurev.pp.13.060162.000331

Paila, Y. D., Richardson, L. G. L., and Schnell, D. J. (2015). New insights into the mechanism of chloroplast protein import and its integration with protein quality control, organelle biogenesis and development. *J. Mol. Biol.* 427 (5), 1038–1060. doi: 10.1016/j.jmb.2014.08.016

Palmer, T., and Stansfeld, P. J. (2020). Targeting of proteins to the twin-arginine translocation pathway. *Mol. Microbiol.* 113 (5), 861–871. doi: 10.1111/mmi.14461

Pasquier, C., Promponas, V. J., Palaios, G. A., Hamodrakas, J. S., and Hamodrakas, S. J. (1999). A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm. *Protein Eng.* 12 (5), 381–385. doi: 10.1093/protein/12.5.381

Rottet, S., Besagni, C., and Kessler, F. (2015). The role of plastoglobules in thylakoid lipid remodeling during plant development. *Biochim. Biophys. Acta* 1847 (9), 889–899. doi: 10.1016/j.bbabio.2015.02.002

Rottet, S., Devillers, J., Glauser, G., Douet, V., Besagni, C., and Kessler, F. (2016). Identification of plastoglobules as a site of carotenoid cleavage. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01855

Rubio, A., Rambla, J. L., Santaella, M., Gomez, M. D., Orzaez, D., Granell, A., et al. (2008). Cytosolic and plastoglobule-targeted carotenoid dioxygenases from crocus sativus are both involved in beta-ionone release. *J. Biol. Chem.* 283 (36), 24816–24825. doi: 10.1074/jbc.M804000200

Shumskaya, M., Bradbury, L. M., Monaco, R. R., and Wurtzel, E. T. (2012). Plastid localization of the key carotenoid enzyme phytoene synthase is altered by isozyme, allelic variation, and activity. *Plant Cell* 24 (9), 3725–3741. doi: 10.1105/tpc.112.104174

Steinmuller, D., and Tevini, M. (1985). Composition and function of plastoglobuli : i. isolation and purification from chloroplasts and chromoplasts. *Planta* 163 (2), 201–207. doi: 10.1007/BF00393507

Sun, J. L., Li, J. Y., Wang, M. J., Song, Z. T., and Liu, J. X. (2021). Protein quality control in plant organelles: current progress and future perspectives. *Mol. Plant* 14 (1), 95–114. doi: 10.1016/j.molp.2020.10.011

van Wijk, K. J., and Kessler, F. (2017). Plastoglobuli: plastid microcompartments with integrated functions in metabolism, plastid developmental transitions, and environmental adaptation. *Annu. Rev. Plant Biol.* 68, 253–289. doi: 10.1146/annurev-arplant-043015-111737

Vidi, P. A., Kanwischer, M., Baginsky, S., Austin, J. R., Csucs, G., Dormann, P., et al. (2006). Tocopherol cyclase (VTE1) localization and vitamin e accumulation in chloroplast plastoglobule lipoprotein particles. *J. Biol. Chem.* 281 (16), 11225–11234. doi: 10.1074/jbc.M511939200

Welsch, R., Wust, F., Bar, C., Al-Babili, S., and Beyer, P. (2008). A third phytoene synthase is devoted to abiotic stress-induced abscisic acid formation in rice and defines functional diversification of phytoene synthase genes. *Plant Physiol.* 147 (1), 367–380. doi: 10.1104/pp.108.117028

White, S. H., and von Heijne, G. (2005). Transmembrane helices before, during, and after insertion. *Curr. Opin. Struct. Biol.* 15 (4), 378–386. doi: 10.1016/j.sbi.2005.07.004

Woolhead, C. A., Thompson, S. J., Moore, M., Tissier, C., Mant, A., Rodger, A., et al. (2001). Distinct Albino3-dependent and -independent pathways for thylakoid membrane protein insertion. *J. Biol. Chem.* 276 (44), 40841–40846. doi: 10.1074/jbc.M106523200

Xu, X., Ouyang, M., Lu, D., Zheng, C., and Zhang, L. (2021). Protein sorting within chloroplasts. *Trends Cell Biol.* 31 (1), 9–16. doi: 10.1016/j.tcb.2020.09.011

You, M. K., Kim, J. H., Lee, Y. J., Jeong, Y. S., and Ha, S.-H. (2017). Plastoglobule-targeting competence of a putative transit peptide sequence from rice phytoene synthase 2 in plastids. *Int. J. Mol. Sci.* 18 (1), 18. doi: 10.3390/ijms18010018

You, M. K., Lim, S. H., Kim, M. J., Jeong, Y. S., Lee, M. G., and Ha, S. H. (2015). Improvement of the fluorescence intensity during a flow cytometric analysis for rice protoplasts by localization of a green fluorescent protein into chloroplasts. *Int. J. Mol. Sci.* 16 (1), 788–804. doi: 10.3390/ijms16010788

Ytterberg, A. J., Peltier, J. B., and van Wijk, K. J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. a surprising site for differential accumulation of metabolic enzymes. *Plant Physiol.* 140 (3), 984–997. doi: 10.1104/pp.105.076083

Zhu, D., Xiong, H., Wu, J., Zheng, C., Lu, D., Zhang, L., et al. (2022). Protein targeting into the thylakoid membrane through different pathways. *Front. Physiol.* 12. doi: 10.3389/fphys.2021.802057

Zita, W., Bressoud, S., Glauser, G., Kessler, F., and Shanmugabalaji, V. (2022). Chromoplast plastoglobules recruit the carotenoid biosynthetic pathway and contribute to carotenoid accumulation during tomato fruit maturation. *PloS One* 17 (12), e0277774. doi: 10.1371/journal.pone.0277774

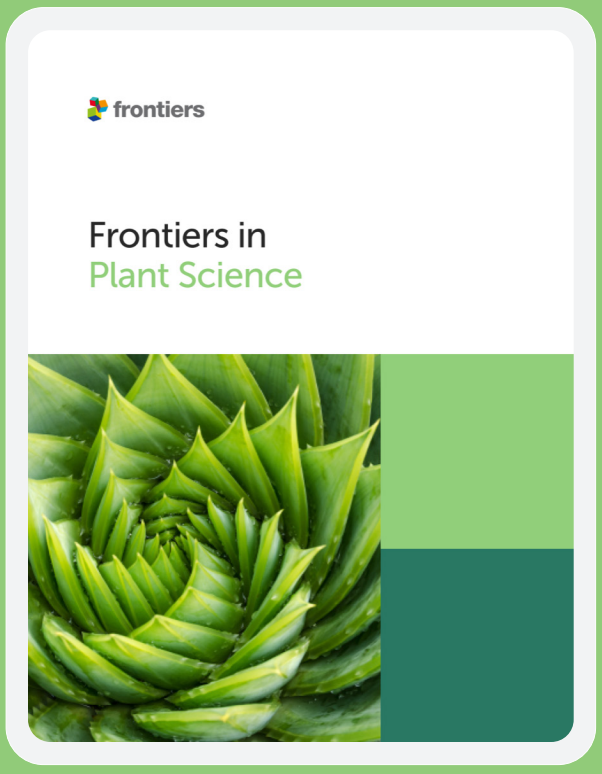
Frontiers in
Plant Science

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