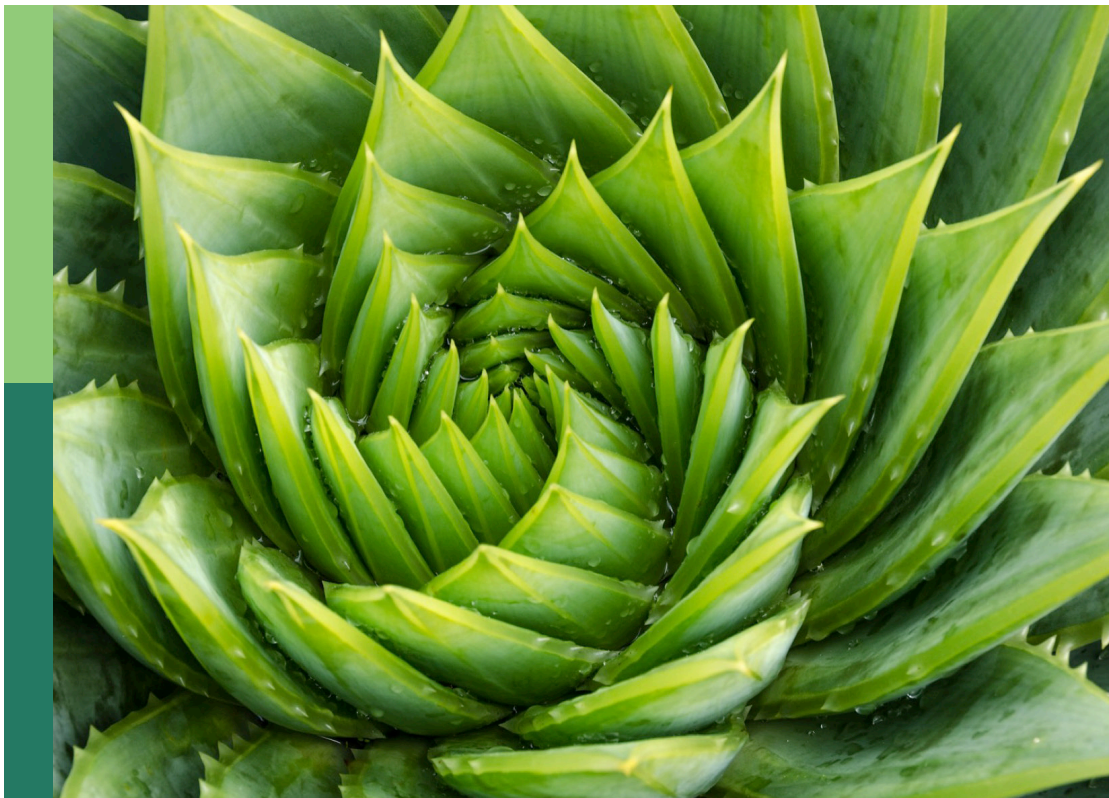
frontiers | Research Topics

Advances in discoveries of plant phytochemicals

Edited by
Rajesh Chandra Misra, Ramesha Thimmappa and Mercedes Bonfill

Published in
Frontiers in Plant Science
Frontiers in Pharmacology



ISSN 1664-8714
ISBN 978-2-8325-4683-3
DOI 10.3389/978-2-8325-4683-3

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

Advances in discoveries of plant phytochemicals

Topic editors

Rajesh Chandra Misra — John Innes Centre, United Kingdom

Ramesha Thimmappa — Amity University, India

Mercedes Bonfill — University of Barcelona, Spain

Citation

Misra, R. C., Thimmappa, R., Bonfill, M., eds. (2024). *Advances in discoveries of plant phytochemicals*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4683-3

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 30 April 2024
DOI 10.3389/fpls.2024.1414150

OPEN ACCESS

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

*CORRESPONDENCE
Rajesh Chandra Misra
rajesh.chandra-misra@jic.ac.uk

RECEIVED 08 April 2024
ACCEPTED 22 April 2024
PUBLISHED 30 April 2024

CITATION
Misra RC, Thimmappa R and Bonfill M (2024) Editorial: Advances in discoveries of plant phytochemicals.
Front. Plant Sci. 15:1414150.
doi: 10.3389/fpls.2024.1414150

Editorial: Advances in discoveries of plant phytochemicals

Rajesh Chandra Misra[1]*, Ramesha Thimmappa[2] and Mercedes Bonfill[3]

[1]Biochemistry and Metabolism Department, John Innes Centre, Norwich, United Kingdom, [2]Amity Institute of Genome Engineering, Amity University, Noida, Uttar Pradesh, India, [3]Department of Biology, Healthcare and Environment, Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain

KEYWORDS

phytochemicals, bioactivity, metabolism, biosynthetic pathways, metabolic engineering, heterologous expression

Editorial on the Research Topic
Advances in discoveries of plant phytochemicals

In the realm of health and medicine, the pursuit of novel therapeutic agents often leads scientists to explore the hidden riches of nature. Among these treasures, plant phytochemicals stand out as a diverse and promising category of compounds with profound implications for human health. Recent advances in the discovery and understanding of these phytochemicals have unveiled a wealth of opportunities that could revolutionize medicine, nutrition, and even environmental as well as agricultural sustainability. Phytochemicals are naturally occurring compounds found in plants, encompassing a wide array of chemical classes such as flavonoids, alkaloids, terpenes, and phenolic compounds. While plants produce phytochemicals primarily for their own defence against environmental stressors, these compounds also exhibit remarkable bioactivity when consumed by humans. From antioxidant and anti-inflammatory properties to potential anticancer effects, the therapeutic potential of phytochemicals has captivated researchers worldwide. The therapeutic potential of plant phytochemicals extends across a wide range of health conditions. For instance, flavonoids, abundant in fruits and vegetables, have been linked to cardiovascular health, cognitive function, and even longevity (Panche et al., 2016). Meanwhile, polyphenols found in tea, cocoa, and red wine have garnered attention for their antioxidant properties and potential role in preventing chronic diseases such as cancer and neurodegenerative disorders (Pandey and Rizvi, 2009). Despite these remarkable advancements, challenges remain in harnessing the full potential of plant phytochemicals. Issues such as bioavailability, toxicity, and standardization pose hurdles to their development as pharmaceuticals and functional foods (Atanasov et al., 2015). Moreover, biologically active phytochemicals are regulated by a variety of factors including genetics, environmental conditions, and developmental stages of the plant that involve highly complex and sophisticated biosynthetic pathways (Misra et al., 2014, 2015; Verma and Shukla, 2015; Isah, 2019). However, advancement in next generation sequencing (NGS) and bioinformatic tools have made the rich metabolism of plants more accessible. Historically, plant pathway elucidation has been a challenge. A comprehensive knowledge of the biosynthetic pathways that generate these high value molecules will assist in their exploitation for a wider variety of applications. For instance,

identification of missing enzyme sets for vinblastine biosynthesis, a potent anticancer drug from *Catharanthus roseus* and elucidation of the complete biosynthetic pathway of QS-21, a promising vaccine adjuvant derived from *Quillaja saponaria* (Caputi et al., 2018; Reed et al., 2023; Martin et al., 2024). Additionally, one of the most significant advances in phytochemical research has been the development of new analytical techniques. Advanced analytical techniques, such as mass spectrometry and nuclear magnetic resonance spectroscopy, have enabled scientists to identify and characterize phytochemicals with unprecedented precision that led to the discovery of new phytochemicals in plants that were previously thought to be absent.

The Research Topic includes twelve original research articles and three review articles, with a special focus on the new discoveries of plant phytochemicals as well as their bioactivities. Indeed, plant phytochemicals have garnered considerable attention for their potential therapeutic benefits across various health conditions. For example, Kim et al., investigates the impact of salicylic acid (SA) treatment on secondary metabolites in soybean roots and their potential anti-LDL (low-density lipoprotein) oxidation effects. The authors found that SA treatment led to significant changes in secondary metabolites in soybean roots in particular, SA stimulated the production of coumestrol, a beneficial compound, and broke down its precursors (coumestrin and malonylcoumestrin). These alterations included increases in certain compounds known for their antioxidant properties. Moreover, the authors observed that extracts from SA-treated soybean roots exhibited a much stronger ability to prevent LDL cholesterol oxidation compared to untreated roots, suggesting potential health benefits related to cardiovascular health. Further, Qi et al., reports the anti-acetylcholinesterase (anti-AChE) potential and alkaloid composition of *Rhizoma Coptidis* (RC) from different *Coptis* species. It employs a combined approach involving spectrum-effect relationship analysis and molecular docking. The study suggests *Coptis teeta* might be the best source of RC for Alzheimer's treatment. Extracts from this species showed the strongest inhibitory activity against acetylcholinesterase (AChE), an enzyme involved in Alzheimer's progression. Three alkaloids, columbine, berberine, and palmatine, were pinpointed as the main contributors to AChE inhibition. These can be used as markers for selecting the best RC source for Alzheimer's treatment. Furthermore, molecular docking simulations supported the findings, indicating strong binding between these key alkaloids and the active site of AChE, which provides insights into the mechanisms underlying their activity. In another study, Joshi et al., inspected if *Cyamopsis tetragonoloba* (guar) seed extract could enhance the antioxidant activity of existing phenolic phytochemicals. The authors found that guar seed extract, at low concentrations, significantly increased the antioxidant activity of epigallocatechin gallate (EGCG), a powerful antioxidant. This suggests guar seed extract has potential as an antioxidant booster. The extract also showed promise in protecting cells from oxidative stress in lab tests. Finally, the authors identified previously unknown metabolites in guar extract, which might explain its antioxidant-enhancing effect. The findings could contribute to understanding the synergistic effects of plant extracts and phenolic compounds in combating oxidative stress, which is crucial for potential applications in food, pharmaceuticals, or nutraceuticals. These discoveries underscore the importance of a plant-rich diet in promoting overall health and well-being. Additionally, Kyriakou et al., examines the chemical composition of watercress, focusing on polyphenolics, glucosinolates, and isothiocyanates present in its aerial parts. The authors isolated and identified these compounds in the aerial parts of watercress using analytical techniques like liquid chromatography with tandem mass spectrometry and were also able to quantify the amounts of each compound present. Interestingly, they found that the content of isothiocyanates (potentially beneficial compounds) depended on the presence of other glucosinolates, not just individual ones. Overall, this research provides a detailed analysis of the major health-promoting compounds in watercress, paving the way for using different watercress parts for potential future therapies.

Furthermore, the exploration of plant phytochemicals holds promise for addressing pressing global challenges, including human health and environmental pollution. Plant-derived compounds have shown antibacterial, antifungal, and antiviral activities, offering new avenues for combating infectious diseases in an era of dwindling antibiotic efficacy (Orhan et al., 2010; Polturak et al., 2023). Additionally, phytochemicals can play a role in sustainable agriculture by serving as alternatives to synthetic pesticides and fertilizers, thus reducing the ecological footprint of farming practices. A study by Liu et al., identified four sesquiterpenes from *A. artemisiifolia* and examined their impact on neighbouring plants. They found that these sesquiterpenes exhibited allelopathic effects, affecting the germination and growth of other plants. This research sheds light on the potential role of sesquiterpenes in the ecology and competitive interactions of *A. artemisiifolia*. In a study by Alruhaili et al., revealed the phytochemical composition and antimicrobial properties of two plants, *Amomum subulatum* and *Amomum xanthioides*, through both experimental and computational approaches. The authors analysed the chemical constituents of these plants including protein, lipids, and essential oils and evaluated their antimicrobial activity against various pathogens. Interestingly, *A. subulatum* had higher levels of carvacrol, a compound known for its antimicrobial properties, compared to *A. xanthioides*. Extracts from both plants showed antioxidant activity, with *A. subulatum* seeds having the strongest effect. Additionally, the study revealed that *A. subulatum* extracts were particularly effective against several harmful bacteria species, which could be explored for therapeutic purposes. Further, the *in silico* analysis provides insights into the mechanisms through which these phytochemicals may exert their antimicrobial effects. A systematic elucidation of natural product biosynthetic pathways, leading to a better understanding of how these valuable compounds are made and provide opportunities for metabolic pathway engineering. For example, Almeida et al., established hairy roots of *Cucurbita pepo* as a platform to modify and increase production of cucurbitacins, which are valuable plant compounds with potential medicinal applications. The authors aim to increase the yield of cucurbitacins by manipulating the metabolic pathways in *Cucurbita pepo* through genetic engineering techniques. They showed that overexpression of *CpCUCbH1* (bHLH transcription factor) can induce cucurbitacins in several Cucurbitaceae species, and also overexpression of the cytochromes P450 *CsCYP88L2* and *McCYP88L7* from *Cucumis sativus* and *Momordica charantia* (respectively), results in accumulation of new analogues of

cucurbitacin with distinct structural modifications that are previously unknown. The study provides initial evidence that a hairy roots platform can be used in modifying and increasing the production of valuable plant specialized metabolites for which the biosynthetic pathway has not been fully characterized. Furthermore, Istiandari et al., investigates the roles of Class I and Class II NADPH cytochrome P450 reductases in triterpenoid biosynthesis within *Lotus japonicus*. The authors found that these two classes of reductases play distinct roles in the biosynthesis process. Class I CPR, encoded by the *LjCPR1* gene, seems to be crucial for plant growth and development, particularly seed development. Class II CPRs, encoded by *LjCPR2* genes, are more involved in the specific production of soyasaponins, a type of triterpenoids. The findings suggest that this difference arises because Class I CPRs are generally involved in basic plant metabolism, while Class II CPRs are specialized for synthesizing particular compounds like soyasaponins. This finding contributes to our understanding of how plants produce complex molecules through the interplay of different enzymes. Finally, Zhang et al., reviewed how advancements in high-throughput omics technologies are revolutionizing the discovery of new drugs from medicinal plants. Omics technologies encompass genomics, transcriptomics, proteomics, and metabolomics, which allow for comprehensive analysis of biological systems. By applying these technologies to medicinal plants, researchers can identify and characterize bioactive compounds with potential therapeutic applications. This integrated approach offers a more efficient and systematic way to discover novel drugs from natural sources compared to traditional methods. The review highlights the benefits and challenges of employing omics technologies in medicinal plants research and underscores their significant role in advancing natural drug discovery.

In conclusion, the recent strides in the discovery and characterization of plant phytochemicals represent a watershed moment in pharmaceuticals. These natural compounds offer a treasure trove of therapeutic opportunities, from disease prevention to environmental stewardship. As scientists continue to unravel the mysteries of the plant kingdom, we can look forward to a future where the healing power of nature is fully realized for the benefit of humanity. It is imperative that we support and invest in further research endeavours to unlock the full potential of these botanical wonders.

Challenges and future prospects

The field of phytochemical research has seen significant advances in recent years, with ongoing discoveries of new plant compounds and a growing understanding of their potential health benefits. Despite the many advances in phytochemical research, there are still some challenges that need to be addressed. One challenge is the lack of standardization in the production of phytochemical extracts. This can make it difficult to compare the results of different studies and to develop reliable and effective phytochemical-based products. Another challenge is the need for more clinical trials to evaluate the safety and efficacy of phytochemicals. Although many phytochemicals have shown promise in preclinical studies, more research is needed to confirm their benefits in humans. However, the advances in the discovery and understanding of plant phytochemicals have the potential to revolutionize the way we approach nutrition and preventive healthcare. Researchers are continuing to make new discoveries about the health benefits of phytochemicals and to develop new and innovative ways to use them to prevent and treat disease. One promising area of research is the development of personalized phytochemical-based therapies. This could lead to more effective and targeted treatments for a variety of diseases. Another promising area of research is the combination therapy where phytochemicals can be used in combination with conventional drugs to improve efficacy and reduce side effects.

Author contributions

RCM: Conceptualization, Resources, Visualization, Writing – original draft, Writing – review & editing. RT: Writing – original draft, Writing – review & editing, Funding acquisition. MB: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. RT acknowledges funding support from the Department of Biotechnology (DBT)-Ramalingaswami Re-entry fellowship from India.

Acknowledgments

The editors thank all reviewers who evaluated manuscripts and contributors for their Research Topic.

Conflict of interest

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

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

Publisher's note

References

Atanasov, A. G., Waltenberger, B., Pferschy-Wenzig, E. M., Linder, T., Wawrosch, C., Uhrin, P., et al. (2015). Discovery and resupply of pharmacologically active plant-derived natural products: a review. *Biotechnol. Adv.* 33, 1582–1614. doi: 10.1016/j.biotechadv.2015.08.001

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. *Science* 360, 1235–1239. doi: 10.1126/science.aat4100

Isah, T. (2019). Stress and defence responses in plant secondary metabolites production. *Biol. Res.* 52, 39. doi: 10.1186/s40659-019-0246-3

Martin, L. B. B., Kikuchi, S., Rejzek, M., Owen, C., Reed, J., Orme, A., et al. (2024). Complete biosynthesis of the potent vaccine adjuvant QS-21. *Nat. Chem. Biol.* 20, 493–502. doi: 10.1038/s41589-023-01538-5

Misra, R. C., Garg, A., Roy, S., Chanotiya, C. S., Vasudev, P. G., and Ghosh, S. (2015). Involvement of an ent-copalyl diphosphate synthase in tissue-specific accumulation of specialized diterpenes in *Andrographis paniculata*. *Plant Sci.* 240, 50–64. doi: 10.1016/j.plantsci.2015.08.016

Misra, R. C., Maiti, P., Chanotiya, C. S., Shanker, K., and Ghosh, S. (2014). Methyl jasmonate-elicited transcriptional responses and pentacyclic triterpene biosynthesis in sweet basil. *Plant Physiol.* 164, 1028–1044. doi: 10.1104/pp.113.232884

Orhan, D. D., Özçelik, B., Özgen, S., and Ergun, F. (2010). Antibacterial, antifungal, and antiviral activities of some flavonoids. *Microbiol. Res.* 165, 496–504. doi: 10.1016/j.micres.2009.09.002

Panche, A. N., Diwan, A. D., and Chandra, S. R. (2016). Flavonoids: an overview. *J. Nutr. Sci.* 5, e47. doi: 10.1017/jns.2016.41

Pandey, K. B., and Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. *Oxid. Med. Cell Longev.* 2, 270–278. doi: 10.4161/oxim.2.5.9498

Polturak, G., Misra, R. C., El-Demerdash, A., Owen, C., Steed, A., McDonald, H., et al. (2023). Discovery of isoflavone phytoalexins in wheat reveals an alternative route to isoflavonoid biosynthesis. *Nat. Commun.* 14, 6977. doi: 10.1038/s41467-023-42464-3

Reed, J., Orme, A., Demerdash, A., Owen, C., Martin, L., Misra, R. C., et al. (2023). Elucidation of the pathway for biosynthesis of saponin adjuvants from the soapbark tree. *Science* 379, 1252–1264. doi: 10.1126/science.adf3727

Verma, N., and Shukla, S. (2015). Impact of various factors responsible for fluctuation in plant secondary metabolites. *J. Appl. Res. Med. Aromat. Plants* 2, 105–113. doi: 10.1016/j.jarmap.2015.09.002

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fpls.2022.996498

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre,
United Kingdom

REVIEWED BY
Youzhi Li,
Hunan Agricultural University,
China
Margot Schulz,
University of Bonn,
Germany

*CORRESPONDENCE
Yulong Feng
fyl@syau.edu.cn
Tong An
m18602463916@163.com

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

RECEIVED 17 July 2022
ACCEPTED 22 August 2022
PUBLISHED 02 September 2022

CITATION
Liu Z, Zhang N, Ma X, Zhang T, Li X, Tian G, Feng Y and An T (2022) Sesquiterpenes from *Ambrosia artemisiifolia* and their allelopathy.
Front. Plant Sci. 13:996498.
doi: 10.3389/fpls.2022.996498

Sesquiterpenes from *Ambrosia artemisiifolia* and their allelopathy

Zhixiang Liu[1,2], Nan Zhang[2], Xiaoqing Ma[2], Tong Zhang[2], Xuan Li[2], Ge Tian[2], Yulong Feng[1,2]* and Tong An[1,2]*

[1]College of Plant Protection, Shenyang Agricultural University, Shenyang, China, [2]College of Biological Science and Technology, Shenyang Agricultural University, Shenyang, China

Ambrosia artemisiifolia, an invasive plant, has seriously harmed the agricultural production, native ecosystems and human health. Allelopathy is an important reason for the successful invasion of this alien plant. However, the chemical basis, action effects, action mechanism and release pathway of its allelopathy remain unclear. To address these problems, four sesquiterpenes (**1**–**4**), consisting of three new sesquiterpenes (**1**–**2**, **4**), were isolated from the whole plant of *A. artemisiifolia* using a variety of column chromatography techniques, and identified using HR-ESIMS, 1D-NMR, 2D-NMR, and ECD. All the compounds exhibited different levels of inhibitory effects on three native plants (*Setaria viridis*, *Digitaria sanguinalis*, *Chenopodium album*) and one model plant (*Arabidopsis thaliana*), especially compound **1**. In addition, the preliminary action mechanism of active compound 1 was revealed by FDA/PI staining assay. Furthermore, the allelopathic substances **1**–**3** were released into environment through the root secretion pathway by UPLC-MS/MS analyses.

KEYWORDS

allelopathy, *Ambrosia artemisiifolia*, invasive plant, isolate, sesquiterpenes

Introduction

In recent decades, with the acceleration of globalization and world trade, the spread of invasive plants has been increasing, causing a serious threat to agricultural production and the ecological environment (Bonnamour et al., 2021). It is urgent to reveal the invasive mechanism of alien plants, which is necessary for risk assessment of invasive plants and to develop effective prevention methods. In recent years, the novel weapons hypothesis (NWH) has become a hot topic of research for investigating the invasion mechanism of alien plants (Callaway and Ridenour, 2004). The hypothesis suggests that the successful invasion of alien plants is largely due to their release of allelopathic substances that are relatively novel to native plants (Bais, 2003). The allelopathic substances are released into the environment through appropriate pathways, such as root secretion and rainwater leaching, to negatively affect the growth of native plants, enabling alien plants to gain advantages and promote their expansion into dominant single populations (Thorpe et al., 2009). Therefore, fully understanding the allelopathic effects of invasive plants on native

plants may help reveal the invasion mechanism of alien plants while also clarifying the chemical relationship between invasive plants and native plants, and this may enable new approaches for the utilization of invasive plants.

Ambrosia artemisiifolia L. (Asteraceae), which originated from North America, is an invasive weed found all over the world. The invader can inhibit the growth of native plants, destroying the original ecosystem and reducing agricultural production (Sun and Roderick, 2019). Furthermore, its pollen can cause a series of allergic reactions and directly affect human health (Zhao et al., 2017). Field observations have shown that *A. artemisiifolia* often forms a dominant single population with few other plants around it. The reason is not only that *A. artemisiifolia* has strong growth characteristics, but also its successful allelopathic effects on native plants. Previous studies reported that various solvent extracts of *A. artemisiifolia* had significant inhibitory effects on the growth and seed germination of *Lactuca sativa*, *Lycopersicon esculentum*, *Zea mays*, and other plants (Lehoczky et al., 2011; Vidotto et al., 2013; Bonea et al., 2018). These studies also confirmed that allelopathy is a major reason for the successful invasion of *A. artemisiifolia.* However, most studies only consider the effects of crude extract of *A. artemisiifolia* on native plants, and the chemical basis, action effects, action mechanism and release pathway of its allelopathy remain unclear. Like most higher plants, *A. artemisiifolia* contains a large number of secondary metabolites. Among them, sesquiterpenoids are the main chemical components, including the eudesmane-type, germacrane-type, bisabolane-type and guaiane-type, which have complex and variable structural frameworks (Taglialatela-Scafati et al., 2012; Ding et al., 2015; An et al., 2019). Studies have shown that sesquiterpenoids play an important role in plant allelopathy, which can significantly inhibit plant growth. Similarly, sesquiterpenoids may also be the potential allelopathic substances for *A. artemisiifolia* (Bennett and Wallsgrove, 1994; Chadwick et al., 2013).

In this study, to clarify the chemical basis of allelopathy, we isolated and identified the secondary metabolites from the whole plant of *A. artemisiifolia*. And the allelopathic effects of isolated compounds on three native plants (*Setaria viridis*, *Digitaria sanguinalis*, *Chenopodium album*) and one model plant (*Arabidopsis thaliana*) were also examined. In addition, the preliminary action mechanism of active compound **1** was revealed by FDA/PI staining assay. Furthermore, the release pathway of allelopathic substances was analyzed by UPLC-MS/MS analyses.

Materials and methods

General

Column chromatography was carried out using silica gel (Qingdao Marine, China), MCI (Mitsubishi, Japan), and Sephadex LH-20 (GE Healthcare, Sweden). RP-HPLC isolation was performed using a 1,260 system (Agilent, United States) coupled with a 250 mm × 10 mm, 5 μm XDB-C_{18} column (YMC, Japan). UV spectra were determined using a 241 spectrophotometer (Perkin Elmer, United States). HR-ESIMS spectra were measured using a 6,545 Q-TOF spectrometer (Agilent, United States). NMR spectra were recorded using an AV-600 instrument (Bruker, Germany). GC analysis was performed using a 7890A system (Agilent, United States). ECD spectra were obtained using a MOS-450 detector (Bio-Logic, France). Cell viability analyses were performed using an A1 laser confocal microscope (Nikon, Japan). UPLC-MS/MS analyses were performed using a 6,545 LC/Q-TOF system (Agilent, United States) coupled with a 50 mm × 2.1 mm, 1.9 μm EC-C_{18} column (Agilent, United States).

Plant material

The whole plant of *A. artemisiifolia* was collected from Shenyang, Liaoning Province, China (123° 48′ E, 42° 05′ N) in August 2020 and identified by Professor Bo Qu (Shenyang Agricultural University). A voucher specimen (20200811) was kept in the herbarium of Shenyang Agricultural University.

Extraction and isolation

The air-dried and powdered *A. artemisiifolia* (50 kg) were extracted with 80% ethanol at room temperature (200 L × 3, 7 days each time). The concentrated ethanol extract (1,200 g) was partitioned continuously with petroleum ether (5 L × 3) and ethyl acetate (5 L × 3). The ethyl acetate fraction (230 g) was then subjected to silica gel column chromatography eluted with a gradient of dichloromethane:methanol (99:1, 98:2, 95:5, 90:10, 80:20, 70:30, *v*/*v*) to yield five subfractions (Fr. A–E). Fr. D (18 g) was subjected to MCI column chromatography eluted with a gradient of methanol:H_2O (10:90, 30:70, 50:50, 70:30, *v*/*v*) to yield five subfractions (Fr. D-1–5). Fr. D-2 (2.1 g) was subjected to Sephadex LH-20 column chromatography eluted with isocratic of acetone to yield eight subfractions (Fr. D-2-1–8). Fr. D-2-3 (130 mg) was isolated using RP-HPLC (210 nm, 5.0 ml/min) eluted with isocratic of methanol:H_2O (20:80, *v*/*v*) to yield compounds **1** (10.2 mg, t_R = 32.4 min), **2** (3.5 mg, t_R = 17.5 min) and **3** (4.6 mg, t_R = 57.9 min), respectively. Similarly, Fr. D-2-5 (64 mg) was isolated using RP-HPLC (210 nm, 5.0 ml/min) eluted with isocratic of methanol:H_2O (15:85, *v*/*v*) to yield compound **4** (6.1 mg, t_R = 78.1 min).

Compound **1**: light yellow powder; UV (methanol) λ_{max} (logε): 219 (0.54) nm; ECD (methanol) λ_{max} ($\Delta\varepsilon$) 209 (−60.57), 228 (−11.81), 246 (+8.90) nm; HR-ESIMS at *m/z* 303.1206 $[M+Na]^+$ (calcd for $C_{15}H_{20}O_5Na$, 303.1208); ^1H and ^{13}C NMR data, see Table 1.

Compound **2**: light yellow oil; UV (methanol) λ_{max} (logε): 214.05 (0.20) nm; ECD (methanol) λ_{max} ($\Delta\varepsilon$) 226 (−54.63) nm; HR-ESIMS at *m/z* 321.1311 $[M+Na]^+$ (calcd for $C_{15}H_{22}O_6Na$, 321.1314); ^1H and ^{13}C NMR data, see Table 1.

TABLE 1 ^{1}H (600MHz) and ^{13}C NMR (150MHz) spectroscopic data of compounds 1–4 in methanol-d_4.

Position	1		2		3		4	
	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
1	76.2	3.42 (1H, dd, 11.8, 4.4)	76.3	3.36 (1H, dd, 12.4, 4.0)	75.6	3.42 (1H, dd, 12.2, 3.9)	131.5	
2	41.1	2.12 (1H, m) 1.55(1H, q, 11.8)	36.5	1.89 (1H, m) 1.68 (1H, q, 12.4)	34.0	1.94 (1H, m) 1.73 (1H, q, 12.2)	30.4	1.40 (1H, m) 1.28 (1H, m)
3	70.5	3.97 (1H, dd, 11.8. 5.1)	76.4	3.48 (1H, dd, 12.4, 4.5)	78.2	4.72 (1H, dd, 12.2, 4.6)	207.2	
4	148.0		76.3		74.9		138.2	
5	53.9	1.72 (1H, d, 11.0)	57.7	1.33 (1H, d, 11.5)	57.6	1.45 (1H, d, 12.0)	165.2	
6	79.8	5.20 (1H, d, 11.0)	81.1	5.30 (1H, d, 11.5)	80.6	5.31 (1H, d, 12.0)	78.8	5.41 (1H, d, 11.1)
7	169.3		169.8		169.5		56.4	2.14 (1H, m)
8	23.1	3.05 (1H, m) 2.50 (1H, td, 14.1, 5.7)	23.3	3.30 (1H, m) 2.49 (1H, td, 14.2, 5.6)	23.2	3.04 (1H, m) 2.50 (1H, td, 14.0, 5.5)	74.1	4.15 (1H, m)
9	37.8	2.23 (1H, m) 1.28 (1H, m)	41.2	2.17 (1H, m) 1.24 (1H, m)	41.0	2.18 (1H, m) 1.27 (1H, m)	42.6	2.88 (1H, dd, 15.2, 3.1) 2.35 (1H, dd, 15.2, 3.1)
10	42.6		42.3		42.2		133.3	
11	123.7		123.6		123.9		42.5	2.75 (1H, m)
12	175.6		174.6		174.5		180.0	
13	54.0	4.28 (2H, s)	54.0	4.28 (2H, s)	54.0	4.29 (2H, s)	14.9	1.32 (3H, d, 7.0)
14	10.9	0.90 (3H, s)	13.5	1.07 (3H, s)	13.5	1.10 (3H, s)	9.6	1.96 (3H, s)
15	106.7	5.37 (1H, s) 5.10 (1H, s)	17.4	1.35 (3H, s)	18.1	1.44 (3H, s)	25.5	1.94 (3H, s)
1′					172.2			
2′					21.0	2.06 (3H, s)		

Compound **3**: light yellow oil; UV (methanol) λ_{max} (logε): 213 (0.32) nm; ECD (methanol) λ_{max} ($\Delta\varepsilon$) 226 (−5.92) nm; HR-ESIMS at *m/z* 363.1417 $[M+Na]^+$ (calcd for $C_{17}H_{24}O_7Na$, 363.1420); ^{1}H and ^{13}C NMR data, see Table 1.

Compound **4**: light yellow oil; UV (methanol) λ_{max} (logε): 306 (0.11) nm; ECD (methanol) λ_{max} ($\Delta\varepsilon$) 231 (+32.86), 310 (+8.16) nm; HR-ESIMS at *m/z* 263.1289 $[M+H]^+$ (calcd for $C_{15}H_{19}O_4$, 263.1283); ^{1}H and ^{13}C NMR data, see Table 1.

ECD calculations

The ECD calculations of compounds **1**–**4** were conducted using Gaussian 09 (Liu et al., 2021). Firstly, the conformational analyses were initially performed using the MMFF94 force field. Then, the obtained conformations were further optimized at the B3LYP/6-31G (d) level. Subsequently, the optimized conformations were calculated using a TDDFT method at the B3LYP/6–311 + G (2d, p) level (methanol). Finally, based on the Boltzmann weighting of each conformer, the calculated ECD curves were generated.

Allelopathic assay

The allelopathic assay was performed as a previously described method with some modifications (Li et al., 2019). Firstly, the tested compounds were dissolved with DMSO and added to different volumes of 1/2 MS medium, respectively, to obtain the medium containing 100, 50, and 25 μM compounds, and the blank control contained the same volume of DMSO. Then, the 2 ml of medium was removed into the each well of 6-well plates, respectively. Subsequently, the seeds of *S. viridis*, *D. sanguinali*, *C. album*, *A. thaliana* were sterilized with 0.1% $HgCl_2$ and washed with sterilized water at least three times, respectively. After the medium had cooled naturally to a solid state, 8 to 10 sterilized seeds were evenly placed in a row on the medium and cultured vertically in an illumination incubator (three replicates per compound). Finally, when the roots of the blank control grew to the bottom of 6-well plates, the lengths of roots were measured. Inhibitory rate (%) was calculated as $(L_C - L_T)/Lc \times 100\%$, where L_C and L_T were the average lengths of blank control and compound-treated roots, respectively. Logran, a common herbicide, was used as a positive control.

Statistical analyses

The data of allelopathic assay were expressed as means ± SD of three replicates. One-way ANOVA was used to compare data between groups when the data followed a normal distribution. Differences were considered to be statistically significant if $p < 0.05$.

Cell viability analyses

The cell viability analyses were performed as a previously described method with some modifications (Pan et al., 2001). After the allelopathic assay, the root tips of *A. thaliana* (0.5 cm) were stained with a mixture of 12.5 μg/ml FDA (fluorescein diacetate) and 5 μg/ml PI (propidium iodide) for 10 min. Then, the roots were washed with distilled water and placed on slides. Finally, the roots were observed using a laser confocal microscope (excitation at 480 nm and emission at 520 nm). Red and green fluorescence represented dead and living cells, respectively.

Collection of root secretion and rainwater leaching

The collection of root secretion and rainwater leaching was performed as a previously described method with some modifications (Wang et al., 2015). The rhizosphere soil of *A. artemisiifolia* (200 g) was randomly collected at 5–10 cm depths and carefully picked out the roots. The soil samples were crushed to pass through a sieve (30 mesh). The sieved soil was extracted ultrasonically with methanol (1 L) at room temperature for 30 min. The extraction were then filtered and concentrated *in vacuo* to obtain the root secretion samples. The aerial part of *A. artemisiifolia* (10 living bodies) was washed with distilled water (5 L) for 10 min. The rinses were then collected and concentrated *in vacuo* to obtain the rainwater leaching samples.

UPLC-MS/MS analyses

The identification and quantification of compounds **1**–**4** in the root secretion and rainwater leaching samples of *A. artemisiifolia* were performed by UPLC-MS/MS. The solution of root secretion and rainwater leaching were filtered through a millipore filter (0.22 μm) before UPLC-MS/MS analyses (flow rate of 0.4 ml/min; injection volume of 1 μl; column temperature of 30°C), respectively. The mobile phase consisted of water containing 0.1% formic acid (A) and methanol (B) was programmed as follows: 0–5 min, B 10%; 5–20 min, B from 10% to 50%. Positive ionization was performed using the following settings: gas temperature, 280°C; drying gas flow, 8 L/ min; nebulizer, 35 psi; sheath gas temperature, 320°C; sheath gas flow, 12 L/ min; fragmentor, 145 V; skimmer, 65 V; oct 1 RF Vpp, 750 V. The identification of compounds in the root secretion and rainwater leaching was determined by comparing the retention times, MS, MS/MS spectra with those of isolated compounds **1**–**4**. The quantification of compounds in the root secretion and rainwater leaching was performed using the same method described above, with the isolated compounds as external standards. The standard curves were plotted using six concentrations (100, 10, 1, 0.1, 0.01, 0.001 μg/ml) and the corresponding ion peak area. The regression equations of compounds **1**–**3** were $y = 22{,}758x - 8415.2$ ($R^2 = 0.9996$), $y = 17{,}790x + 7201.4$ ($R^2 = 0.9996$) and $y = 20{,}732x - 5553.5$ ($R^2 = 0.9997$), respectively.

Results and discussion

Compound **1** was obtained as a light yellow powder with $C_{15}H_{20}O_5$ (six degrees of unsaturation) according to HR-ESIMS data (*m/z* 303.1206 $[M + Na]^+$, calcd for $C_{15}H_{20}O_5Na$, 303.1208). Its 1D-NMR spectra (Table 1) showed one carbonyl group [δ_C 175.6 (C-12)], two double bond groups [δ_H 5.37 (1H, s, H-15), 5.10 (1H, s, H-15); δ_C 169.3 (C-7), 148.0 (C-4), 123.7 (C-11), 106.7 (C-15)], one primary alcohol group [δ_H 4.28 (2H, s, H-13); δ_C 54.0 (C-13)], two secondary alcohol groups [δ_H 3.97 (1H, dd, $J = 11.8$, 5.1 Hz, H-3), 3.42 (1H, dd, $J = 11.8$, 4.4 Hz, H-1); δ_C 76.2 (C-1), 70.5 (C-3)], and one methyl group [δ_H 0.90 (3H, s, H-14); δ_C 10.9 (C-14)], respectively.

In the HMBC spectrum (Figure 1), the key correlations of H-3/C-2, C-4; H-6/C-5; H-8/C-7, C-9; H-14/C-1, C-5, C-9, C-10; H-15/C-3, C-5 established a eudesmane-type sesquiterpene moiety (An et al., 2019). The key correlations of H-6/C-11, C-12; H-13/C-7, C-11, C-12 established a isosiphonodin moiety (Caloprisco et al., 2002). In addition, the key correlations of H-6/C-5; H-8/C-7 revealed that the isosiphonodin moiety was connected to the eudesmane-type sesquiterpene moiety through C-6 and C-7. In the NOESY spectrum (Figure 2), the key correlations of H-1/H-3, H-5; H-14/H-6 indicated that H-1, H-3, H-5 and H-6, H-14 were oriented on the opposite side. Furthermore, the calculated ECD curve of (1*R*, 3*S*, 5*S*, 6*R*, 10*R*)-**1b** matched well with the experimental result of **1** (Figure 3). Thus, the planar and stereoscopic structure of compound **1** was constructed and named Eudesmanol A (Figure 4).

Compound **2** was obtained as a light yellow oil with $C_{15}H_{22}O_6$ (five degrees of unsaturation) according to HR-ESIMS data (*m/z* 321.1311 $[M + Na]^+$, calcd for $C_{15}H_{22}O_6Na$, 321.1314). Comparison of its 1D NMR spectra with Eudesmanol A showed that they were very similar except for one additional methyl group [δ_H 1.35 (3H, s, H-15); δ_C 17.4 (C-15)], one additional tertiary alcohol group [δ_C 76.3 (C-4)], and one missing terminal double bond group. In the HMBC spectrum (Figure 1), the key correlations of H-15/C-4, C-5 revealed that the methyl and hydroxyl groups were connected to Eudesmanol A by C-4. In the NOESY spectrum (Figure 2), the key correlations of H-1/H-3, H-5; H-6/H-14, H-15 indicated that H-1, H-3, H-5 and H-6, H-14, H-15 were oriented on the opposite side. Furthermore, the calculated ECD curve of (1*R*, 3*S*, 4*S*, 5*S*, 6*R*,

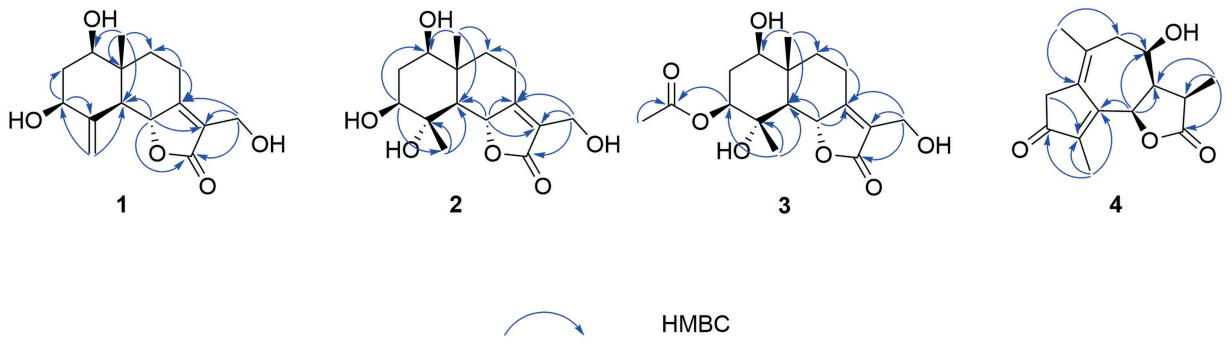


FIGURE 1
Key HMBC correlations of compounds **1–4**.

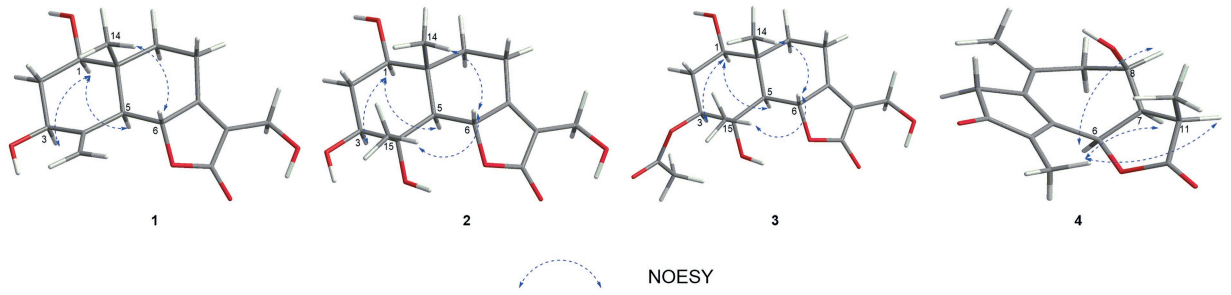


FIGURE 2
Key NOESY correlations of compounds **1–4**.

10*R*)-**2a** matched well with the experimental result of **2** (Figure 3). Thus, the planar and stereoscopic structure of compound **2** was constructed and named Eudesmanol B (Figure 4).

Compound **3** was obtained as a light yellow oil with $C_{17}H_{24}O_7$ (six degrees of unsaturation) according to HR-ESIMS data (*m/z* 363.1417 $[M+Na]^+$, calcd for $C_{17}H_{24}O_7Na$, 363.1420). Comparison of its 1D NMR spectra with a known compound 3*β*-Acetoxy-l*β*, 4*α*, 13-trihydroxyeudesm-7 (11)-*en*-6*α*, l2-olide showed that they were very similar except for the different deuterated solvents. In the HMBC spectrum (Figure 1), the key correlations revealed that compound **3** has the same planar structure as 3*β*-Acetoxy-l*β*, 4*α*, 13-trihydroxyeudesm-7 (11)-*en*-6*α*, l2-olide (Abdel-Mogi et al., 1989). In the NOESY spectrum (Figure 2), the key correlations of H-1/H-3, H-5; H-6/H-14, H-15 indicated that H-1, H-3, H-5 and H-6, H-14, H-15 were oriented on the opposite side. Furthermore, the calculated ECD curve of (1*R*, 3*S*, 4*S*, 5*S*, 6*R*, 10*R*)-**3a** matched well with the experimental result of **3** (Figure 3), which revealed that compound **3** has the same stereoscopic structure as 3*β*-Acetoxy-l*β*, 4*α*, 13-trihydroxyeudesm-7 (11)-*en*-6*α*, l2-olide. Thus, the planar and stereoscopic structure of compound **3** was constructed (Figure 4).

Compound **4** was obtained as a light yellow oil with $C_{15}H_{18}O_4$ (seven degrees of unsaturation) according to HR-ESIMS data (*m/z* 263.1289 $[M+H]^+$, calcd for $C_{15}H_{19}O_4$, 263.1283). Its 1D-NMR spectra (Table 1) showed two carbonyl groups [δ_C 207.2 (C-3), 180.0 (C-12)], two double bond groups [δ_C 165.2 (C-5), 138.2 (C-4), 133.3 (C-10), 131.5 (C-1)], one secondary alcohol group [δ_H 4.15 (1H, m, H-8); δ_C 74.1 (C-8)], three methyl groups [δ_H 1.96 (3H, s, H-14), 1.94 (3H, s), 1.32 (3H, d, J = 7.0 Hz); δ_C 25.5 (C-15), 14.9 (C-13), 9.6 (C-14)], respectively.

In the HMBC spectrum (Figure 1), the key correlations of H-2/C-4; H-6/C-5, C-7, C-8; H-9/C-8; H-14/C-3, C-4, C-5; H-15/C-1, C-9 established a guaiane-type sesquiterpene moiety (Ding et al., 2015). The key correlations of H-6/C-7; H-13/C-7, C-11, C-12 established a *α*-methylbutyzolactone moiety (Simonaitis and Pitts, 1969). In addition, the key correlations of H-6/C-5, C-7, C-8 revealed that the *α*-methylbutyzolactone moiety was connected to the guaiane-type sesquiterpene moiety by C-6 and C-7. In the NOESY spectrum (Figure 2), the key correlations of H-7/H-6; H-8/H-6; H-11/H-6 indicated that H-6, H-7, H-8, H-11 and H-13 were oriented on the opposite side. Furthermore, the calculated ECD curve of (6*R*, 7*R*, 8*R*, 11*R*)-**4a** matched well with the experimental result of **4** (Figure 3). Thus, the planar and stereoscopic structure of compound **4** was constructed and named Guaianin (Figure 4).

Moreover, we assessed the allelopathic effects of compounds **1**–**4** from *A. artemisiifolia* on the root elongation of *S. viridis*, *D. sanguinali*, *C. album*, *A. thaliana*. As shown in

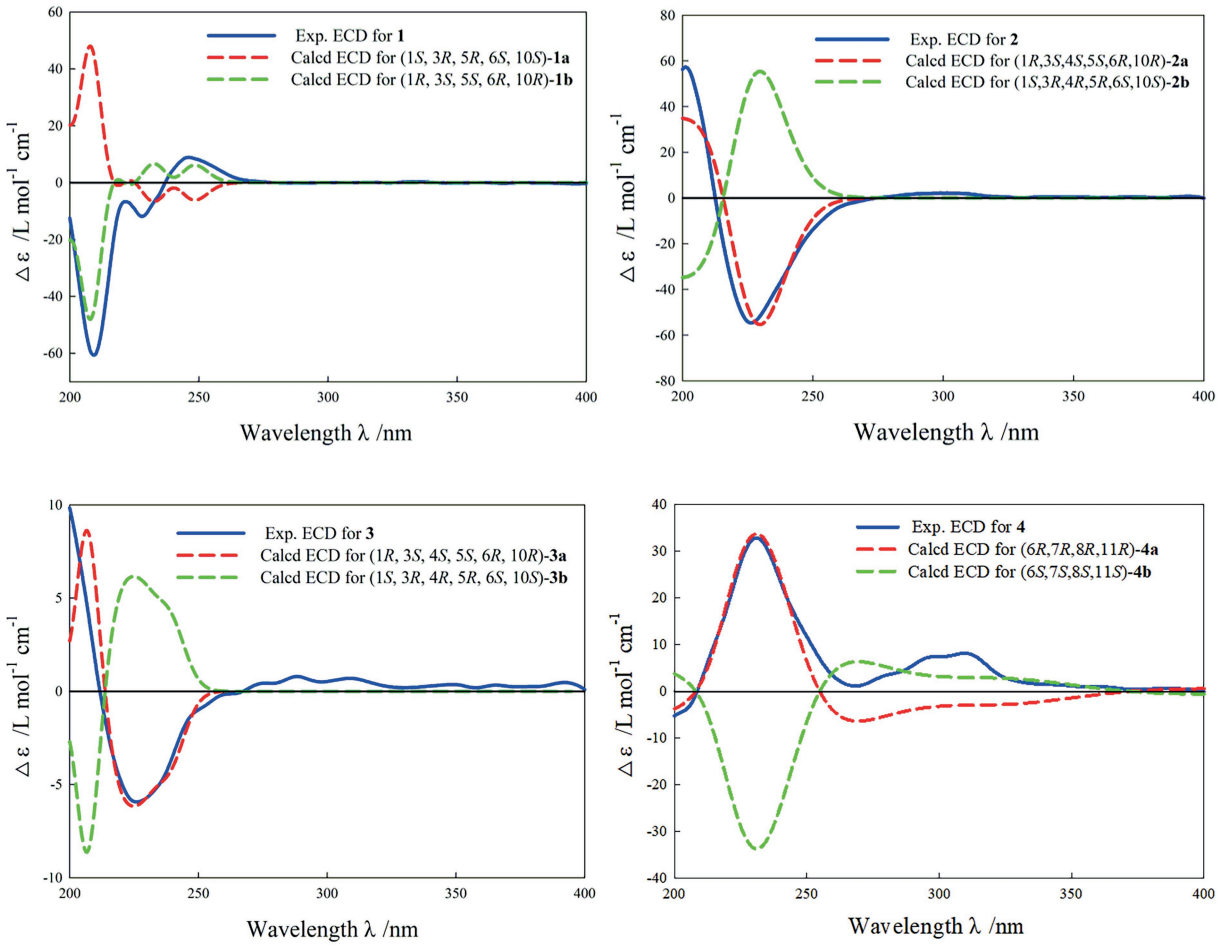


FIGURE 3
Experimental and calculated ECD curve of compounds **1–4**.

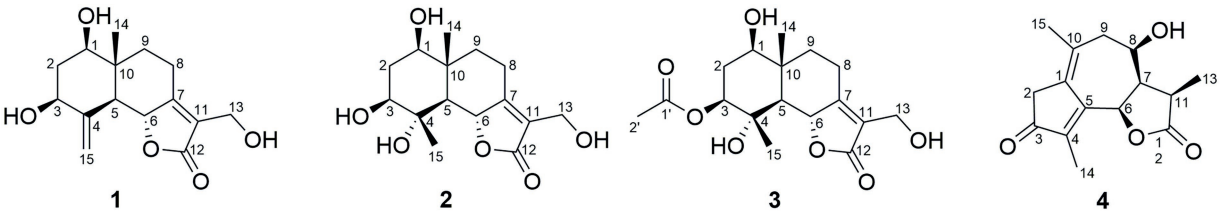


FIGURE 4
Structures of compounds **1–4**.

Figure 5, all of the compounds exhibited different levels of allelopathic effects. Compounds **1–3** exhibited moderate allelopathic effects with inhibition rates ranging from 15.60% to 73.42% at 100 μM. In addition, compound **4** exhibited slight allelopathic effects, and promoted the root elongation of three kinds of *S. viridis*, *D. sanguinali*, *A. thaliana* at low concentrations. Notably, compound **1** showed potent allelopathic effects (73.42% ± 8.54% on *S. viridis*, 51.23% ± 4.12% on *D. sanguinali*, 69.88% ± 8.09% on *C. album*, 59.34% ± 5.80% on *A. thaliana*, respectively) with more than 50% inhibitory rate at 100 μM, which approached the results observed for Logran. Comparisons of the structure–activity relationships of compounds **1–3** showed that the eudesmane-type sesquiterpenes with a terminal double bond group at C-4 may have greater allelopathic effects than the compounds with hydroxyl and methyl groups at C-4.

When plants are subjected to external stress, the cell viability changes. Thus, cell viability can be used as an important indicator for evaluating the allelopathic effects of compounds on plants (Yan et al., 2016). The preliminary action mechanism of active compound **1** was revealed by FDA/PI staining assay. As shown in Figure 6, the root tips of *A. thaliana* began to show red fluorescence at 25 μM of compound **1**, indicating that compound **1** could cause the partial cell death of *A. thaliana* at a low concentration. Additionally, when the concentration of compound **1** was 100 μM, the red fluorescence was dominant and green fluorescence was reduced in the root tips of *A. thaliana*. The results showed that the cell viability of root tips of *A. thaliana* decreased with increasing concentrations of compound **1**. Therefore, we speculated that compound **1** could play an allelopathic role by decreasing the cell viability of plants. However, the reason why the plant cell viability was decreased by compound **1** needs to be explored further.

It is well-known that allelopathic substances can play their allelopathic roles only when they are released into environment through appropriate pathways (Thorpe et al., 2009). To determine the allelopathic release pathway of the sesquiterpenes **1**–**4**, UPLC-MS/MS analyses were carried out to detect root secretion and rainwater leaching of *A. artemisiifolia*. According to their TIC chromatogram (Figure 7), compounds **1**–**3** had retention times of 4.42, 1.67, and 8.29 min, respectively, and they were detected in the root secretion of *A. artemisiifolia*., but no compounds were detected in rainwater leaching. The concentrations of compounds **1**–**3** in the rhizosphere soil were also determined as 0.55 ± 0.06, 0.44 ± 0.03, 0.48 ± 0.04 μg/g, respectively. The results indicated that

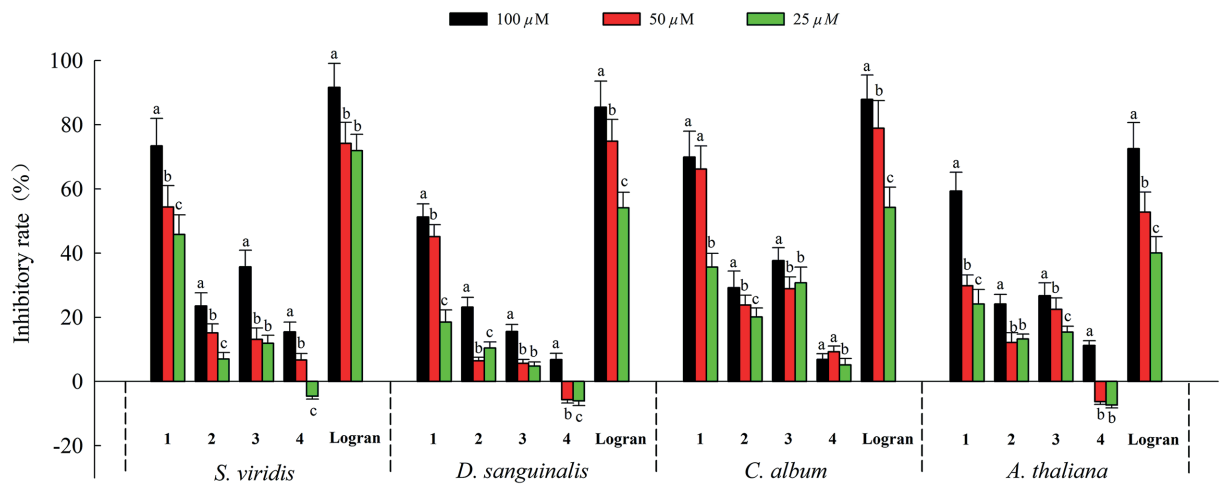


FIGURE 5
Allelopathic effects of compounds **1**–**4** on root growth of weeds. Different letters indicate significant differences between concentration treatments (p<0.05).

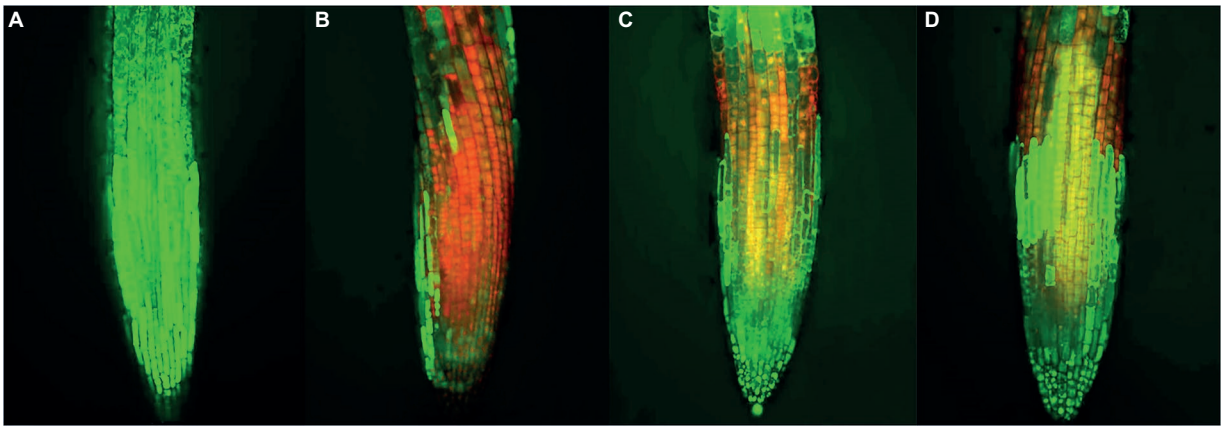


FIGURE 6
Effects of compound **1** on cell viability of the root tips of *Arabidopsis thaliana*. at concentrations of **(A)** 0, **(B)** 100, **(C)** 50, and **(D)** 25μM, respectively. Red and green fluorescence represented dead and living cells, respectively.

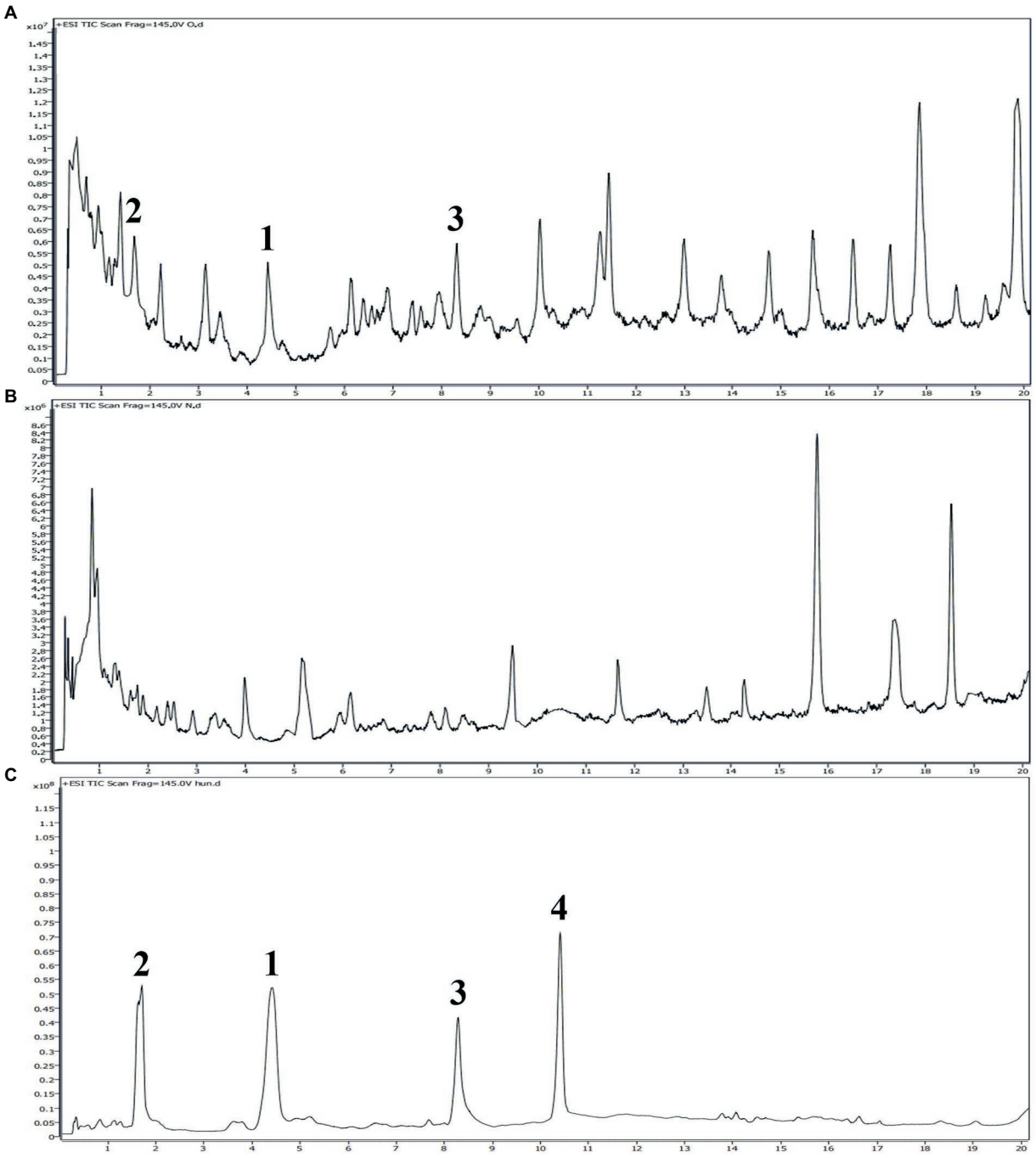


FIGURE 7
UPLC-MS/MS analyses of the root secretion, rainwater leaching, and isolated compounds from *Ambrosia artemisiifolia*. **(A)** TIC chromatogram of the root secretion from *A. artemisiifolia*; **(B)** TIC chromatogram of the rainwater leaching from *A. artemisiifolia*; **(C)** TIC chromatogram of compounds **1**–**4** from *A. artemisiifolia*.

the active compounds **1**–**3** were released into environment through root secretion to negatively affect the growth of native plants. Since the content of compounds **1**–**3** in the rhizosphere soil is low, we speculated that *A. artemisiifolia* could accumulate these compounds to achieve effective inhibitory concentrations only under certain conditions (e.g., when competing with other weeds; Kong et al., 2018). However, their dynamic release process remains unclear in the actual environment. Compound **4** was not detected in root secretion, possibly because of its rapid degradation by soil microorganisms or low water solubility. In addition, the lack of compounds in the rainwater leaching may be caused by the short washing times with distilled water, the replacement of distilled

water for rainwater, or the absence of compounds **1**–**4** in the aerial part of *A. artemisiifolia*.

In recent years, invasive plants have brought serious ecological, economic and social problems around the world. To effectively control these alien plants, it is of great theoretical and practical significance to clarify their invasion mechanisms (Enders et al., 2020). Numerous studies have shown that allelopathy is a driving factor in the successful invasion of alien plants (Yuan et al., 2021). For example, alien plants can release allelopathic substances that are relatively novel to native plants, thus negatively affecting on the growth of native plants. Sesquiterpenoids are abundant in many invasive plants, which play an important role in plant allelopathy and can significantly inhibit plant growth (Bennett and Wallsgrove, 1994; Chadwick et al., 2013) In our study, three eudesmane-type sesquiterpenes (**1**–**3**) with potent allelopathic effects were isolated and identified from *A. artemisiifolia*. These active compounds can be released into environment through the root secretion pathway to affect the cell viability of surrounding plants, thus promoting the invasion of *A. artemisiifolia*. We speculated that native plants have not yet adapted to these compounds (**1**–**3**) because of their novel structures and the lack of coevolution between alien and native plants. However, further research is needed to determine whether these compounds are already present in the original areas of invasive plants and how they specifically affect local plants. In addition, sesquiterpenoids generally have strong phytotoxic activities against many weeds (Anese et al., 2015; Masi et al., 2020). Moreover, allelopathic substances have the advantages of safety, easy degradation and no resistance to weed control (Macías et al., 2019). Therefore, the allelopathic sesquiterpenes (**1**–**3**) from *A. artemisiifolia* can be used to develop novel botanical herbicides and reduce the use of synthetic herbicides.

Conclusion

In summary, four sesquiterpenes (**1**–**4**), consisting of three eudesmane-type sesquiterpenes, one guaiane-type sesquiterpene, were isolated from the whole plant of *A. artemisiifolia* by a variety of column chromatography techniques. Their planar and stereoscopic structures were identified using HR-ESIMS, 1D-NMR, 2D-NMR, and ECD. All the compounds exhibited different levels of allelopathic effects on three native plants (*S. viridis*, *D. sanguinalis*, *C. album*) and one model plant (*A. thaliana*), but in particular, compound **1** significantly inhibited the root elongation of plants at 100 μM. In addition, active compound **1** decreased cell viability by exerting allelopathic effects, as observed by FDA/PI staining assay. Furthermore, the eudesmane-type sesquiterpenes (**1**–**3**) were mainly released into environment through the root secretion pathway, which was revealed by UPLC-MS/MS analyses. Our findings not only helped to reveal the invasion mechanism of *A. artemisiifolia* from the perspective of allelopathy, but also supported a promising strategy for its exploitation in herbicides.

Data availability statement

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

Author contributions

ZL designed the research, performed the experiments, analyzed the data, and wrote the manuscript. NZ, XM, TZ, XL, and GT performed the experiments and analyzed the data. YF and TA reviewed and revised the draft. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Natural Science Foundation of China (31700472).

Acknowledgments

We wish to thank Shihong Luo, Juan Hua, and Mingchao Liu from Shenyang Agricultural University for their assistance in allelopathic assay and UPLC-MS/MS analyses.

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.2022.996498/full#supplementary-material

References

Abdel-Mogi, M., Jakupovi, J., Dawidar, A. M., Metwallyb, M. A., and Abou-Elzahabb, M. (1989). Glaucolides from *Achillea fragrantissima*. *Phytochemistry* 28, 3528–3530. doi: 10.1016/0031-9422(89)80381-4

An, J. P., Ha, T. K. Q., Kim, H. W., Ryu, B., Kim, J., Park, J., et al. (2019). Eudesmane glycosides from *Ambrosia artemisiifolia* (common ragweed) as potential neuroprotective agents. *J. Nat. Prod.* 82, 1128–1138. doi: 10.1021/acs.jnatprod.8b00841

Anese, S., Jatobá, L. J., Grisi, P. U., Gualtieri, S. C. J., Santos, M. F. C., and Berlinck, R. G. S. (2015). Bioherbicidal activity of drimane sesquiterpenes from *Drimys brasiliensis* Miers roots. *Ind. Crop Prod.* 74, 28–35. doi: 10.1016/j.indcrop.2015.04.042

Bais, H. P. (2003). Allelopathy and exotic plant invasion: from molecules and genes to species interactions. *Science* 301, 1377–1380. doi: 10.1126/science.1083245

Bennett, R. N., and Wallsgrove, R. M. (1994). Secondary metabolites in plant defence mechanisms. *New Phytol.* 127, 617–633. doi: 10.1111/j.1469-8137.1994.tb02968.x

Bonea, D., Bonciu, E., Niculescu, M., and Olaru, A. L. (2018). The allelopathic, cytotoxic and genotoxic effect of *Ambrosia artemisiifolia* on the germination and root meristems of *Zea mays*. *Caryologia* 71, 24–28. doi: 10.1080/00087114.2017.1400263

Bonnamour, A., Gippet, J. M. W., Bertelsmeier, C., and Gurevitch, J. (2021). Insect and plant invasions follow two waves of globalisation. *Ecol. Lett.* 24, 2418–2426. doi: 10.1111/ele.13863

Callaway, R. M., and Ridenour, W. M. (2004). Novel weapons: invasive success and the evolution of increased competitive ability. *Front. Ecol. Environ.* 2, 436–443. doi: 10.1890/1540-9295(2004)002[0436:nwisat]2.0.co;2

Caloprisco, E., Fourneron, J. D., Faure, R., and Demarne, F. E. (2002). Unusual lactones from *Cananga odorata* (Annonaceae). *J. Agric. Food Chem.* 50, 78–80. doi: 10.1021/jf0105079

Chadwick, M., Trewin, H., Gawthrop, F., and Wagstaff, C. (2013). Sesquiterpenoids lactones: benefits to plants and people. *Int. J. Mol. Sci.* 14, 12780–12805. doi: 10.3390/ijms140612780

Ding, W. B., Huang, R., Zhou, Z. S., and Li, Y. Z. (2015). New sesquiterpenoids from *Ambrosia artemisiifolia* L. *Molecules* 20, 4450–4459. doi: 10.3390/molecules20034450

Enders, M., Havemann, F., Ruland, F., Bernard-Verdier, M., Catford, J. A., Gómez-Aparicio, L., et al. (2020). A conceptual map of invasion biology: integrating hypotheses into a consensus network. *Glob. Ecol. Biogeogr.* 29, 978–991. doi: 10.1111/geb.13082

Kong, C. H., Zhang, S. Z., Li, Y. H., Xia, Z. C., Yang, X. F., Meiners, S. J., et al. (2018). Plant neighbor detection and allelochemical response are driven by root-secreted signaling chemicals. *Nat. Commun.* 9:3867. doi: 10.1038/s41467-018-06429-1

Lehoczky, E., Glya, G., Szab, R., and Szalai, A. (2011). Allelopathic effects of ragweed (*Ambrosia artemisiifolia* L.) on cultivated plants. *Commun. Agric. Appl. Biol. Sci.* 76, 545–549.

Li, Y. L., Zhu, R. X., Li, G., Wang, N. N., Liu, C. Y., Zhao, Z. T., et al. (2019). Secondary metabolites from the endolichenic fungus *Ophiosphaerella korrae*. *RSC Adv.* 9, 4140–4149. doi: 10.1039/c8ra10329a

Liu, Z. X., Wang, M. Q., Tian, M. X., Yuan, L. L., Yu, B. M., Qu, B., et al. (2021). Pyrrole alkaloids from *Solanum rostratum* and their chemical defense function against *Henosepilachna vigintioctomaculata*. *Fitoterapia* 155:105031. doi: 10.1016/j.fitote.2021.105031

Macías, F. A., Mejías, F. J. R., and Molinillo, J. M. G. (2019). Recent advances in allelopathy for weed control: from knowledge to applications. *Pest Manag. Sci.* 75, 2413–2436. doi: 10.1002/ps.5355

Masi, M., Pannacci, E., Santoro, E., Zermane, N., Superchi, S., and Evidente, A. (2020). Stoechanones A and B, phytotoxic copaane sesquiterpenoids isolated from *Lavandula stoechas* with potential herbicidal activity against *Amaranthus retroflexus*. *J. Nat. Prod.* 83, 1658–1665. doi: 10.1021/acs.jnatprod.0c00182

Pan, J. W., Zhu, M. Y., and Chen, H. (2001). Aluminum-induced cell death in root-tip cells of barley. *Environ. Exp. Bot.* 46, 71–79. doi: 10.1016/s0098-8472(01)00083-1

Simonaitis, R., and Pitts, J. N. (1969). Molecular structure and photochemical reactivity. XI. Wavelength and methyl substituent effects in the photochemistry of lactones in the liquid phase. *J. Am. Chem. Soc.* 91, 108–112. doi: 10.1021/ja01029a023

Sun, Y., and Roderick, G. K. (2019). Rapid evolution of invasive traits facilitates the invasion of common ragweed, *Ambrosia artemisiifolia*. *J. Ecol.* 107, 2673–2687. doi: 10.1111/1365-2745.13198

Taglialatela-Scafati, O., Pollastro, F., Minassi, A., Chianese, G., Petrocellis, L. D., Marzo, V. D., et al. (2012). Sesquiterpenoids from common ragweed (*Ambrosia artemisiifolia* L.), an invasive biological polluter. *Eur. J. Org. Chem.* 2012, 5162–5170. doi: 10.1002/ejoc.201200650

Thorpe, A. S., Thelen, G. C., Diaconu, A., and Callaway, R. M. (2009). Root exudate is allelopathic in invaded community but not in native community: field evidence for the novel weapons hypothesis. *J. Ecol.* 97, 641–645. doi: 10.1111/J.1365-2745.2009.01520.X

Vidotto, F., Tesio, F., and Ferrero, A. (2013). Allelopathic effects of *Ambrosia artemisiifolia* L. in the invasive process. *Crop Prot.* 54, 161–167. doi: 10.1016/j.cropro.2013.08.009

Wang, Y., Luo, S. H., Hua, J., Liu, Y., Jing, S. X., Li, X. N., et al. (2015). Capitate glandular trichomes of *Paragutzlaffia henryi* harbor new phytotoxic labdane diterpenoids. *J. Agric. Food Chem.* 63, 10004–10012. doi: 10.1021/acs.jafc.5b04113

Yan, Z. Q., Wang, D. D., Cui, H. Y., Zhang, D. H., Sun, Y. H., Jin, H., et al. (2016). Phytotoxicity mechanisms of two coumarin allelochemicals from *Stellera chamaejasme* in lettuce seedlings. *Acta Physiol. Plant.* 38:248. doi: 10.1007/s11738-016-2270-z

Yuan, L., Li, J. M., Yu, F. H., Oduor, A. M. O., and Kleunen, M. V. (2021). Allelopathic and competitive interactions between native and alien plants. *Biol. Invasions* 23, 3077–3090. doi: 10.1007/s10530-021-02565-w

Zhao, F., Durner, J., Winkler, J. B., Traidl-Hoffmann, C., Strom, T. M., Ernst, D., et al. (2017). Pollen of common ragweed (*Ambrosia artemisiifolia* L.): Illumina-based de novo sequencing and differential transcript expression upon elevated NO_2/O_3. *Environ. Pollut.* 224, 503–514. doi: 10.1016/j.envpol.2017.02.032

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fpls.2022.1000705

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre,
United Kingdom

REVIEWED BY
Orsolya Kinga Gondor,
Hungarian Academy of Sciences (MTA),
Hungary
Masaru Nakayasu,
Kyoto University,
Japan

*CORRESPONDENCE
Ki Hun Park
khpark@gnu.ac.kr

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

RECEIVED 22 July 2022
ACCEPTED 26 August 2022
PUBLISHED 26 September 2022

CITATION
Kim JH, Shah AB, Lee YH, Baiseitova A, Ban YJ and Park KH (2022) Changes in secondary metabolites in soybean (*Glycine max* L.) roots by salicylic acid treatment and their anti-LDL oxidation effects. *Front. Plant Sci.* 13:1000705. doi: 10.3389/fpls.2022.1000705

Changes in secondary metabolites in soybean (*Glycine max* L.) roots by salicylic acid treatment and their anti-LDL oxidation effects

Jeong Ho Kim, Abdul Bari Shah, Yong Hyun Lee, Aizhamal Baiseitova, Yeong Jun Ban and Ki Hun Park*

Division of Applied Life Science (BK21 Plus), IALS, Gyeongsang National University, Jinju, South Korea

Abundance of metabolites in plant is a critical factor toward being functional food stuff. Salicylic acid (SA) treatment led significant changes in levels of the secondary metabolites in soybean roots. Notably, the exposure of 3mM of SA aqueous solution to soybean plants for 24h resulted in distinctive increases in the levels of coumestrol (16-fold, 0.3–4.8mg/g DW) and daidzein (7-fold, 1.2–8.9mg/g DW) in roots part. These changes were systematically investigated by LC-ESI-TOF/MS analysis to afford a clear difference of PLS-DA score, heatmap, and box plots. Quantitative analysis showed that SA treatment played to stimulate biosynthesis of coumestrol as well as hydrolysis of its glycosides (coumestrin and malonylcoumestrin). The highly improved anti-LDL oxidation effect was observed in the SA treated soybean roots in the three different assay systems. It might be rationalized by the increased levels of coumestrol and daidzein.

KEYWORDS

soybean roots, coumestrol, daidzein, salicylic acid, anti-LDL oxidation

Introduction

The health benefits of plant-based resources may be related to the presence and abundance of bioactive compounds in plants (Baenas et al., 2014). The abundance level of bioactive compounds can be increased with different tools such as molecular breeding (MB), genetically modified organisms (GMOs), and the application of elicitation tools. The MB and GMO approaches have practical limits in terms of the time they require to implement and their biosafety, respectively (Prakash et al., 2011).

A large number of studies have proven that elicitation seems to be a promising alternative to other conventional techniques for raising the levels of bioactive components in plants (Pérez-Tortosa et al., 2012; Anusuya and Sathiyabama, 2016; Wang et al., 2020). In fact, elicitors are widely used for increasing the level of secondary metabolites in plant

cells, tissues, or the whole plant, and a diverse range of elicitors, such as phytohormones, polysaccharides, amino acids, and organic acids, have been used (Liu et al., 2019). Elicitations have been reported to enhance the level of phenolic compounds (Yamamoto et al., 2020), glucosinolates (Schreiner et al., 2011), vitamins (No et al., 2003), and γ-aminobutyric acid (Hao et al., 2016). However, few studies concerned with sufficiently enhancing the levels of bioactive compounds such that they are useful for nutraceutical purposes, have been reported.

Plant roots are located in the region close to the rhizosphere, where various interactions with microbial communities occur through phytochemicals (Bakker et al., 2013). In particular, soybean (*glycine max* L.) is a representative plant that interacts with microorganisms in the rhizosphere (Sugiyama, 2019). Soybean roots contain a series of isoflavones and pterocarpans with diverse biological benefits (Yuk et al., 2011; Ban et al., 2020). This study focused on increasing the bioactive metabolites in soy roots with the aid of salicylic acid (SA) belonging to plant hormones (Brunswick, 1992). SA is well known to elicit the production of secondary metabolites in response to many pathogens (Maruri-López et al., 2019; Ding and Ding, 2020). There is a report that SA influenced changes in phenolic metabolites of soybean seedlings (Durango et al., 2018).

Coumestrol is a phytoalexin in the soybean plant and is associated with insect attacks and senescence to which SA as a phytohormone connects (Morandi et al., 1984; Boué et al., 2000; Mun et al., 2021). In addition, coumestrol is considered as a valuable phytochemical with diverse biological activities such as its anti-low density lipoprotein (anti-LDL) oxidation (Jin et al., 2006), anti-inflammation (Yuk et al., 2017), anti-cancer (Zafar et al., 2017), anti-obesity (Kim et al., 2020b), and skin protection (Park et al., 2015) benefits. Daidzein is another important phytochemical that is highly sought after in the nutraceutical field owing to its antioxidant properties (Dwiecki et al., 2009), its use as a plant-based alternative to estrogen (Vitale et al., 2013), and its transformation potential to equal (Legette et al., 2014).

On the other hand, low density lipoprotein (LDL) is easily oxidized to oxLDL by reactive oxygen species (ROS; Chen et al., 2003). LDL oxidation prompts an early atherosclerosis process through several steps including endothelial cell damage, a form of cell accumulation, and the synthesis of autoantibodies (Pirillo, 2013; Mundi et al., 2018). Thus, substances with anti-LDL oxidative properties have been attractive natural sources of nutraceuticals. Pterocarpans, including coumestrol, were reported to display significant anti-LDL oxidation activities (Jin et al., 2006). These results, together with results from previous reports, suggest that coumestrol enriched soybean roots would be potentially useful because of their anti-LDL oxidation effects.

In this study, we observed the changes in the metabolites in soybean roots by treating the roots with SA to afford coumestrol and daidzein enriched soybean roots (CDESR). The detailed changes in the metabolites were systematically identified by LC-ESI-Q-TOF/MS and reported as the PLS-DA score, heatmap, and box plots. The anti-LDL oxidation effects of the CDESR were assessed with three different assay systems.

Materials and methods

Plant material and experimental design

Soybean [*G. max* (L.) Merrill] seeds were obtained from the National Institute of Crop Science (NICS), Miryang, Republic of Korea. The soybean plants were cultivated in pots in a greenhouse over a period of about 30 days until the plant reached the V3 growth stage (three nodes on the main stem with fully developed leaves). Experiments were conducted by exposing the plants to two different chemicals [SA and sodium salicylate (SS)]. After rinsing the roots with water, the plants were soaked in an aqueous solution of SA or SS at five concentrations: 1, 2, 3, 4, or 5 mM, following which the plants were placed in a chamber maintained at 25°C and a 16 h photoperiod for the duration of the treatment. The control plants were soaked in aqueous solution without SA and maintained for same time with treatment groups (Supplementary Figure S1). The plant roots were harvested 5 times (0, 3, 6, 12, and 24 after treatment) for time dependent effects. Samples for metabolomic analysis were collected randomly with 9 replications. The harvested samples were dried at a temperature of 35°C, milled to a powder in an electric mill and stored at-20°C for further experiments.

HPLC-DAD analysis of phenolic metabolites

The milled soybean root samples (1.0 g) were extracted using 80% ethanol (50 ml) with sonication for 3 h at 30°C. After extraction, the samples were centrifuged at 5000 rpm for 15 min. The supernatants were filtered with a syringe filter (0.2 μm) and analyzed by HPLC. HPLC analysis of the extracts was conducted using an Agilent 1,200 series HPLC system including an Agilent G1311 A quaternary pump (Agilent Technologies, Santa Clara, CA) and Agilent G1315B diode array detector connected to a Zolbax Bonus-RP (150 mm × 4.6 mm, 5 μm) column (Agilent Technologies, Santa Clara, CA). Data were collected and analyzed with Chem Station software (Agilent Technologies). A linear gradient elution was carried out with water/0.1% acetic acid (A) and acetonitrile (B). The procedures for the gradient elution were as follows: 5 min, 85% A; 20 min, 80% A; 50 min, 50% A; 65 min, 40% A; 75 min, 0% A. The flow rate was 1.0 ml/min, 10 μl of each sample was injected into the column, and the absorbance was detected at 254 or 340 nm using the Agilent diode array detector. The phenolic compounds in the extracts were identified by comparison of the retention times (t_R) of the compounds with the HPLC profiles of commercial standards and quantified by using their calibration curves obtained by determining the peak areas from the chromatograms (Supplementary Figure S2). The standards

of daidzin, daidzein, genistin, genistein, and coumestrol were purchased from Sigma-Aldrich Co., Ltd. (St. Louis, MO, United States), and those of malonyldaidzin and malonylgenistin were purchased from MedChemExpress (NJ, United States).

UPLC–Q-TOF/MS analysis

The metabolomic profiles in various extracts of soybean roots were analyzed on an ultra-performance liquid chromatography (UPLC) system manufactured by Waters Technologies (Waters, Milford, MA, United States). The metabolites were separated in an Acquity UPLC BEH C 18 column (2.1 × 100 mm, 1.7 μm, Waters) with a flow rate of 0.35 ml/min at 30°C. The mass spectrometer was operated in positive electrospray ionization (ESI) mode and the TOF-MS and MS/MS results of the metabolites were collected in the *m*/*z* 50–1,500 range with a scan time of 0.2 s. The other mass spectrometric conditions were configured as follows: temperature of source, 100°C; desolvation, 400°C; sampling cones voltage, 30 V; capillary voltage, 3 kV; flow rate of desolvation, 800 l/h; collision energy, 10 to 30 eV. The mobile phase was comprised of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in acetonitrile) using a gradient elution of 0–1 min, 3% B; 1–5 min, 3—15% B; 5–10 min, 15—25% B; 10–12 min, 25—45% B; and 12–20 min, 45–100% B. For quality control (QC), a mixture of all samples was injected after every 6 samples. The MS data including the retention time (t_R) and *m*/*z* were obtained using MassLynx software (Waters) and the ion intensities were acquired by MarkerLynx software (Waters).

Data processing and statistical analysis

The data sets from the UPLC-ESI-Q-TOF/MS analysis were collected, aligned, and normalized by MarkerLynx software (Waters). The peak of each metabolite was collected as follows: peak width at 5% height of 0.5 s, and peak-to-peak baseline noise of 1,000. The respective metabolites were identified based on Chemspider,[1] the METLIN database (metlin.scripps.edu), and human metabolome databases.[2] The aligned and normalized LC/MS data sets were multivariately analyzed with SIMCA-P$^+$ version 16.0.2 (Umetrics, Umeå, Sweden). Partial least squares discriminant analysis (PLS-DA) was carried out to visualize discrimination between the control group and the groups exposed to SA. Hotelling's T2 test was employed to exclude outliers from the 95% confidence region. The quality of the PLS-DA model was evaluated using the following three parameters: R^2X, R^2Y, and Q^2Y. The improvement of fit was quantified by R^2X and R^2Y and the predictive ability was determined by Q^2Y. The PLS-DA models were validated using 7-fold cross validation and the reliability was evaluated by a permutation test ($n = 200$). The metabolites contributing to the discrimination in the sample groups were found and then determined by variable significance in the projection (VIP) value >0.7, intended by PLS-DA with one-way analysis of variance (ANOVA) with Duncan's test ($p < 0.05$) using SPSS 17.0 (SPSS Inc., United States). The selected metabolites with significant differences ($p < 0.05$) were also visualized in a heat map described using *R* with gplots.

Measurements of thiobarbituric acid reactive substances

The thiobarbituric acid reactive substances (TBARS) assay was conducted according to an existing method with minor alteration (Jin et al., 2006). The absorbance of the product of the reaction between malondialdehyde (MDA) and thiobarbituric acid (TBA) was measured at 532 nm. Briefly, 10 μl of 3 mg/ml of LDL (ProSpec, Rehovot, Israel) solution in 220 μl of PBS buffer (10 mM, pH 7.4), 10 μl of 0.25 mM $CuSO_4$, and 10 μl of different concentrations of compounds and soybean root extracts or probucol as a positive control were mixed in a 1.5 ml Eppendorf tube and incubated at 37°C for 24 h. After incubation, 100 μl of trichloroacetic acid (20% w/v) and 100 μl of 0.67% thiobarbituric acid (0.67% w/v dissolved in 0.05 N NaOH) were added to the mixture and heated in the water bath at 100°C for 30 min. After heating, the mixture was cooled on ice and centrifuged for 10 min at 5000 rpm. The absorbance of the supernatant was determined and the IC_{50} values of each sample were calculated using the inhibition rate obtained by Equation (1).

$$\text{Inhibition rate}(\%) = 100 \times \left[\left(\text{rate of control} - \text{rate of inhibitor} \right) / \text{rate of control} \right] \quad (1)$$

Measurements of conjugated diene formation

The formation of conjugated diene (CDs) from oxLDL was measured using a previously reported method (Xu et al., 2007). First, 1 ml of 50 μg/ml of LDL solution in PBS buffer (10 mM, pH 7.4) was incubated with different concentrations of compounds, soybean root extract, or probucol as a positive control. The control experiments were carried out under equivalent assay conditions but without adding the above compounds or extracts. The oxidation commenced immediately after adding $CuSO_4$ such that the final concentration was 10 μM. The formation of CDs was monitored at 234 nm for 240 min in increments of 5 min, continuously, using a SpectraMax® M3 multi-mode microplate

1 www.chemspider.com

2 www.hmdb.ca

reader (Molecular Devices, CA, United States). On the acquired absorbance spectra, the lag time of each sample was observed as the endpoint indicated in the lag phase.

Relative electrophoretic mobility

The relative electrophoretic mobility (REM) of native LDL and oxidized LDL induced by Cu^{2+} was measured according to procedures in previous reports (Kim et al., 2020a). Before electrophoresis, the reaction mixtures including the 10 μM $CuSO_4$ and the compounds, samples of soybean root extract, or probucol as a positive control at various concentrations were added to the LDL solution (500 μg/ml), then incubated at 37°C for 16 h in the dark. The native LDL and premixed LDL samples were loaded onto 0.5% agarose gel in TAE buffer (40 mM Tris-acetate with 1 mM EDTA, pH 8.0) and electrophoresed for 50 min at 100 V. After the electrophoresis, the gel was fixed in 40% ethanol with 10% acetic acid for 30 min, stained with 0.15% coomassie brilliant blue R250, and the LDL bands were visualized by a destaining solution (methanol:acetic acid:water, 3:2:15, v/v).

Results and discussion

Effects of the application of salicylic acid to soybean roots

Soybean plants secrete unique metabolites from their roots to maintain a symbiotic relationship with rhizobia (Han et al., 2020) and mycorrhiza fungi (Pawlowski et al., 2020). The major metabolites in the roots are mostly isoflavones and their derivatives, all of which reportedly are beneficial for human health (Wang et al., 2013). In this study, soybean roots were observed to undergo a distinctive ingredient change upon exposure to SA. We applied SA aqueous solutions ranging 1 to 5 mM at the V3 growth stage (30 days after sowing). The level of metabolites were changed according to SA concentrations and exposure time (Supplementary Figures S3–S5). The optimal condition were observed 3 mM of SA solution and 24 h of exposure time. The visible appearance of the soybean plant did not differ much during its exposure to SA under these conditions as shown in Figure 1; Supplementary Figure S1.

Treatment with SA afforded a dramatic increase in the levels of daidzein and coumestrol in the roots, as was evident from the red band on the HPLC chromatogram (Figures 2A,B) and the results of the quantitative analysis (Table 1). Their specific changes are discussed in the section on metabolomic analysis (*vide infra*). We assumed that daidzein (peak 5) and coumestrol (peak 9) emerged as a result of the hydrolysis of their corresponding glycosides, namely daidzin (peak 1), malonyldaidzin (peak 4), coumestrin (peak 3), and malonylcoumestrin (peak 8). In particular, the content of coumestrol was increased 16-fold to be 4.76 mg/g DW from 0.30 mg/g DW. Enzymatic hydrolysis experiments indicated that the contents of coumestrin (peak 3) and malonylcoumestrin (peak 8) were equivalent to 2.14 mg/g DW of coumestrol as shown in Supplementary Figure S6. This may be explained by the activation of β-glucosidase in the chemical defense process of the plant by signal molecules including SA (Morant et al., 2008). The SA application also stimulated production of coumestrol which was *ca.* 55% of total amount (2.62 mg/g DW). Meanwhile, SS also effectively enhanced the daidzein and coumestrol content, but the induction yields were 55% and 71%, respectively, compared with the yields after SA treatment (Figures 3A,B). In the presence of SA, the soybean roots started accumulating coumestrol and daidzein, the concentrations of which increased rapidly and continuously for 24 h as shown in Supplementary Figures S4, S5. Figure 3C shows the dependence of the coumestrol content on the

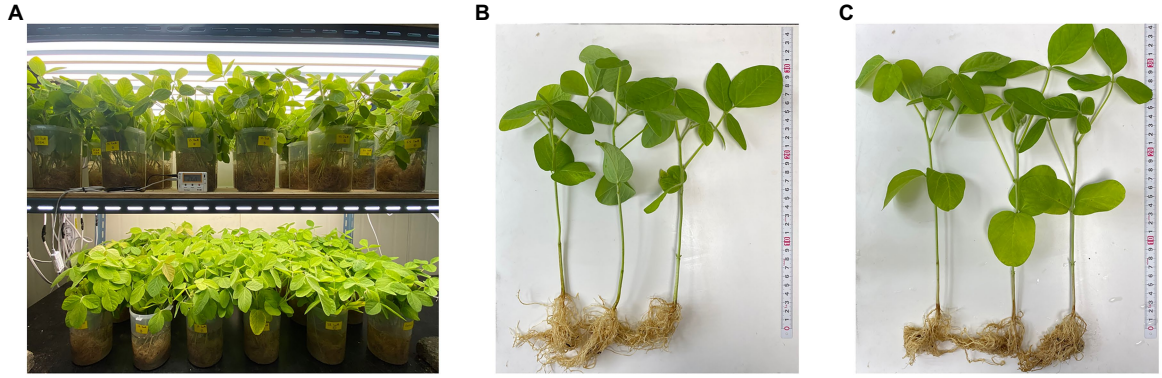


FIGURE 1
(A) Process of salicylic acid (SA) treatment on soybean plants. **(B)** Control soybean plants. **(C)** Soybean plants after treatment with 3mM of SA for 24h.

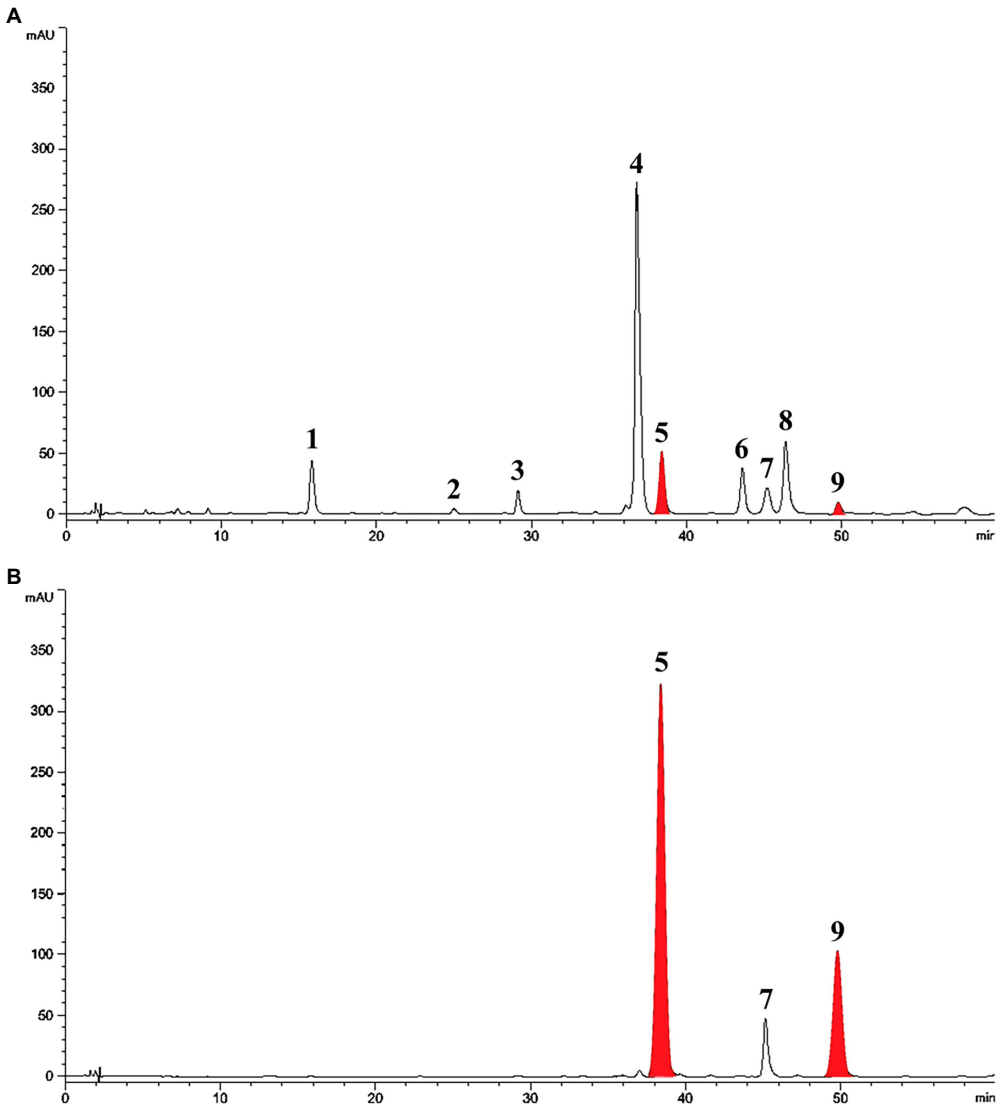


FIGURE 2
HPLC chromatographic patterns of soybean root extracts at 254nm. Peak 1, daidzin; peak 2, genistin; peak 3, coumestrin; peak 4, malonyldaidzin; peak 5, daidzein; peak 6, malonylgenistin; peak 7, genistein; peak 8, malonylcoumestrin; peak 9, coumestrol. **(A)** Control soybean roots. **(B)** SA treated soybean roots.

exposure time to SA: 0 h (0.30 mg/g DW), 3 h (1.35 mg/g DW), 6 h (2.75 mg/g DW), 12 h (4.73 mg/g DW), and 24 h (4.76 mg/g DW). Quantification of phytoestrogens in Legumes (Franke et al., 1994) indicated that most of plants possess a trace amount of coumestrol. The most coumestrol abundant plant sources are clover sprout (5.6 mg/g DW) and alfalfa sprout (0.7 mg/g DW), but they have no daidzein. When we consider the levels of daidzein and coumestrol together, the CDESR might be the plant source with the highest level of coumestrol and daidzein.

SA is well known to be induced in plants in response to many pathogens and then this elicits the production of secondary metabolites (Ding and Ding, 2020). Many case studies of the application of SA have been conducted with the aim of enhancing the production of stilbenes (Xu et al., 2015), alkaloids (Figueroa-Pérez et al., 2015), anthraquinones (Lee et al., 2013), terpenoids (Wen Xu, 2012), etc. Our results would be a comparative case study regarding SA application to introduce secondary metabolites in plants.

Metabolomic analysis of soybean roots with UPLC-ESI-Q-TOF/MS

The metabolomic analysis was designed to determine the changes in the metabolites in soybean roots resulting from treatment with SA. In this regard, 80% ethanol was found to be appropriate solvent for maximal and well-balanced extraction

TABLE 1 Contents (mg/g DW) of dietary phytoestrogens in soybean roots.

Compound	Control[a]	CDESR[b]
Daidzin	0.92[c] ± 0.22	ND[d]
Malonyldaidzin	14.77 ± 2.55	ND
Daidzein	1.16 ± 0.2	8.87 ± 1.06
Genistin	0.12 ± 0.01	ND
Malonylgenistin	1.86 ± 0.15	ND
Genistein	0.11 ± 0.03	0.88 ± 0.07
Coumestrin	NA[d]	ND
Malonylcoumestrin	NA	ND
Coumestrol	0.30 ± 0.04	4.76 ± 0.81

[a]Not treated.
[b]Coumestrol and daidzein enriched soybean roots treated with 3 mM of salicylic acid for 24 h.
[c]All values are mean ± SD of in six independent experiments; Not detected.
[d]Not commercially available.

in view of diverse polarities of metabolites. An analysis of the profiles of metabolites using UPLC-ESI-Q-TOF/MS in positive ion mode led to the detection of a total of 176 metabolites in soybean roots. A comparison of the representative base peak intensity (BPI) chromatogram within 20 min showed that the metabolites in the roots were of a similar type, but the levels of many metabolites differed as shown in Figures 4A–G. The system stability and quality of the obtained data were evaluated to be reliable by correlating the QC samples (Figure 4C). The PLS-DA score plots were prepared from the metabolites with VIP > 0.7 among the sample groups which contributed to separation (Supplementary Table S1).

The 25 major metabolites were annotated by using their molecular weight ($[M+H]^+$), the major fragment of their TOF/MS spectral data, and a comparison with previous reports (Table 2; Jin et al., 2006; Yuk et al., 2011; Ban et al., 2020). First, ten isoflavone peaks were annotated by identical molecular ion

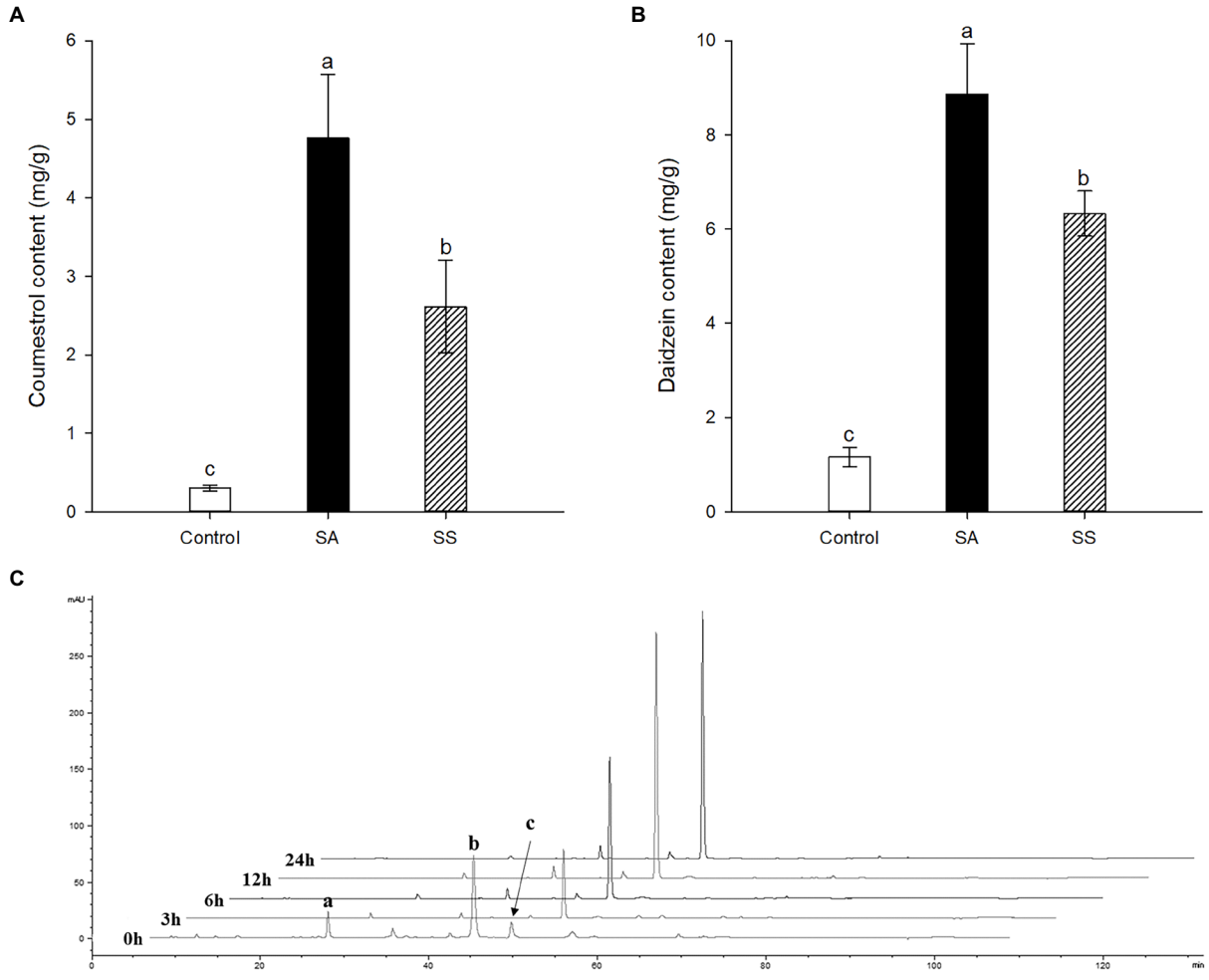


FIGURE 3
(A) Comparison of coumestrol contents between SA and sodium salicylate (SS) treatments. **(B)** Comparison of daidzein contents between SA and SS treatments. **(C)** Changes in HPLC chromatographic patterns by treatment time of SA at 340nm. Peak a, coumestrin; peak b, malonylcoumestrin; peak c, coumestrol. Data are the mean±SD of determinations performed in five replicates in 5 samples. Different letters indicate significant differences, as determined by Tukey's (LSD) test with $p<0.05$.

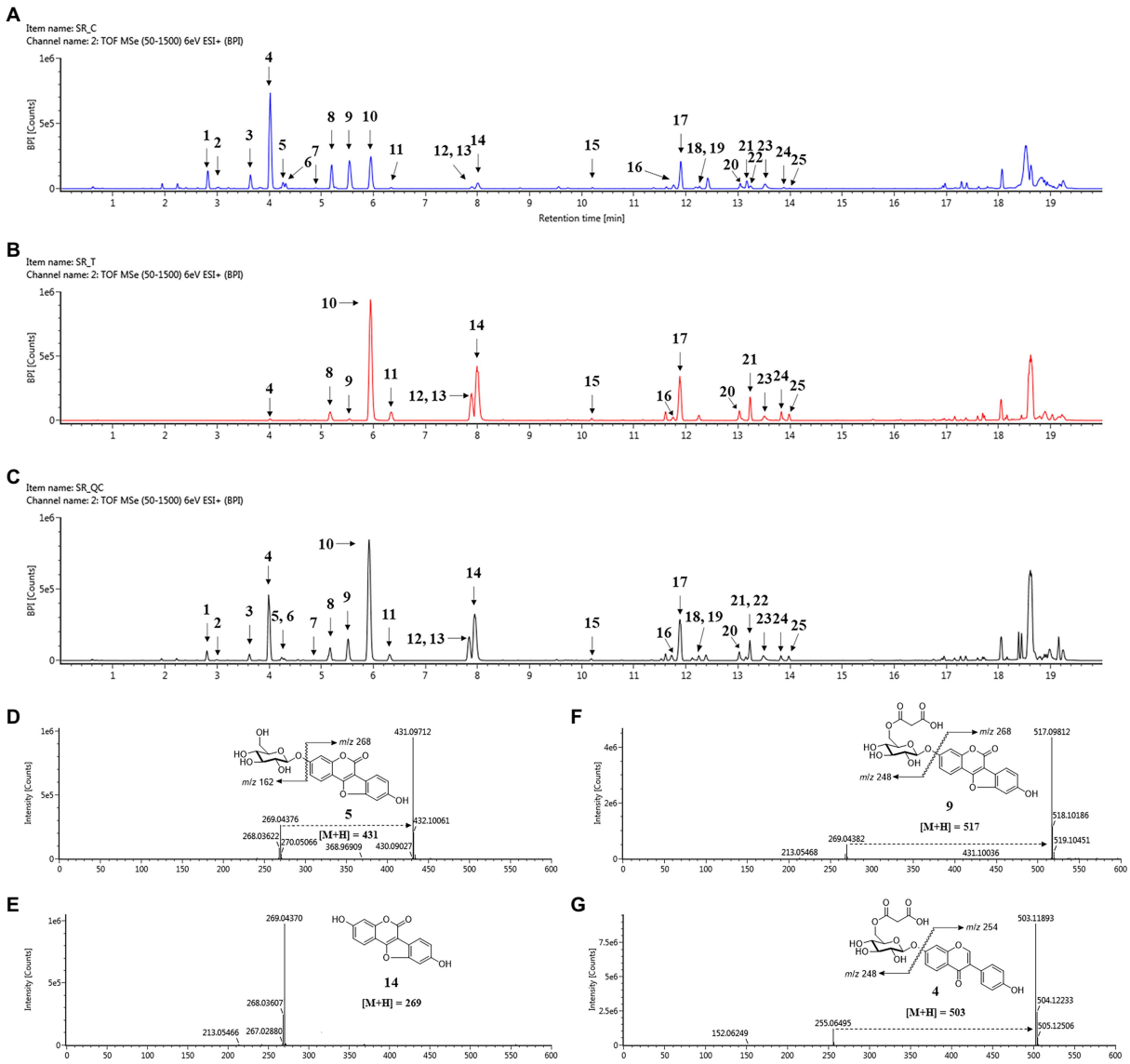


FIGURE 4
Changes in secondary metabolites of soybean roots by SA treatment. **(A)** BPI chromatogram of control soybean roots. **(B)** BPI chromatogram of SA treated soybean roots. **(C)** BPI chromatogram of quality control (QC). **(D–G)** Positive ion mass spectra of representative metabolites of soybean roots: **(D)** Coumestrin, **(E)** Malonylcoumestrin, **(F)** Coumestrol, **(G)** malonyldaidzin. Peak 1, daidzin; peak 2, glycitin; peak 3, genistin; peak 4, malonyldaidzin; peak 5, coumestrin; peak 6, malonylglycitin; peak 7, acetyldaidzin; peak 8, malonylgenistin; peak 9, malonylcoumestrin; peak 10, daidzein; peak 11, glycitein; peak 12, isotrifoliol; peak 13, genistein; peak 14, coumestrol; peak 15, soyasaponin Ab; peak 16, glyceollin I; peak 17, soyasaponin I; peak 18, soyasaponin II; peak 19, soyasaponin III; peak 20, soyasaponin αg; peak 21, soyasaponin βg; peak 22, phaseol; peak 23, soyasaponin βα; peak 24, γg; peak 25, glyceollin IV.

$[M+H]^+$ at *m/z* 417.1176 (peak 1, daidzin), *m/z* 447.1284 (peak 2, glycitin), *m/z* 433.1128 (peak 3, genistin), *m/z* 503.1189 (peak 4, malonyldaidzin), *m/z* 533.1296 (peak 6, malonylglycitin), *m/z* 459.1288 (peak 7, acetyldaidzin), *m/z* 519.1136 (peak 8, malonylgenistin), *m/z* 255.0646 (peak 10, daidzein), *m/z* 285.0753 (peak 11, glycitein), and *m/z* 271.0595 (peak 13, genistein). The seven pterocarpans were confirmed by identical molecular ion $[M+H]^+$ at *m/z* 431.0971 (peak 5, coumestrin), *m/z* 517.0981 (peak 9, malonylcoumestrin), *m/z* 299.0543 (peak 12, isotrifoliol), *m/z* 269.0437 (peak 14, coumestrol), *m/z* 339.1220 (peak 16, glyceollin I), *m/z* 337.1064 (peak 22, phaseol), and *m/z* 355.1534 (peak 25, glyceollin IV). The eight soyasaponin peaks were annotated by identical molecular ion $[M+H]^+$ at *m/z* 1437.6557 (peak 15, soyasaponin Ab), *m/z* 943.5263 (peak 17, soyasaponin I), *m/z* 797.4693 (peak 18, soyasaponin III), *m/z* 913.5161 (peak 19, soyasaponin II), *m/z* 1085.5523 (peak 20, soyasaponin αg), *m/z* 1069.5565 (peak 21, soyasaponin βg), *m/z* 1039.5466 (peak 23, soyasaponin βα), and 923.5010 (peak 24, soyasaponin γg). The specific MS data are summarized in Table 2; Supplementary Figure S7.

TABLE 2 Peak assignments of secondary metabolites in soybean roots by UPLC–ESI-TOF/MS.

Peak	t_R (min)	Elemental composition	Molecular ion $[M+H]^+$ (m/z)	Observed fragment ions $[M+H]^+$ (m/z)	Idendification
1	2.82	$C_{21}H_{20}O_9$	417.1176	255.0647	Daidzin
2	3.08	$C_{22}H_{22}O_{10}$	447.1284	285.0756	Glycitin
3	3.64	$C_{21}H_{20}O_{10}$	433.1128	271.0595	Genistin
4	4.02	$C_{24}H_{22}O_{12}$	503.1189	255.0649	Malonyldaidzin
5	4.27	$C_{21}H_{18}O_{10}$	431.0971	269.0437	Coumestrin
6	4.31	$C_{25}H_{24}O_{13}$	533.1296	285.0757	Malonylglycitin
7	4.91	$C_{23}H_{22}O_{10}$	459.1288	255.0653	Acetyldaidzin
8	5.20	$C_{24}H_{22}O_{13}$	519.1136	271.0596	Malonylgenistin
9	5.54	$C_{25}H_{24}O_{12}$	517.0981	269.0438	Malonylcoumestrin
10	5.95	$C_{15}H_{10}O_4$	255.0646		Daidzein
11	6.34	$C_{16}H_{12}O_5$	285.0753		Glycitein
12	7.86	$C_{16}H_{10}O_6$	299.0543	271.0603, 255.0284	Isotrifoliol
13	7.89	$C_{15}H_{10}O_5$	271.0595		Genistein
14	8.00	$C_{15}H_8O_5$	269.0437		Coumestrol
15	10.20	$C_{67}H_{104}O_{33}$	1437.6557	975.5156, 439.3580, 169.0488	Soyasaponin Ab
16	11.76	$C_{20}H_{18}O_5$	339.1220	321.1113, 305.0811	Glyceollin I
17	11.90	$C_{48}H_{78}O_{18}$	943.5263	423.3620	Soyasaponin I
18	12.26	$C_{42}H_{68}O_{14}$	797.4693	423.3621	Soyasaponin III
19	12.26	$C_{47}H_{76}O_{17}$	913.5161	617.4060, 441.3727, 423.3621	Soyasaponin II
20	13.08	$C_{54}H_{84}O_{22}$	1085.5523	571.2404, 543.2803, 423.3619	Soyasaponin αg
21	13.17	$C_{54}H_{84}O_{21}$	1069.5565	563.2436, 423.3619	Soyasaponin βg
22	13.24	$C_{20}H_{16}O_5$	337.1064	281.0441, 253.0503	Phaseol
23	13.46	$C_{53}H_{82}O_{20}$	1039.5466	743.4381, 423.3616	Soyasaponin βα
24	13.76	$C_{48}H_{74}O_{17}$	923.5010	423.3623	Soyasaponin γg
25	13.88	$C_{21}H_{22}O_5$	355.1534	337.1428, 335.1276, 245.1172	Glyceollin IV

The quality parameters of the model used for the PLS-DA analysis were calculated as having the following values: $R^2X = 0.648$, $R^2Y = 0.995$, and $Q^2 = 0.983$. The permutation validation was reliable and verified the PLS-DA analysis: R^2 intercept = 0.673, Q^2 intercept = 0.169, and *value of* $p < 0.05$. The groups of metabolites in the soybean roots were significantly different on the two-component PLS-DA score plots (Figures 5A,B). The profiling overview of these different metabolites in the pairwise comparison of the two groups is shown in the form of a heatmap (Figure 5C). The control group and that subjected to SA treatment were clearly separated into two clusters as indicated by the green-red color scale, which was obtained from the *z*-score transformed raw data of the metabolites. The 17 metabolites were selected to be significantly affected by exposure to SA based on higher VIP values than 1.0. The 17 different metabolites were assigned to the flavonoid and soyasaponin categories. The metabolites with increased levels were annotated, namely daidzein (**10**), glycitein (**11**), isotrifoliol (**12**), genistein (**13**), coumestrol (**14**), soyasaponin I (**17**), soyasaponin III (**18**), and soyasaponin II (**19**). The decreased metabolites were daidzin (**1**), genistin (**3**), malonyldaidzin (**4**), coumestrin (**5**), malonylglycitin (**6**), malonylgenistin (**8**), malonylcoumestrin (**9**), soyasaponin βg (**21**), and soyasaponin βα (**23**). The box plots in Figure 5D show that daidzein and coumestrol predominantly accumulated in the course of the SA treatment and that the levels of their corresponding glycosides in the soybean roots were significantly lower as a result. The levels of the soyasaponin metabolites were also affected in that those of three of them (**17**, **18**, and **19**) increased and two of them (**21** and **23**) decreased. The metabolomic analysis revealed that treatment with SA promoted significant changes in the levels of metabolites in soybean roots towards enhancing the value of these roots for nutraceutical purposes.

Anti-LDL oxidation effects of soybean roots

In a previous study (Jin et al., 2006), coumestrol showed significant anti-LDL oxidation activity in four different assay systems, namely the TBARS, lag time of CD formation, relative electrophoretic mobility (REM) experiment, and fragmentation pattern of ApoB-100 protein. Daidzein also showed moderate inhibition against LDL oxidation. Coumestrol and daidzein enriched soybean roots (CDESR) were obtained by exposing the roots to SA. The anti-LDL oxidation effects of an 80% ethanol

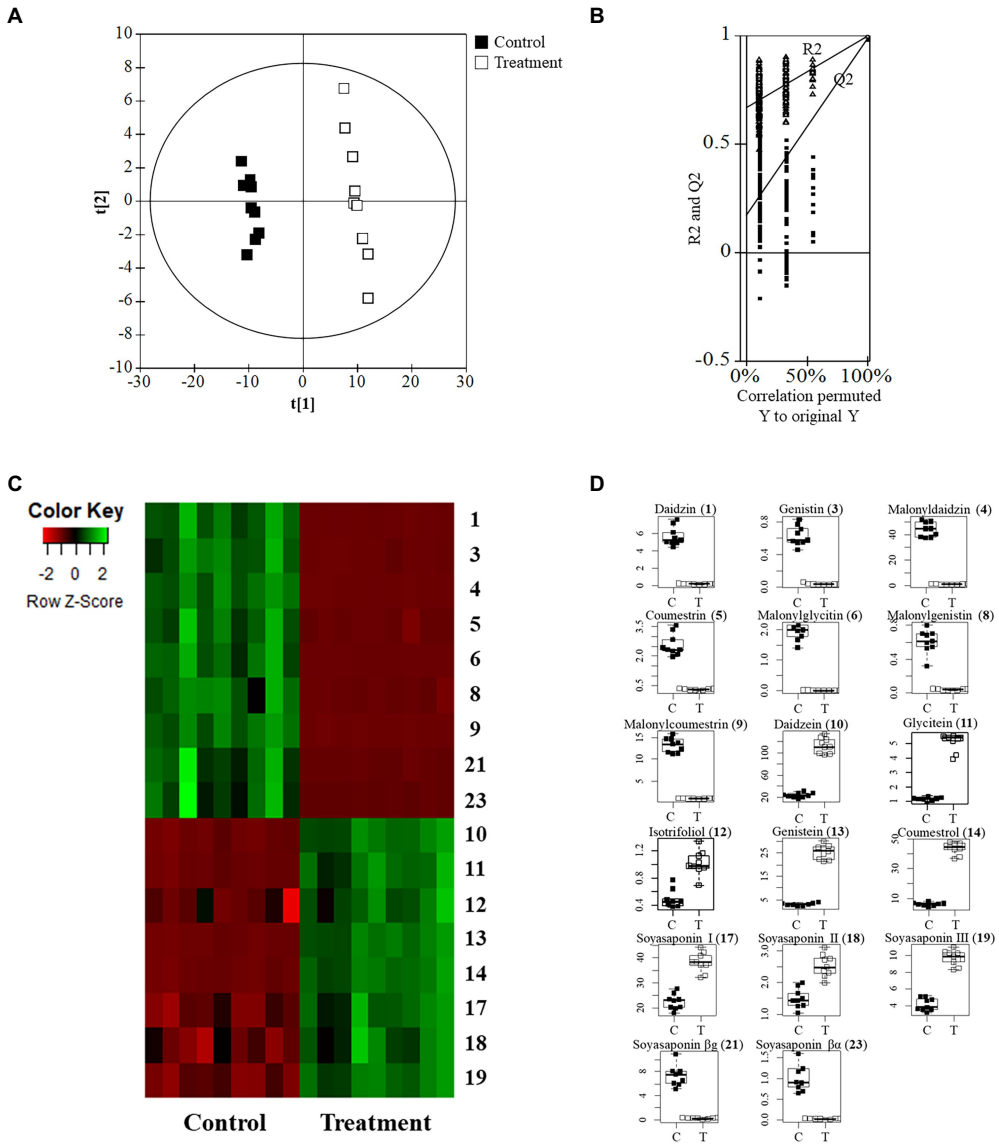


FIGURE 5
Metabolomic analysis of soybean roots by LC-TOF/MS. **(A)** PLS-DA plots of untreated (control) and SA treated roots. C, control group; T, treatment group. **(B)** Performance of the permutation tests validated from the PLS-DA model. **(C)** Heatmap for the 17 metabolites ($p<0.05$) between control and treatment. **(D)** Boxplots of representative metabolites.

extract of CDESR were examined with TBARS, the lag time of CD formation, and REM experiments. All experiments were carried out with human LDL in the presence of 10 μM Cu^{2+} as an oxidation initiator.

First, the anti-LDL oxidation effects of coumestrol, daidzein, control roots, and CDESR were assessed with a TBARS assay using a concentration of 40 μg/ml (Figure 6A). The CDESR blocked LDL-oxidation efficiently compared with the control roots and daidzein. The CDESR (IC_{50} = 36.1 μg/ml) were 3-fold more effective than the control roots (IC_{50} = 108.9 μg/ml) as shown in Supplementary Table S2. As reported previously (Jin et al., 2006), coumestrol completely inhibited LDL-oxidation at 40 μg/ml. The difference between CDESR and the control roots were doubly confirmed by the dose-dependence curve (Figure 6B).

The anti-LDL oxidation effects of all the entries were additionally assessed by examining their CD formation at 5 μg/ml for 240 min (Figure 6C). The anti-LDL oxidation potencies were demonstrated by determining the lag time in the process of incubation with 10 μM $CuSO_4$. The control oxLDL recorded a lag time of 50 min, whereas probucol as a positive control extended the lag time to 60 min. Similar to the results of the TBARS assay, the anti-LDL oxidation of CDESR was more efficient in terms of its lag time (190 min), exceeding that of the control (115 min). The lag time became longer in proportion to the concentrations of CDESR (1.25–10 μg/ml) dose-dependently, as shown in Figure 6D.

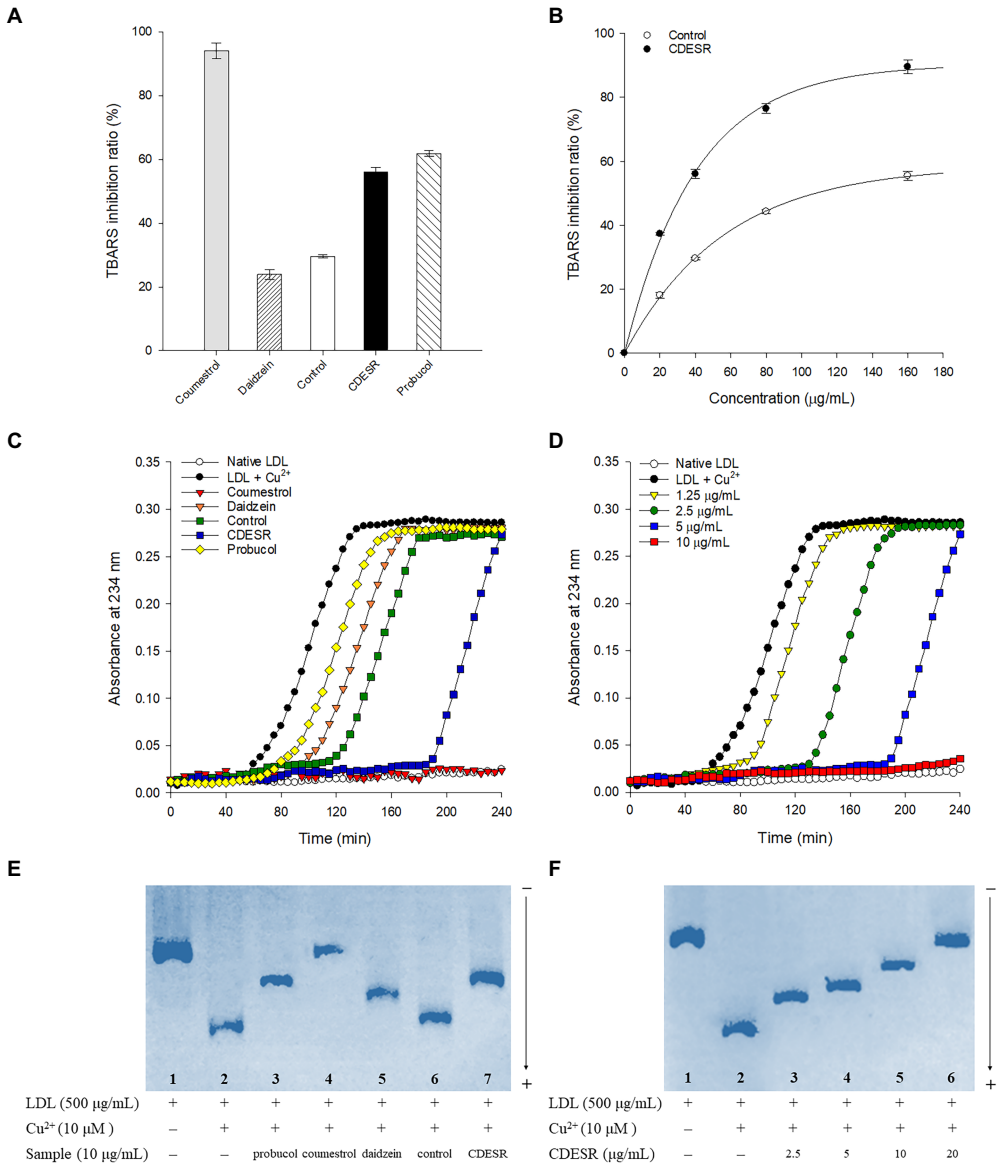


FIGURE 6
(A) Inhibitions of Cu^{2+}-induced lipid peroxidation by 40μg/ml of coumestrol, daidzein, control soybean roots extract, and CDESR (coumestrol and daidzein enriched soybean roots) extract in the TBARS assay. Probucol was used as a positive control. Data represent the mean±SE (n=3). **(B)** Dose-dependent effects of control roots and CDESR in the TBARS assay. Data represent the mean±SE (n=3). **(C)** Inhibitory effects (extension of lag time) of 5μg/ml of samples on conjugated diene generation in Cu^{2+}-induced LDL oxidation. **(D)** Dose-dependent effects of CDESR on conjugated diene generation. **(E)** Inhibitory effects on Cu^{2+} mediated LDL oxidation by relative electrophoretic mobility (REM). Lane 1, native LDL (absence of Cu^{2+}); lane 2, oxidized LDL; lane 3–7, 10μg/ml of probucol, coumestrol, daidzein, control root extract, and CDESR extract. **(F)** Dose-dependent effects of CDESR by REM assay. Lane 1, native LDL; lane 2, oxidized LDL; lane 3–6, CDESR (2.5μg/ml– 0μg/ml).

In a third assessment of the anti-LDL oxidation properties, REM experiments were carried out to examine additional parameters. As shown in Figure 6E, oxLDL moved to the bottom because of its negative charge. The lanes (1–7) in Figure 6E were designed as follows: lane 1 (native LDL), lane 2 (oxidized LDL), and lane 3–7 (samples at 10 μg/ml), and all samples were incubated for 16 h. The mobility of LDL was clearly reduced by coumestrol, daidzein, and probucol as in the previous study (Jin et al., 2006). The efficacy of the CDESR inhibited LDL-oxidation was 78.2% compared with that of the control roots of 17.3% at 10 μg/ml. Dose-dependent inhibitions of CDESR were observed in REM experiments of 4ss0.8% (2.5 μg/ml), 56.3% (5 μg/ml), 78.2% (10 μg/ml), and 95.8% (20 μg/ml) as shown in Figure 6F; Supplementary Figure S8. The improved antioxidant potential of CDESR could be attributed to the dramatic increase in the coumestrol (0.3–4.76 mg/g DW) and daidzein (1.16–8.87 mg/g DW) levels as a consequence of the application of SA.

Conclusion

In summary, the application of SA led to significant changes in the levels of secondary metabolites in soybean roots. Coumestrol and daidzein were revealed to be the metabolites of which the concentrations changed the most to reach levels as high as 4.76 mg/g DW and 8.87 mg/g DW, respectively. Metabolomic analyses were carried out by UPLC-ESI-Q-TOF/MS to reveal details of the changes in the metabolites by using the PLS-DA score, heatmap, and box plots. In particular, SA treatment played to stimulate a production of coumestrol as well as hydrolysis of its glycosides (coumestrin and malonylcoumestrin) in soybean roots. The SA treated soybean roots, CDESR, showed much improved anti-LDL oxidation effects than control roots based on the TBARS, CD formation, and REM experiments. This was rationalized by the distinctive increase in the coumestrol and daidzein levels in soybean roots as a result of SA application.

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

JK, YL, YB, and KP conceived and designed the experiments. AB and YB analyzed the data. JK wrote the manuscript. JK, AS, AB, YB, and KP were involved in the related discussion. AB and AS helped to improve the quality of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the Bio and Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science and ICT (No. NRF2020M3A9I3038523) and Rural Development Administration (PJ015732), Republic of Korea. The BK21 plus program supported scholarships for all students. The researchers received no external funding.

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.2022.1000705/full#supplementary-material

References

Anusuya, S., and Sathiyabama, M. (2016). Effect of chitosan on growth, yield and curcumin content in turmeric under field condition. *Biocatal. Agric. Biotechnol.* 6, 102–106. doi: 10.1016/j.bcab.2016.03.002

Baenas, N., García-Viguera, C., and Moreno, D. A. (2014). Elicitation: a tool for enriching the bioactive composition of foods. *Molecules* 19, 13541–13563. doi: 10.3390/molecules190913541

Bakker, P. A. H. M., Berendsen, R. L., Doornbos, R. F., Wintermans, P. C. A., and Pieterse, C. M. J. (2013). The rhizosphere revisited: root microbiomics. *Front. Plant Sci.* 4, 1–7. doi: 10.3389/fpls.2013.00165

Ban, Y. J., Song, Y. H., Kim, J. Y., Baiseitova, A., Lee, K. W., Kim, K. D., et al. (2020). Comparative investigation on metabolites changes in soybean leaves by ethylene and activation of collagen synthesis. *Ind. Crop. Prod.* 154:112743. doi: 10.1016/j.indcrop.2020.112743

Boué, S. M., Carter, C. H., Ehrlich, K. C., and Cleveland, T. E. (2000). Induction of the soybean phytoalexins coumestrol and glyceollin by aspergillus. *J. Agric. Food Chem.* 48, 2167–2172. doi: 10.1021/jf9912809

Brunswick, N. (1992). Salicylate, a new plant Hormone1. *Plant Physiol.* 99, 799–803. doi: 10.1104/pp.99.3.799

Chen, K., Thomas, S. R., and Keaney, J. F. (2003). Beyond LDL oxidation: ROS in vascular signal transduction. *Free Radic. Biol. Med.* 35, 117–132. doi: 10.1016/S0891-5849(03)00239-9

Ding, P., and Ding, Y. (2020). Stories of salicylic acid: a plant defense hormone. *Trends Plant Sci.* 25, 549–565. doi: 10.1016/j.tplants.2020.01.004

Durango, D., Murillo, J., Echeverri, F., Escobar, G., and Quiñones, W. (2018). Isoflavonoid composition and biological activity of extracts from soybean seedlings treated by different elicitors. *An. Acad. Bras. Cienc.* 90, 1955–1971. doi: 10.1590/0001-3765201820170785

Dwiecki, K., Neunert, G., Polewski, P., and Polewski, K. (2009). Antioxidant activity of daidzein, a natural antioxidant, and its spectroscopic properties in organic solvents and phosphatidylcholine liposomes. *J. Photochem. Photobiol. B Biol.* 96, 242–248. doi: 10.1016/j.jphotobiol.2009.06.012

Figueroa-Pérez, M. G., Gallegos-Corona, M. A., Ramos-Gomez, M., and Reynoso-Camacho, R. (2015). Salicylic acid elicitation during cultivation of the peppermint plant improves anti-diabetic effects of its infusions. *Food Funct.* 6, 1865–1874. doi: 10.1039/c5fo00160a

Franke, A. A., Custer, L. J., Cerna, C. M., and Narala, K. K. (1994). Quantitation of phytoestrogens in legumes by HPLC. *J. Agric. Food Chem.* 42, 1905–1913. doi: 10.1021/jf00045a015

Han, Q., Ma, Q., Chen, Y., Tian, B., Xu, L., Bai, Y., et al. (2020). Variation in rhizosphere microbial communities and its association with the symbiotic efficiency of rhizobia in soybean. *ISME J.* 14, 1915–1928. doi: 10.1038/s41396-020-0648-9

Hao, J., Wu, T., Li, H., Wang, W., and Liu, H. (2016). Dual effects of slightly acidic electrolyzed water (SAEW) treatment on the accumulation of γ-aminobutyric acid (GABA) and rutin in germinated buckwheat. *Food Chem.* 201, 87–93. doi: 10.1016/j.foodchem.2016.01.037

Jin, H. L., Byong, W. L., Jin, H. K., Jeong, T. S., Mim, J. K., Woo, S. L., et al. (2006). LDL-antioxidant pterocarpans from roots of *Glycine max* (L) Merr. *J. Agric. Food Chem.* 54, 2057–2063. doi: 10.1021/jf052431c

Kim, S. N., Ahn, S. Y., Song, H. D., Kwon, H. J., Saha, A., Son, Y., et al. (2020b). Antiobesity effects of coumestrol through expansion and activation of brown adipose tissue metabolism. *J. Nutr. Biochem.* 76:108300. doi: 10.1016/j.jnutbio.2019.108300

Kim, J. Y., Wang, Y., Li, Z. P., Baiseitova, A., Ban, Y. J., and Park, K. H. (2020a). Xanthine oxidase inhibition and anti-LDL oxidation by Prenylated Isoflavones from *Flemingia philippinensis* root. *Molecules* 25:3074. doi: 10.3390/molecules25133074

Lee, Y. S., Ju, H. K., Kim, Y. J., Lim, T. G., Uddin, M. R., Kim, Y. B., et al. (2013). Enhancement of anti-inflammatory activity of Aloe vera adventitious root extracts through the alteration of primary and secondary metabolites via salicylic acid elicitation. *PLoS One* 8:e82479. doi: 10.1371/journal.pone.0082479

Legette, L. L., Prasain, J., King, J., Arabshahi, A., Barnes, S., and Weaver, C. M. (2014). Pharmacokinetics of equol, a soy isoflavone metabolite, changes with the form of equol (dietary versus intestinal production) in ovariectomized rats. *J. Agric. Food Chem.* 62, 1294–1300. doi: 10.1021/jf400097m

Liu, H. K., Kang, Y. F., Zhao, X. Y., Liu, Y. P., Zhang, X. W., and Zhang, S. J. (2019). Effects of elicitation on bioactive compounds and biological activities of sprouts. *J. Funct. Foods* 53, 136–145. doi: 10.1016/j.jff.2018.12.019

Maruri-López, I., Aviles-Baltazar, N. Y., Buchala, A., and Serrano, M. (2019). Intra and extracellular journey of the phytohormone salicylic acid. *Front. Plant Sci.* 10, 1–11. doi: 10.3389/fpls.2019.00423

Morandi, D., Bailey, J. A., and Gianinazzi-Pearson, V. (1984). Isoflavonoid accumulation in soybean roots infected with vesicular-arbuscular mycorrhizal fungi. *Physiol. Plant Pathol.* 24, 357–364. doi: 10.1016/0048-4059(84)90009-2

Morant, A. V., Jørgensen, K., Jørgensen, C., Paquette, S. M., Sánchez-Pérez, R., Møller, B. L., et al. (2008). β-Glucosidases as detonators of plant chemical defense. *Phytochemistry* 69, 1795–1813. doi: 10.1016/j.phytochem.2008.03.006

Mun, B. G., Kim, H. H., Yuk, H. J., Hussain, A., Loake, G. J., and Yun, B. W. (2021). A potential role of Coumestrol in soybean leaf senescence and its interaction with Phytohormones. *Front. Plant Sci.* 12, 1–11. doi: 10.3389/fpls.2021.756308

Mundi, S., Massaro, M., Scoditti, E., Carluccio, M. A., Van Hinsbergh, V. W. M., Iruela-Arispe, M. L., et al. (2018). Endothelial permeability, LDL deposition, and cardiovascular risk factors-a review. *Cardiovasc. Res.* 114, 35–52. doi: 10.1093/cvr/cvx226

No, H. K., Lee, K. S., Kim, I. D., Park, M. J., Kim, S. D., and Meyers, S. P. (2003). Chitosan treatment affects yield, ascorbic acid content, and hardness of soybean sprouts. *J. Food Sci.* 68, 680–685. doi: 10.1111/j.1365-2621.2003.tb05731.x

Park, G., Baek, S., Kim, J. E., Lim, T. G., Lee, C. C., Yang, H., et al. (2015). Flt3 is a target of coumestrol in protecting against UVB-induced skin photoaging. *Biochem. Pharmacol.* 98, 473–483. doi: 10.1016/j.bcp.2015.08.104

Pawlowski, M. L., Vuong, T. D., Valliyodan, B., Nguyen, H. T., and Hartman, G. L. (2020). Whole-genome resequencing identifies quantitative trait loci associated with mycorrhizal colonization of soybean. *Theor. Appl. Genet.* 133, 409–417. doi: 10.1007/s00122-019-03471-5

Pérez-Tortosa, V., López-Orenes, A., Martínez-Pérez, A., Ferrer, M. A., and Calderón, A. A. (2012). Antioxidant activity and rosmarinic acid changes in salicylic acid-treated *Thymus membranaceus* shoots. *Food Chem.* 130, 362–369. doi: 10.1016/j.foodchem.2011.07.051

Pirillo, A. (2013). LOX-1, OxLDL, and atherosclerosis. *Mediat. Inflamm.* 2013, 1–12. doi: 10.1155/2013/152786

Prakash, D., Verma, S., Bhatia, R., and Tiwary, B. N. (2011). Risks and precautions of genetically modified organisms. *ISRN Ecol.* 2011, 1–13. doi: 10.5402/2011/369573

Schreiner, M., Krumbein, A., Knorr, D., and Smetanska, I. (2011). Enhanced glucosinolates in root exudates of brassica rapa ssp. rapa mediated by salicylic acid and methyl jasmonate. *J. Agric. Food Chem.* 59, 1400–1405. doi: 10.1021/jf103585s

Sugiyama, A. (2019). The soybean rhizosphere: metabolites, microbes, and beyond—a review. *J. Adv. Res.* 19, 67–73. doi: 10.1016/j.jare.2019.03.005

Vitale, D. C., Piazza, C., Melilli, B., Drago, F., and Salomone, S. (2013). Isoflavones: estrogenic activity, biological effect and bioavailability. *Eur. J. Drug Metab. Pharmacokinet.* 38, 15–25. doi: 10.1007/s13318-012-0112-y

Wang, M., Ding, Y., Wang, Q., Wang, P., Han, Y., Gu, Z., et al. (2020). NaCl treatment on physio-biochemical metabolism and phenolics accumulation in barley seedlings. *Food Chem.* 331:127282. doi: 10.1016/j.foodchem.2020.127282

Wang, Q., Ge, X., Tian, X., Zhang, Y., Zhang, J., and Zhang, P. (2013). Soy isoflavone: the multipurpose phytochemical (review). *Biomed. Reports* 1, 697–701. doi: 10.3892/br.2013.129

Wen Xu, Y. (2012). Effects of salicylic acid on monoterpene production and antioxidant systems in *Houttuynia cordata*. *African J. Biotechnol.* 11, 1364–1372. doi: 10.5897/ajb11.1524

Xu, B. J., Yuan, S. H., and Chang, S. K. C. (2007). Comparative studies on the antioxidant activities of nine common food legumes against copper-induced human low-density lipoprotein oxidation in vitro. *J. Food Sci.* 72, S522–S527. doi: 10.1111/j.1750-3841.2007.00464.x

Xu, A., Zhan, J. C., and Huang, W. D. (2015). Effects of ultraviolet C, methyl jasmonate and salicylic acid, alone or in combination, on stilbene biosynthesis in cell suspension cultures of *Vitis vinifera* L. cv. Cabernet sauvignon. *Plant Cell Tissue Organ Cult.* 122, 197–211. doi: 10.1007/s11240-015-0761-z

Yamamoto, R., Ma, G., Zhang, L., Hirai, M., Yahata, M., Yamawaki, K., et al. (2020). Effects of salicylic acid and methyl jasmonate treatments on flavonoid and carotenoid accumulation in the juice sacs of Satsuma mandarin in vitro. *Appl. Sci.* 10, 1–13. doi: 10.3390/app10248916

Yuk, H. J., Curtis-Long, M. J., Ryu, H. W., Jang, K. C., Seo, W. D., Kim, J. Y., et al. (2011). Pterocarpan profiles for soybean leaves at different growth stages and investigation of their glycosidase inhibitions. *J. Agric. Food Chem.* 59, 12683–12690. doi: 10.1021/jf203326c

Yuk, H. J., Lee, J. W., Park, H. A., Kwon, O. K., Seo, K. H., Ahn, K. S., et al. (2017). Protective effects of coumestrol on lipopolysaccharide-induced acute lung injury via the inhibition of proinflammatory mediators and NF-κB activation. *J. Funct. Foods* 34, 181–188. doi: 10.1016/j.jff.2017.04.027

Zafar, A., Singh, S., and Naseem, I. (2017). Cytotoxic activity of soy phytoestrogen coumestrol against human breast cancer MCF-7 cells: insights into the molecular mechanism. *Food Chem. Toxicol.* 99, 149–161. doi: 10.1016/j.fct.2016.11.034

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 31 October 2022
DOI 10.3389/fpls.2022.1020309

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Marcus Scotti,
Federal University of Paraíba, Brazil
Wei Wang,
Hunan University of Chinese Medicine, China
Yan-Hui Fu,
Hainan Normal University, China

*CORRESPONDENCE
Lina Xia
xialina@cdutcm.edu.cn
Yuntong Ma
Mayuntong@cdutcm.edu.cn

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

RECEIVED 16 August 2022
ACCEPTED 06 October 2022
PUBLISHED 31 October 2022

CITATION
Qi L, Zhong F, Liu N, Wang J, Nie K, Tan Y, Ma Y and Xia L (2022) Characterization of the anti-AChE potential and alkaloids in Rhizoma Coptidis from different *Coptis* species combined with spectrum-effect relationship and molecular docking. *Front. Plant Sci.* 13:1020309. doi: 10.3389/fpls.2022.1020309

Characterization of the anti-AChE potential and alkaloids in Rhizoma Coptidis from different *Coptis* species combined with spectrum-effect relationship and molecular docking

Luming Qi[1,2], Furong Zhong[3], Nannan Liu[1,2], Jie Wang[1,2], Kaidi Nie[1,2], Youli Tan[4], Yuntong Ma[3]* and Lina Xia[1,2]*

[1]School of Health Preservation and Rehabilitation, Chengdu University of Traditional Chinese Medicine, Chengdu, China, [2]State Administration of Traditional Chinese Medicine Key Laboratory of Traditional Chinese Medicine Regimen and Health, Chengdu University of Traditional Chinese Medicine, Chengdu, China, [3]School of Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu, China, [4]Department of Pharmacy, Affiliated Sport Hospital of CDSU, Chengdu Sport University, Chengdu, China

Coptis species are the main source of Rhizoma Coptidis (RC) drugs, which have always been used to treat Alzheimer's disease in the clinical experience of ancient China. However, many species of this genus have been largely underutilized until now. With this fact, this research has been designed to investigate for the first time the anti-acetylcholinesterase (AChE) property of different extracts for RC drugs from four *Coptis* species (*C. chinensis*, *C. deltoidea*, *C. teeta* and *C. omeiensis*) and to quantify the main alkaloids. Petroleum ether, ethyl acetate and n-butanol fractions of RC drugs were sequentially collected using an accelerated solvent extraction technique. Spectrum-effect relationship and molecular docking were applied to analyse the relationships between alkaloids and AChE inhibitory activity. The N-butanol extract was proven to be the main active fraction, and *C. teeta* may be the best source of RC drugs for Alzheimer's disease treatment, with significantly lower IC 20, IC 50 and IC 80 values for AChE inhibition. The UPLC/QqQ-MS quantitative analysis showed that the accumulations of 10 alkaloids in RC drugs from different sources greatly varied. Three data processing methods (Random forest, Boruta and Pearson correlation) comprehensively analysed the spectrum-effect relationship and revealed that columbamine, berberine and palmatine were the most important AChE inhibitors that could be used as quality markers to select RC drugs for Alzheimer's disease treatment. In addition, the dominant compounds were successfully docked against AChE to verify the binding affinity and interactions with the active site. The present study can contribute to the reasonable development and utilization of RC drugs from different sources, especially to provide certain evidence for their application in the treatment of Alzheimer's disease.

KEYWORDS

Rhizoma Coptidis, *Coptis* species, alkaloids, AChE inhibition, Spectrum-effect relationship, Molecular docking

Introduction

Alzheimer's disease is the most common age-related disease with chronic memory and cognitive decline as the main manifestations, followed by psychiatric symptoms, behavioural disorders and impairment of activities in daily lives (Ju and Tam, 2022). With the increasing ageing of the world, the incidence rate of this disease continues to increase, which has attracted extensive attention from scientific researchers. Acetylcholinesterase (AChE) inhibitors, including tacrine hydrochloride, huperzine A, and donepezil, are currently used drugs to relieve the symptoms of this disease (Taqui et al., 2022). Although these drugs can delay disease progression, most of them are accompanied by certain side effects (Refaay et al., 2022) Therefore, the screening of AChE inhibitors from natural products has become a research hotspot to better treat Alzheimer's disease (Lin et al., 2020b; Ahmed et al., 2021).

The root and rhizome material of the *Coptis* species Rhizoma Coptidis (RC) is a top-grade drug that has been documented to have an excellent effect on improving memory and treating brain diseases in ancient clinical experience in China. Modern pharmacological studies have proven that this drug has a positive therapeutic potential on Alzheimer's disease, and several alkaloids have been identified as the main AChE inhibitors (Zhao et al., 2016; Cao et al., 2018; Lin et al., 2020a). However, these studies mainly focused on the extract of RCdrugs from *C. chinensis* Franch, and other species of the *Coptis* genus have never been investigated regarding their anti-AChE properties. In particular, some studies have demonstrated that RC drugs separated from different *Coptis* plants always have different therapeutic effects. For example, Feng et al. (Feng et al., 2011) compared the antibacterial effect of different RC drugs and demonstrated that drugs from *C. chinensis* had the highest inhibitory effect against *Staphylococcus aureus*. Our research team has proven that RC drugs from *C. deltoidea* have a better hypoglycaemic effect than those from the other three *Coptis* species (Chen et al., 2021). In summary, the anti-AChE potential of RC drugs from different *Coptis* species has never been investigated, and the relationships between this property and alkaloids are unclear.

Because natural medicinal materials are subject to a series of metabolic processes that produce multipolar chemical profiles, an efficient extraction method is a key step for their chemical and pharmacological analysis (Vezzulli et al., 2022). Accelerated solvent extraction techniques with the advantages of easy operation, high speed and efficiency are more effective for pre-treating these complex biological systems. Based on an efficient preparation procedure, the activity assay and chemical analysis can be more accurately achieved. In addition, spectrum-effect relationship analysis and molecular docking are popular methods to analyse the association between chemical compounds and the biological activity of natural products (Shi et al., 2016; da Silva Barbosa et al., 2020; Chang et al., 2021; Lu et al., 2022; Amir Rawa et al., 2022). Spectrum-effect relationship analysis refers to constructing the relationship of the chemical profiles with specific biological activity and this association can be used to reflect the internal quality of natural drugs (Zhang et al., 2018). Molecular docking is a commonly used technique based on structural molecular biology and computer-assisted drug design that can effectively predict the binding mechanism of a ligand with a protein based on the three-dimensional structure (Mtemeli et al., 2022). These techniques have been proven effective for the in-depth analysis of the association mechanism between biological activity and chemical compositions regarding natural drugs.

Thus, the present study was designed to investigate the anti-AChE potential of n-butanol, ethyl acetate and petroleum ether extracts of RC drugs from four *Coptis* species (*C. chinensis*, *C. deltoidea*, *C. omeiensis* and xtitC. teeta) based on an *in vitro* activity assay and to quantify the main alkaloids in the active fraction using ultra-performance liquid chromatography with triple quadrupole mass spectrometry (UPLC/QqQ-MS). The accelerated solvent extraction technique was used to obtain different extraction fractions. Furthermore, spectrum-effect relationships and molecular docking techniques were jointly applied in an auxiliary manner to analyse the internal association between alkaloid profiles and anti-AChE activity. We hope that our conclusion can provide an effective foundation for the rational application of RC drugs from different *Coptis* species, especially for the treatment of Alzheimer's disease.

Materials and methods

RC drugs materials

Four *Coptis* species *C. chinensis*, *C. deltoidea*, *C. omeiensis* and *C. teeta* were collected in their harvest time (n=6). The sampling site of first three plants is Heishan Village, Hongya County, Sichuan Province of China (Longitude: 103.1611, Latitude: 29.5097), and the sampling site of *C. teeta* is Pihe Township, Fugong County, Yunnan Province of China (Longitude: 98.9236, Latitude: 26.5375). All materials were collected from 5-year-old plants which were identified by Pro. Ma of Chengdu University of Traditional Chinese Medicine.

Roots and rhizomes of these plants were firstly separated and washed by deionized water. Then, these materials were dried in an oven at a 60°C condition and smashed using a pulverizer. The obtained powder were filtered through a 100-mesh sieve. Finally, these RC drugs were labelled and stored in a cool condition for the subsequent analysis. The images of these materials are presented in Figure 1.

FIGURE 1
The images of RC materials. **(A)** *C. chinensis* species and its root and rhizome; **(B)** *C. deltoidea* species and its root and rhizome; **(C)** *C. omeiensis* species and its root and rhizome; **(D)** *C. teeta* species and its root and rhizome.).

Reagents

Chemical standards of magnoflorine, groenlandicine, demethyleneberberine, columbamine, epiberberine, coptisine, jatrorrhizine, berberrubine, palmatine and berberine were purchased from Chroma-Biotechnology Co., Ltd. (Chengdu, China). Their purities are higher than 98%. AChE and phosphate buffered saline were purchased from Sigma-Aldrich Shanghai Trading Co. Ltd. (Shanghai, China). 5,5'-Dithiobis-(2-nitrobenzoic acid) (DTNB), S-Acetylthiocholine iodide (ATCHI) and sodium dodecyl sulfate (SDS) were purchased from Adamas Reagent Co., Ltd. (Shanghai, China). Methanol, acetonitrile and formic acid of chromatographic grade were obtained from Thermo Fisher Scientific (Shanghai, China). Deionized water was produced by an ultrapure water system (Millipore, USA). Other solvents such as n-butanol, ethyl acetate petroleum ether and dimethylsulfoxide (DMSO) were analytical grade and came from Kelong Chemical Co., Ltd. (Chengdu, Shanghai).

Accelerated solvent extraction

Accelerated solvent extraction was achieved by a Speed Extractor E-916 instrument (BÜCHI, Switzerland) according to the developed method (Chen et al., 2021). Briefly, 1 g per sample of RC drugs was weighed and placed in a 40-mL extraction cell mixed with approximately 50 g quartz sand. A series of extracted fractions of petroleum ether, ethyl acetate and n-butanol (water-saturated) with different polarities sequentially set, with temperatures and pressures of 90°C and 100 bar, respectively. Finally, the collected extracts were evaporated by a rotary evaporator. The fractions of petroleum ether and ethyl acetate were redissolved in 10% dimethylsulfoxide at a concentration of 0.2 g/mL, while the fractions of water-saturated n-butanol were redissolved in 50% methanol at a concentration of 0.2 g/mL. Related parameters of this technique were exhibited in Table S1.

Evaluation of AChE inhibitory activity

An *in vitro* AChE inhibition test was conducted based on a previous report with slight modification (Kaufmann et al., 2016). The experiment was performed in a 96-well microplate and mainly included three parts: 1) Sequentially adding AChE (0.1 U/mL) and RC extracts into phosphate buffer; 2) Combining DTNB (2.5 mmol/mL) and ATCHI (10 mmol/mL) for the reaction at 37°C for 10 min; 3) Prohibiting the reaction at 37°C for 10 min using SDS (1%) solutions. After the reaction, a SpectraMax iD3 microplate reader (Molecular Devices, USA) was applied to measure the absorbance at 405 nm. The inhibition (%) ratio was calculated by the following formula: Inhibition (%) = $[1-(AB_a-AB_b)/(AB_c-AB_d)]\times100\%$ (where AB_a is the absorbance of the mixture with sample and AChE; AB_b is the absorbance of the mixture only with sample solution; AB_c is the absorbance of the mixture only with AChE; AB_d is the absorbance of the mixture without both sample solution and AChE). The same volume of phosphate buffer was used to make up the missing solutions. All samples were analysed in triplicate. The IC 20, IC 50 and IC 80 values were calculated using a logistic regression method to display the anti-AChE activity of different extracts. Prior to the sample determination, the AChE inhibition activity of the blank reconstitution solvent was determined to ensure the accuracy of the final results. Detailed operation process was exhibited in Table S2.

UPLC/QqQ-MS analysis

A UPLC/QqQ-MS system with a Waters ACQUITY UPLC H-Class connected online to a Waters Xevo Triple quadrupole (TQD) mass spectrometer (Waters, Milford, MA, USA) was applied to determine the alkaloid profiles of RC extracts according to our developed method (Zhong et al., 2020). The extracts were chromatographically separated by an ACQUITY UPLC BEH C18 (100 × 2.1 mm, 5 m) with a column temperature of 25°C and a flow rate of 0.4 mL/min. The mobile phases were 0.1% formic acid water (A) and acetonitrile (B). The following gradient program was set: 0-2 min, 85%-76% A; 2-6 min, 76%-75.5% A; 6-8 min, 75.5%-75.4% A; and 8-10 min, 75.4%-75% A. The injection volume and detection wavelength were set as 1 L and 320 nm, respectively.

The Xevo TQD mass spectrometer was conducted in positive ion mode. High purity nitrogen and helium was applied as nebulizing gas and collision gas, respectively. The mass spectrometry conditions are listed below: capillary voltage: 2.5 KV; cone voltage: 25 V; source temp: 120°C; desolvation temp: 500°C; desolvation gas: 1000 L/hr; cone gas: 50 L/hr; Full scan range: 100-1200 amu; scan mode: MSe. The RC extracts were determined based on the multiple reaction monitoring (MRM) mode and the cone voltages and collision energies were optimized according to the alkaloid reference standards.

Spectrum-effect relationship analysis

Spectrum-effect relationship analysis was applied to interpret the role of the main alkaloids, which explained the anti-AChE bioactivity of RC extracts by defining the AChE inhibition rate as the dependent variable "*Y*" and defining the alkaloid profiles as the independent variable "*X*". Because data processing methods have an important influence on spectrum-effect relationship results, we combined the random forest (RF), Boruta and Pearson correlation based on different mathematical principles to obtain a more accurate result.

RF is an ensemble machine learning method that addresses relationships between abundant "*X*" variables and "*Y*" response in a high-dimensional space (Speiser et al., 2019). For this algorithm, a set number of trees (n_{tree}) that are independent of one another is selected as an individual classifier. Next, a random subspace of variables (m_{try}) regarding each individual classifier was defined to minimize the model error. An increase in mean squared error (*IncMSE*) based on the permutation importance represents the importance of each variable in the RF model. By randomly assigning a value to each variable, the error of the RF model will increase if this variable is important (Kursa and Rudnicki, 2010). Based on the RF model, Boruta (Altmann et al., 2010) introduces the shadow attribute to evaluate the importance of each "*X*" variable to explain the dependent variable "*Y*". The shadow attributes are used as a reference to calculate the Z score to decide which variables are truly important. Pearson correlation (Yang et al., 2021) is a method to measure the linear correlation of two variables. The correlation coefficient is between 0 and 1 and represents the relevance between two variables. A numerical value of 1 means that the two variables are positively correlated, while a coefficient value of 0 means that the two variables are negatively correlated.

Finally, the *IncMSE*, Z score and Pearson coefficient were output to interpret the importance of each alkaloid responsible for the anti-AChE activity based on the developed spectrum-effect relationship model. We performed fuzzy aggregation connective operators (minimum, maximum, average and product) to make the fusion for the outputs from three different mathematical algorithms (Obisesan et al., 2017). For the three highest ranked attributes, the multiple regression technique was used to construct the linear equation to directly exhibit the relationship between alkaloids and AChE inhibition activity.

Simulation of molecular docking

Molecular docking aims to search for active small molecules with binding potency to protein receptors based on the complementary laws of geometry, energy and chemical environment (Mtemeli et al., 2022). According to the minimum binding free energy principle, the scoring function ranks the binding ability between ligands and receptors, which represents the bioactivity of small molecules. This high-throughput technique performs at a virtual level without wasting solvent and monomer components and has been extensively applied to screen bioactive molecules from herbal medicines.

The crystal structure of AChE in complex with donepezil (PDB ID: 6O4W, resolution: 2.35 Å) (Gerlits et al., 2019) was downloaded from the RCSB PDB database. The three-dimensional structures of alkaloids were obtained from the SciFinder database. AutoDock Vina programs (Eberhardt et al., 2021) were applied to perform the molecular docking model. For the molecular docking process, we selected co-crystallized ligands of donepezil for AChE as the reference to ensure the accuracy of the active site. The molecular docking process is briefly described as follows: 1) optimizing the receptor file by deleting water molecules, adding hydrogen and charges, merging nonpolar hydrogen atoms, calculating atomic local charges, etc., and saving it in the pdbqt format; 2) optimizing the ligand files by adding hydrogen and charge and saving it as a pdbqt format file; 3) setting the number of grid points and coordinates of the centre point of the docking and saving it as a GPF file; 4) performing the molecular docking and analysing the interaction between receptor and ligands. The related tools and functions are detailed in Table S3.

Results

Anti-AChE potency of different extracts for four RC drugs

Natural AChE inhibitors have recently attracted more interest for the treatment of the symptoms of Alzheimer's disease. Three petroleum ether, ethyl acetate and n-butanol fractions of RC drugs from four *Coptis* species of *C. chinensis*, *C. deltoidea*, *C. omeiensis* and *C. teeta* were initially obtained using an accelerated solvent extraction technique. Inhibitory assays were conducted *in vitro* to evaluate the anti-AChE properties of different extracts. The results are shown in Figure 2. The petroleum ether and ethyl acetate fractions show only weak inhibition with IC 50 values of 76.25-200.59 mg/mL and 20.36-158.26 mg/mL, respectively (Figures 2A, B).

The n-butanol fraction exhibits the best AChE inhibition activity, which is more than a thousand times stronger than the petroleum ether and ethyl acetate fractions (Figure 2C). Previous studies have suggested that this fraction mainly contains isoquinoline alkaloids that contribute to a better AChE inhibition activity (Zhao et al., 2016; Cao et al., 2018; Lin et al., 2020a). Additionally, the AChE inhibition activity of RC drugs from different *Coptis* species greatly varies. The RC drug from *C. teeta* exhibits the best inhibition activity with a significantly lower IC 50 (22.88 ± 3.05 g/mL) than *C. chinensis* (28.64 ± 2.79 g/mL, P<0.01) and a significantly lower IC 80 (85.95 ± 11.56 g/mL) than *C. omeiensis* (113.86 ± 13.69 g/mL, P <0.05). The RC drugs from *C. chinensis* and *C. omeiensis* exhibit considerable inhibitory activity, with not significantly different IC 50 (28.64 ± 2.79 g/mL vs. 24.91 ± 2.72 g/mL) and IC 80 (98.90 ± 4.40 g/mL vs. 113.86 ± 13.69 g/mL). However, IC 20 between these RC drugs presents a significant variation (7.89 ± 0.98 g/mL vs. 5.35 ± 0.42 g/mL, P<0.05). Comparatively, the RC drugs from *C. deltoidea* show the weakest inhibitory activity, with an extremely significant variation compared to the drugs from the other three *Coptis* plants (P<0.001).

We also determined the anti-AChE activity of huperzine A, which is used as a positive control drug to treat Alzheimer's disease. Its IC 20, IC 50 and IC 80 are 22.63 ± 1.52 g/mL, 74.05 ± 15.26 g/mL and 177.25 ± 8.47 g/mL, respectively. The AChE inhibition activity of huperzine A is comparable to the N-butanol fraction of RC drugs from *C. deltoidea*. The N-butanol fraction of RC drugs from the other three *Coptis* species manifests a much stronger anti-AChE potential than huperzine A according to our *in vitro* inhibition assay.

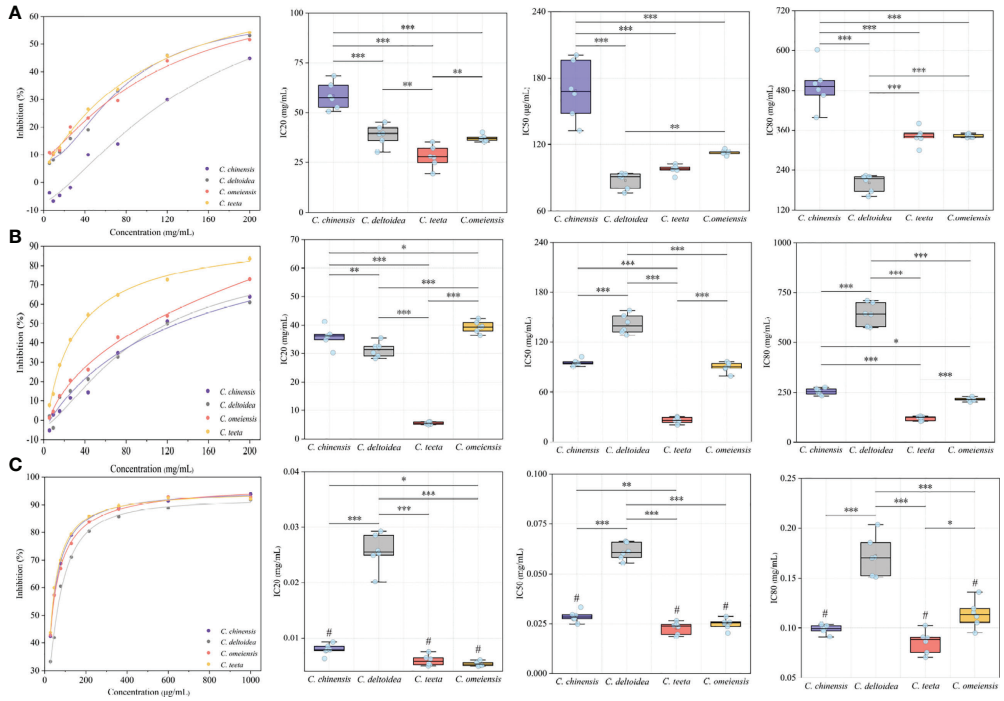


FIGURE 2
The results of *in vitro* inhibitory assays. This figure presents the AChE inhibitory curve and their IC 20, IC 50 and IC 80 of petroleum ether, ethyl acetate and n-butanol fractions of RC drugs for four *Coptis* species. The inhibitory curves were constructed by diluting the test solution in a 60% gradient. The letters of **(A–C)** represent the inhibitory of petroleum ether, ethyl acetate and n-butanol fractions of different RC drugs, respectively. The mark of "#" means the AChE inhibition of this extract is stronger than that of positive drug. The meaning of *, **, *** symbol is P<0.05, P<0.01, P<0.001, respectively.

Quantitative analysis of the main alkaloids in n-butanol extracts

The n-butanol fraction of RC drugs exhibits the best anti-AChE property. In our previous research, we qualitatively identified 15 alkaloids, which accounted for most chemical constituents in RC drugs (Chen et al., 2021). Herein, we focused on the precise quantification of 10 common alkaloids in the n-butanol fraction based on the MRM mode of UPLC/QqQ-MS technique. For each compound, a calibration curve was constructed by plotting its peak area against the standard concentration (Table 1). Other methodological parameters are also optimized using reference standards and exhibited in Tables S4, S5. These results indicate that the developed UPLC/QqQ-MS method can be used to simultaneously and precisely determine 10 alkaloids in the n-butanol fraction of RC drugs.

The UPLC chromatogram is presented in Figure S1. The MRM profiles of 10 alkaloids are shown in Figure 3. Combining their retention time and ion fragment characteristics, all target alkaloids are excellently separated and detected in the n-butanol fraction of RC drugs from different *Coptis* species. Based on the calibration curves, their concentrations were calculated. The hierarchical clustering algorithm holistically presents an excellent separation among four RC drugs from different *Coptis* species, and this result demonstrates that their concentrations of 10 alkaloids are apparently different (Figure 4A). Biplot simultaneously displays the relationships among scores and loadings of classification models of different RC drugs (Figure 4B). Coptisine and berberrubine can be used as identification markers for RC drugs from *C. teeta*, and groenlandicine can be used as an identification marker for RC drugs from *C. deltoidea*. Regarding RC drugs from *C. chinensis*, palmatine and columbamine are more valuable indicators to discriminate these drugs from others.

Figure 4C specifically shows the distribution of 10 alkaloids in RC drugs from different *Coptis* species. Berberine (48.27-79.26 mg/g) is identified as the most accumulated compound in RC drugs, followed by coptisine (17.26-24.47 mg/g), palmatine (5.26-19.98 mg/g), jatrorrhizine (4.25-13.76 mg/g) and epiberberine (0.61-14.61 mg/g). These conclusions are consistent with the previous papers (Zhong et al., 2018; Qi et al., 2018). In addition, RC drugs from *C. teeta* accumulate the highest berberine (73.95 ± 3.96 mg/g), followed by drugs originating from *C. omeiensis* (68.03 ± 2.82 mg/g) and *C. chinensis* (62.69 ± 3.63 mg/g). This compound in RC drugs from *C. deltoidea* (52.75 ± 2.78 mg/g) is extremely low compared with other medicinal materials ($P<0.001$). RC drugs from *C. teeta* also synthesized the highest *coptisine*, which was significantly higher than other RC drugs ($P<0.05$). This conclusion is inconsistent with previous research that RC drugs from *C. chinensis* have the highest coptisine (Chen et al., 2017). This variation may be caused by the difference in extraction techniques. The trace compounds demethyleneberberine (0.22 ± 0.02 mg/g, $P<0.001$) and berberrubine (0.02 ± 0.002 mg/g, $P<0.001$) also accumulated the most in RC drugs from *C. teeta*. RC drugs from *C. chinensis* accumulated the highest palmatine (16.36 ± 1.88 mg/g, $P<0.001$), columbamine (3.65 ± 0.57 mg/g, $P<0.05$) and epiberberine (12.98 ± 1.28 mg/g, $P<0.001$) levels. In addition, magnoflorine (7.94 ± 0.46 mg/g, $P<0.001$), groenlandicine (10.95 ± 1.87 mg/g, $P<0.001$) and jatrorrhizine (11.17 ± 1.83 mg/g) were the most abundant in RC drugs from *C. deltoidea*.

Differences in chemical compound concentrations of herbal medicines cause their different therapeutic efficacies. The alkaloid variation in the four RC extracts is the main reason that explains their different AChE inhibitory activities. Furthermore, the relationship between alkaloid profiles and anti-AChE activity was investigated to reveal the detailed mechanism underlying the efficacy difference.

Spectrum-effect analysis between anti-AChE properties and alkaloid profiles

To obtain a convincing result, we used three data processing methods with different mathematical principles to exhibit the

TABLE 1 The methodological parameters of the UPLC–MS/MS method.

Alkaloid Compounds	Standard Curves	Linear Interval (g/mL)	Correction Coefficient	Limit of Detection (g/mL)	Limit of Quantitation (g/mL)
Magnoflorine	y=4772370.39x+190.05	2.28-39.48	0.9998	0.24	0.81
Groenlandicine	y=13875850.89x-5361.91	0.44-45.64	0.9996	0.36	1.21
Demethyleneberberine	y=134061x+39.415	0.068-3.19	0.9996	0.04	0.09
Columbamine	y=31008479.81x-1094.99	1.69-42.00	0.9999	0.37	1.22
Epiberberine	y=19062767.74x+2286.64	1.21-86.80	0.9999	1.00	3.32
Coptisine	y=19040,809.50x-273.95	3.99-98.82	0.9999	0.87	1.53
Jatrorrhizine	y=14793030.39x+696.38	2.07-51.2	0.9999	0.48	1.61
Berberrubine	y=31230x-38.964	1.39-40.35	0.9991	0.11	0.34
Palmatine	y=19209155.01x+6732.86	3.13-77.60	0.9998	0.97	3.22
Berberine	y=18878818.98x-8645.61	38.71-470.00	0.9998	6.93	23.9

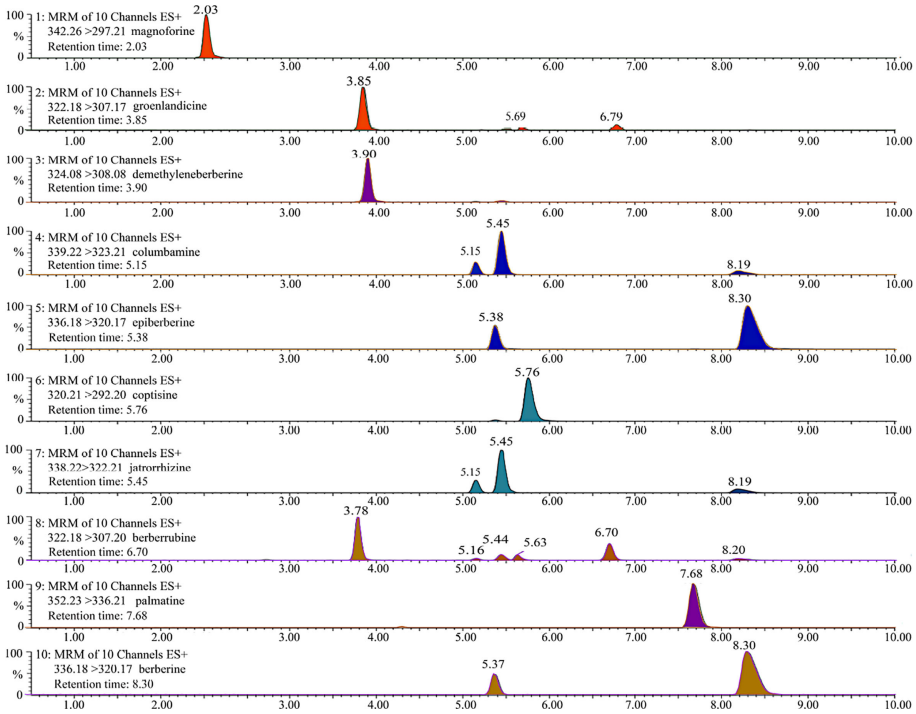


FIGURE 3
The representative MRM profiles of 10 alkaloids in RC drugs based on UPLC/QqQ-MS technique. These compounds were detected in positive ion mode, and all target alkaloids are excellently separated according to their retention time and ion fragment characteristics.

spectrum-effect relationship. The corresponding results are exhibited in Figure 5. A reliable RF model (*MSE*=0.00525) that explains 92.42% of the variables for the alkaloid matrix was first established with n_{tree} and m_{ttry} values of 500 and 3, respectively. The *IncMSE* score indicates that columbamine was most highly correlated with the AChE inhibitory activity, followed by berberine and palmatine (Figure 5A). Boruta was applied by setting shadow attributes, and the Z scores was used to represent the importance of these alkaloids. The *Z* scores of all alkaloids are higher than the shadow attributes, which indicates that they are highly correlated with AChE inhibitory activity. Comparatively, columbamine and berberine are the most important variables responsible for the anti-AChE activity (Figure 5B). The result from this algorithm is close to that from RF.

Furthermore, we performed a Pearson correlation analysis to exhibit the relational degree between AChE inhibition and 10 alkaloids and their own associations (Figure 5C). The results show that berberine and coptisine are the first two key compounds responsible for the anti-AChE activity, followed by palatine and columbamine. The results from this algorithm were obviously different from those of RF and Boruta because their internal mathematical operation methods were different. In addition, columbamine and palatine have the highest correlation (relational degree=0.99) because the former is the precursor material for the biosynthesis of the latter (Liu et al., 2022). Modern research has proved that some key enzymes required for the berberineand coptisine biosynthesis are identical, which can explain their high correlation with accumulations (relational degree=0.97) (Liu et al., 2021).

To take these correlation scores together, we made a data fusion for the minimum, maximum, average and product fuzzy aggregation connective operators based on different outputs from these data processing methods (Obisesan et al., 2017). The voting results indicate that the top three high-contributing alkaloids for AChE inhibitory activity are columbamine, berberine and palmatine (Figure 5D).Lin et al., (2020a) performed monomer experiments and indicated the relatively higher anti-AChE potential of berberine and columbamine. We also determined the AChE inhibition of eight alkaloid monomers. The results indicate that the inhibitory activities of these compounds were as follows: palmatine (1.49 ± 0.4 g/mL)> columbamine (2.29 ± 0.7 g/mL)> berberine (3.22 ± 0.9 g/mL)> coptisine (4.34 ± 1.5 g/mL)> groenlandicine (5.84 ± 2.0 g/mL)> jatrorrhizine (6.36 ± 2.2 g/mL)> epiberberine (10.14 ± 3.4 g/mL) > magnoflorine (21.26 ± 4.5 g/mL) (Figure S2). The previous published papers also showed that the AChE inhibition of

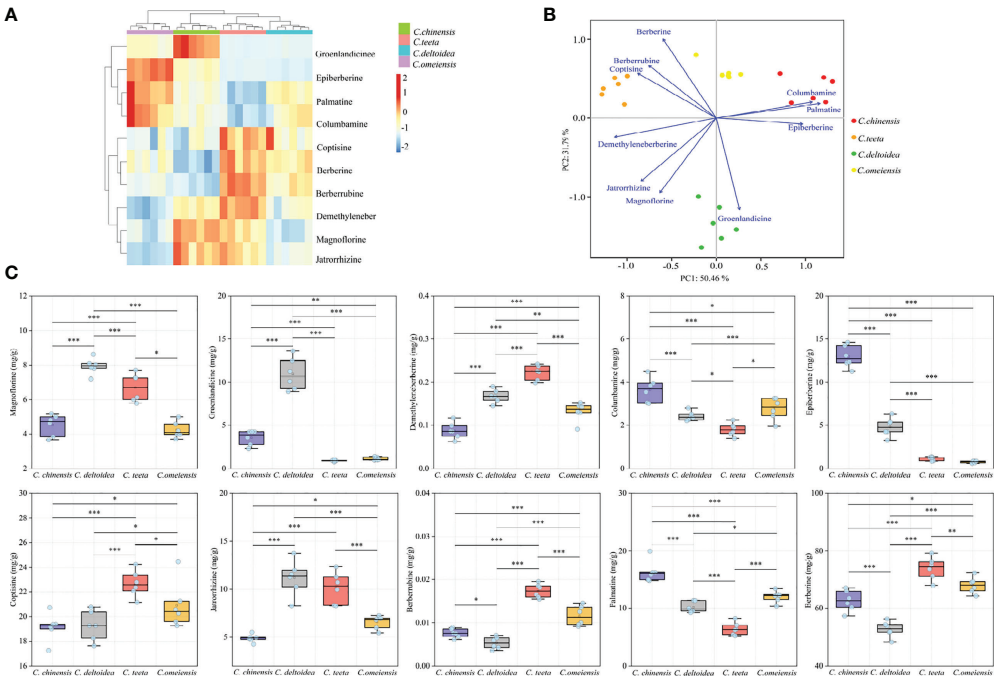


FIGURE 4
The quantitative results of 10 alkaloids in RC drugs. **(A)** hierarchical clustering presents the variation of 10 alkaloids in different RC drugs; **(B)** biplot displays the relationships among scores and loadings of classification model; **(C)** the specific content of each alkaloid. The meaning of *, **, *** symbol is P<0.05, P<0.01, P<0.001, respectively.

demethyleneberberine and berberrubine are much weaker than berberine (Huang et al., 2010; Roselli et al., 2016). These results proved that the spectrum-effect analysis was effective in interpreting the relationship between AChE inhibition activity and alkaloids. All tested alkaloids have much lower IC 50 than huperzine A, which indicates the huge anti-AChE potential of these alkaloids. Finally, the multiple regression technique presented the linear equation of AChE inhibition (YAChE) with berberine (X_b), columbamine (X_c) and palmatine (X_p). After removing insignificant coefficients, the formula is exported as YAChE $=0.60521X_b+3.3591X_c-1.9141X_p+36.878X_c^2$ $-72.817X_cX_p+34.125\ X_p^2+0.1297$, $R^2=0.94$. All model parameters are significant, revealing that the fitting effect of this equation is excellent (Tables S6, S7).

Molecular docking

To obtain better insight into the properties of columbamine, berberine and palmatine against AChE, molecular docking was performed to analyse their ligand-active site interactions. Searching the effective active site in the receptor for drug candidates with a low binding energy is a critical step for molecular docking (Feinstein and Brylinski, 2015). The centre coordinates (*X*: 94.58; *Y*: 94.72; *Z*: 17.01) of the docking box for AChE protein were optimized according to the location of the co-crystallized ligand of donepezil (Brewster et al., 2018). The three-dimensional structures of these alkaloids after energy minimization are exhibited in Fig S3. AutoDock Vina software was used to perform molecular docking, and the first docking poses were output according to the rank of scoring function.

Table 2 exhibits binding affinities, including hydrogen bonds, hydrophobic interactions and pi-pi interactions, which ensure stable interactions between alkaloids and AChE. The *in vitro* results show that the binding energies of columbamine, berberine and palmatine are -9.9, -10.6 and -9.8 Kcal/mol, respectively, which indicates that these small molecules can dock with AChE in a natural state. The visualized images are presented in Figure 6. All of these bioactive components interact with enzyme crystals at long and narrow hydrophobic pockets, which is similar to donepezil. Tyr-337 and Trp-286, which can maintain the geometry of the binding gorge and provide electrostatic balance, are important residues of the AChE protein (Kua et al., 2003). Columbamine, berberine and palmatine can interact with these residues *via* hydrophobic and - stacking interactions. Additionally, columbamine is stabilized by hydrophobic interactions with Tyr-341, Trp-286 and Phe-338, hydrogen bonds withPhe-295, Ser-293 and Arg-296, and - stacking interactions with Tyr-341 (Figure 6A). Berberine is stabilized by hydrophobic interactions with Tyr-341 and Phe-338, hydrogen bonds with Phe-295 and Arg-296, and - stacking interactions with Tyr-341 (Figure 6B). Palmatine

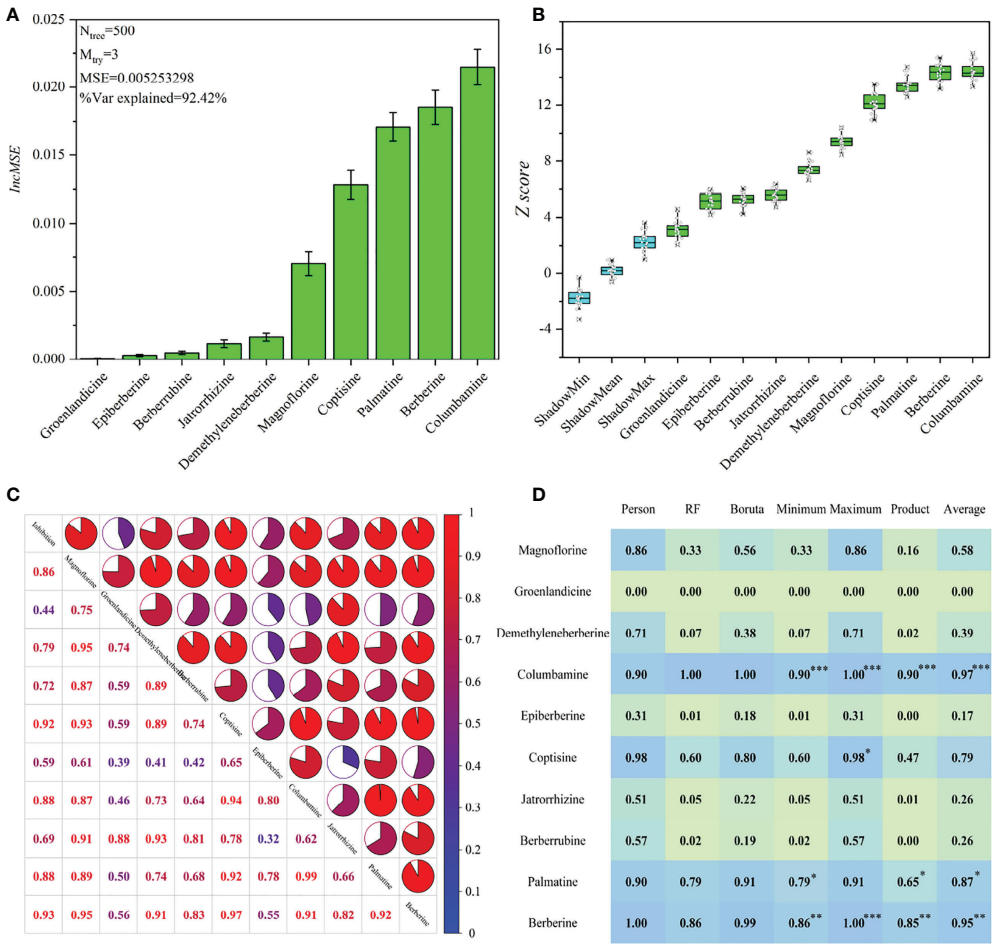


FIGURE 5
The results of spectrum-effect relationship based on different data processing methods. **(A)** the ranking result based on RF model; **(B)** the ranking result based on Boruta model; **(C)** the result of Pearson correlation analysis; **(D)** data fusion results, and ***, ** and * mean the first, second and third ranking of these components, respectively.

presents three strong hydrophobic interactions with Tyr-341, Tyr-72 and Phe-338 and a - stacking interaction with Tyr-341 (Figure 6C).

Based on X-ray crystallography in previous study (Gerlits et al., 2019), the specific reversible central AChE inhibitor of donepezil can interact with this receptor by forming stacking interactions, hydrophobic interactions and hydrogen bonds with amino acid residuesTrp-86, Trp-286, Tyr-337, Phe-338, Tyr-341, Ser-293, Phe-295, and Tyr-72. Our docking results show that columbamine, berberine and palmatine mainly interact with

TABLE 2 Binding sites and action forces between components and AchE protein.

Ligand Molecules	Binding Energy (Kcal/mol)	Hydrophobic Interactions	Hydrogen Bonds	π-π stacking Interaction
Columbamine	-9.9	TRP-286, PHE-338, TYR-341	SER-293, PHE-295, ARG-296	TRP-286, TYR-337, TYR-341
Berberine	-10.6	TRP-286, TYR-337, PHE-338, TYR-341	PHE-295, ARG-296	TYR-341
Palmatine	-9.8	TYR-72, TYR-337, PHE-338, TYR-341	–	TRP-286, TYR-341
Donepezil	-6.61	TRP-86, TYR-337, PHE-338, PHE-341	PHE-295, TYR-337	TRP-86, TRP-286, TYR-341

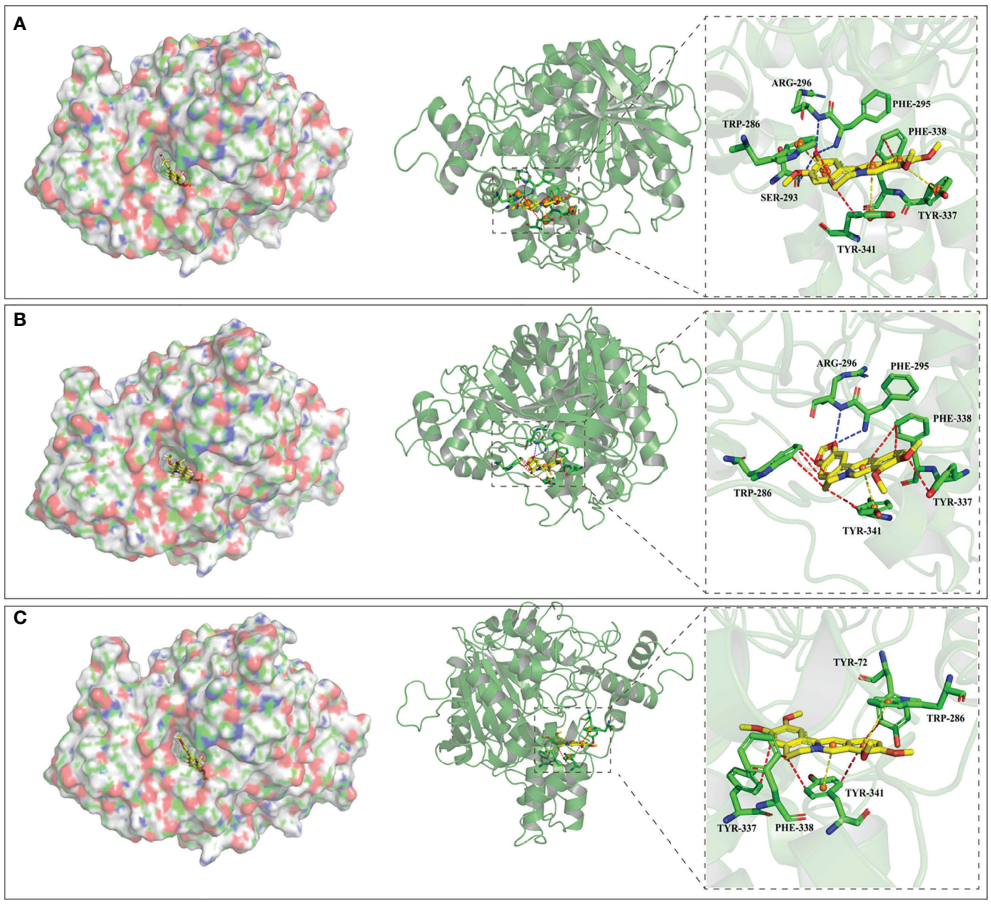


FIGURE 6
The visualization of molecular docking results. The letters of **(A–C)** represent the visualization result columbamine, berberine and palmatine, respectively. The left parts exhibit these compounds interact with AChE at a long andnarrow hydrophobic pockets, and the rights parts exhibit the detail of these interactions. Red, blue and yellow lines represent hydrophobic interaction, hydrogen bond and - stacking interaction, respectively.

these residues to dock towards the AChE protein. These computer simulation results explain the molecular mechanism of columbamine, berberine and palmatine against AChE, which is also consistent with the *in vitro* assay. In summary, these components are potential AChE inhibitors and can be considered quality markers to evaluate the anti-AChE activity of RC drugs from different *Coptis* species.

Discussion

Acetylcholine deficiency is an important factor that promotes the occurrence and progression of Alzheimer's disease. AChE is a serine hydrolase that can terminate nerve impulse transmission by transforming acetylcholine into acetate and choline in the central and peripheral nervous systems (Stanciu et al., 2019). Considering the side effects of current drugs for AChE inhibition, it is necessary to screen natural AChE inhibitors with fewer toxic side effects to treat neurological diseases. Natural products have a long application for the treatment of many ailments, and the ingredients originally isolated from them are relatively safe (Guan et al., 2017; Abate et al., 2021; Wang et al., 2022).

The roots and rhizomes of *Coptis* species have been used as RC drugs for thousands of years in traditional oriental medicine. Several ancient pharmacological writings in China record that these drugs can effectively treat brain diseases. Currently, RC drugs have been clinically applied to prevent and treat Alzheimer's disease. The investigation regarding the AChE inhibitory activity of *Coptis* species is insufficient. Thus, we designed an analysis strategy to investigate for the first time the anti-AChE potential of different extracts for RC drugs from four Coptis species and to analyse the role of the main alkaloids on this potency (Figure 7).

In vitro inhibitory assays proved that n-butanol was the optimized fraction for AChE inhibition, which is consistent

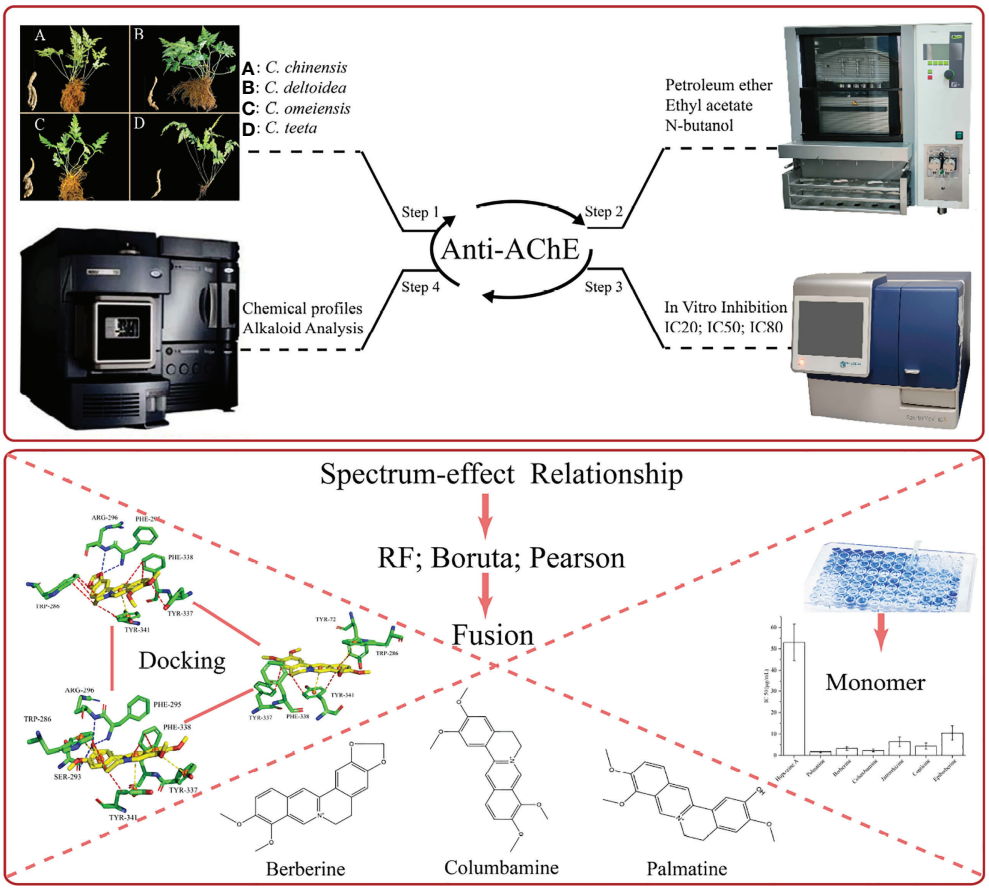


FIGURE 7
The overall analysis strategy of this article. The above section mainly describes the experimental process, and the following section mainly describes the analysis process.

with the previous studies (Zhao et al., 2016; Cao et al., 2018; Lin et al., 2020a). *C. teeta* can be used as a better source of RC drugs than the commonly used *C. chinensis* due to its relatively high AChE inhibition. Meanwhile, we focused on the rapid absolute quantitation of 10 common alkaloids in RC drugs based on our previous qualitative research. Compared with conventional liquid chromatography, the MRM mode of UPLC/QqQ-MS has excellent selectivity and sensitivity for these target compounds (Li et al., 2021). The result showed that these alkaloids in different drugs greatly varied. We also identified several trace compounds as markers to discriminate RC drugs from different sources, such as groenlandicine and magnoflorine for RC drugs from *C. deltoidea*, and berberrubine and demethyleneberberine for RC drugs from *C. teeta*. Overall, these findings can provide scientific evidence for the application of RC drugs to treat brain diseases.

Many studies have demonstrated that spectrum-effect relationship and molecular docking are powerful tools to interpret the mechanism of multi-component systems in herbal medicines that exert their therapeutic effect. We performed these techniques to analyse the relationship between alkaloids and AChE inhibition of RC drugs. Three data processing methods were applied to develop the spectrum-effect model, and a data fusion strategy was used to calculate a comprehensive result. Boruta algorithm indicates the importance of 10 alkaloids for AChE inhibition activity due to their higher *Z*-score than the shadow attributes. Comparatively, columbamine, berberine and palmatine were identified as the best AChE inhibitors. Considering the concentrations of these compounds in RC drugs, berberine is the most critical factor that determines the anti-AChE potential of RC drugs from different sources. Furthermore, we constructed a linear relationship between AChE inhibition activity and the first three alkaloids. This equation may be used to guide the RC drug selection to treat Alzheimer's disease. Compared with the co-crystallized ligands of donepezil, molecular docking identified the key binding forces between these alkaloids and the AChE protein at the mechanistic level. All of these results demonstrated that alkaloids

are the substance foundation for the anti-AChE potential, and columbamine, berberine and palmatine are quality markers to evaluate this activity in clinical application.

Conclusion

The present study developed *in vitro* inhibitory assays to investigate the anti-AChE potential of different extracts of RC drugs from *Coptis* species and the UPLC/QqQ-MS method to quantify the main alkaloids in the n-butanol fraction. Spectrum-effect relationship and molecular docking were applied to interpret the associations between these alkaloids and AChE inhibitory properties. Columbamine, berberine and palmatine were screened and identified as the quality markers responsible for the AChE inhibition activity. We hope that these results can contribute to the effective development and utilization of RC drugs from different sources, especially to provide certain evidence for their application in the treatment of Alzheimer's disease. Of course, the clinical application of these medicinal materials warrants further validation.

Data availability statement

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

Author contributions

LQ conceived and designed the research and wrote the manuscript. FZ and NL collected the experimental materials. JW, KN and YT performed the experiments. LX and YM revised the manuscript. All authors contributed to this article and approved the submitted version.

Funding

This work were financially supported by the National Natural Science Foundation of China (U19A201).

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.2022.1020309/full#supplementary-material

References

Abate, G., Zhang, L., Pucci, M., Morbini, G., Mac Sweeney, E., Maccarinelli, G., et al. (2021). Phytochemical analysis and anti-inflammatory activity of different ethanolic phyto-extracts of artemisia annual. *Biomolecules* 11, 975. doi: 10.3390/biom11070975

Ahmed, S., Khan, S. T., Zargaham, M. K., Khan, A. U., Khan, S., Hussain, A., et al. (2021). Potential therapeutic natural products against alzheimer's disease with reference of acetylcholinesterase. *Biomed. Pharmacother.* 139, 111609. doi: 10.1016/j.biopha.2021.111609

Altmann, A., Toloşi, L., Sander, O., and Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics* 26, 1340–1347. doi: 10.1093/bioinformatics/btq134

Amir Rawa, M. S., Nurul Azman, N. A., Mohamad, S., Nogawa, T., and Wahab, H. A. (2022). *In vitro* and in silico anti-acetylcholinesterase activity from macaranga tanarius and syzygium jambos. *Molecules* 27, 2648. doi: 10.3390/molecules27092648

Brewster, J. T., Dell'Acqua, S., Thach, D. Q., and Sessler, J. L. (2018). Classics in chemical neuroscience: Donepezil. *ACS Chem. Neurosci.* 10, 155–167. doi: 10.1021/acschemneuro.8b00517

Cao, T. Q., Ngo, Q.-M. T., Seong, S. H., Youn, U. J., Kim, J. A., Kim, J., et al. (2018). Cholinesterase inhibitory alkaloids from the rhizomes of *Coptis* chinensis. *Bioorg. Chem.* 77, 625–632. doi: 10.1016/j.bioorg.2018.01.038

Chang, Y., Zhang, D., Yang, G., Zheng, Y., and Guo, L. (2021). Screening of anti-lipase components of artemisia argyi leaves based on spectrum-effect relationships and hplc-ms/ms. *Front. Pharmacol.* 12, 675396. doi: 10.3389/fphar.2021.675396

Chen, H., Fan, G., and He, Y. (2017). Species evolution and quality evaluation of four *Coptis* herbal medicinal materials in southwest china. *3 Biotech.* 7, 1–8. doi: 10.1007/s13205-017-0679-8

Chen, Y., Qi, L., Zhong, F., Li, Y., Ke, W., and Ma, Y. (2021). Integrated metabolomics and ligand fishing approaches to screen the hypoglycemic ingredients from four *Coptis* medicines. *J. Pharm. Biomed. Anal.* 192, 113655. doi: 10.1016/j.jpba.2020.113655

da Silva Barbosa, D. C., Holanda, V. N., de Assis, C. R. D., de Oliveira Farias, J. C. R., Henrique doNascimento, P., da Silva, W. V., et al. (2020). Chemical composition and acetylcholinesterase inhibitory potential, in silico, of myrciaria floribunda (h. west ex willd.) o. berg fruit peel essential oil. *Ind. Crops Products* 151, 112372.

Eberhardt, J., Santos-Martins, D., Tillack, A. F., and Forli, S. (2021). Autodock vina 1.2. 0: New docking methods, expanded force field, and python bindings. *J. Chem. Inf. Model.* 61, 3891–3898. doi: 10.1021/acs.jcim.1c00203

Feinstein, W. P., and Brylinski, M. (2015). Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. *J. Cheminform.* 7, 1–10. doi: 10.1186/s13321-015-0067-5

Feng, X., Yan, D., Zhao, K.-J., Luo, J.-Y., Ren, Y.-S., Kong, W.-J., et al. (2011). Applications of microcalorimetry in the antibacterial activity evaluation of various rhizoma coptidis. *Pharm. Biol.* 49, 348–353. doi: 10.3109/13880209.2010.523428

Gerlits, O., Ho, K.-Y., Cheng, X., Blumenthal, D., Taylor, P., Kovalevsky, A., et al. (2019). A new crystal form of human acetylcholinesterase for exploratory room-temperature crystallography studies. *Chemico Biol. Interact.* 309, 108698. doi: 10.1016/j.cbi.2019.06.011

Guan, X.-R., Zhu, L., Xiao, Z.-G., Zhang, Y.-L., Chen, H.-B., and Yi, T. (2017). Bioactivity, toxicity and detoxification assessment of dioscorea bulbifera l.: a comprehensive review. *Phytochem. Rev.* 16, 573–601. doi: 10.1007/s11101-017-9505-5

Huang, L., Shi, A., He, F., and Li, X. (2010). Synthesis, biological evaluation, and molecular modeling of berberine derivatives as potent acetylcholinesterase inhibitors. *Bioorg. Medicinal Chem.* 18, 1244–1251. doi: 10.1016/j.bmc.2009.12.035

Ju, Y., and Tam, K. Y. (2022). Pathological mechanisms and therapeutic strategies for alzheimer's disease. *Neural Regeneration Res.* 17, 543. doi: 10.4103/1673-5374.320970

Kaufmann, D., Kaur Dogra, A., Tahrani, A., Herrmann, F., and Wink, M. (2016). Extracts from traditional chinese medicinal plants inhibit acetylcholinesterase, a known alzheimer's disease target. *Molecules* 21, 1161.

Kua, J., Zhang, Y., Eslami, A. C., Butler, J. R., and McCammon, J. A. (2003). Studying the roles of w86, e202, and y337 in binding of acetylcholine to acetylcholinesterase using a combined molecular dynamics and multiple docking approach. *Protein Sci.* 12, 2675–2684. doi: 10.1110/ps.03318603

Kursa, M. B., and Rudnicki, W. R. (2010). Feature selection with the boruta package. *J. Stat. Softw.* 36, 1–13. doi: 10.18637/jss.v036.i11

Lin, Y., Guo, H.-C., Kuang, Y., Shang, Z.-P., Li, B., Chen, K., et al. (2020a). Ache inhibitory alkaloids from *Coptis* chinensis. *Fitoterapia* 141, 104464. doi: 10.1016/j.fitote.2019.104464

Lin, Y., Zhao, W.-R., Shi, W.-T., Zhang, J., Zhang, K.-Y., Ding, Q., et al. (2020b). Pharmacological activity, pharmacokinetics, and toxicity of timosaponin aiii, a natural product isolated from anemarrhena asphodeloides bunge: A review. *Front. Pharmacol.* 11, 764. doi: 10.3389/fphar.2020.00764

Liu, X.-M., Tan, J.-P., Cheng, S.-Y., Chen, Z.-X., Ye, J.-B., Zheng, J.-R., et al. (2022). Comparative transcriptome analysis provides novel insights into the molecular mechanism of berberine biosynthesis in coptis chinensis. *Scientia Hortic.* 291, 110585. doi: 10.1016/j.scienta.2021.110585

Liu, Y., Wang, B., Shu, S., Li, Z., Song, C., Liu, D., et al. (2021). Analysis of the *Coptis* chinensis genome reveals the diversification of protoberberine-type alkaloids. *Nat. Commun.* 12, 1–13. doi: 10.1038/s41467-021-23611-0

Li, X., Zhang, C.-T., Ma, W., Xie, X., and Huang, Q. (2021). Oridonin: a review of its pharmacology, pharmacokinetics and toxicity. *Front. Pharmacol.* 12, 645824. doi: 10.3389/fphar.2021.645824

Lu, Y.-F., Li, D.-X., Zhang, R., Zhao, L.-L., Qiu, Z., Du, Y., et al. (2022). Chemical antioxidant quality markers of chrysanthemum morifolium using a spectrum-effect approach. *Front. Pharmacol.* 13. doi: 10.3389/fphar.2022.809482

Mtemeli, F., Ndlovu, J., Mugumbate, G., Makwikwi, T., and Shoko, R. (2022). Advances in schistosomiasis drug discovery based on natural products. *All Life* 15, 608–622. doi: 10.1080/26895293.2022.2080281

Obisesan, K. A., Jiménez-Carvelo, A. M., Cuadros-Rodriguez, L., Ruisánchez, I., and Callao, M. P. (2017). Hplc-uv and hplc-cad chromatographic data fusion for the authentication of the geographical origin of palm oil. *Talanta* 170, 413–418. doi: 10.1016/j.talanta.2017.04.035

Qi, L., Ma, Y., Zhong, F., and Shen, C. (2018). Comprehensive quality assessment for rhizoma *Coptidis* based on quantitative and qualitative metabolic profiles using high performance liquid chromatography, fourier transform near-infrared and fourier transform mid-infrared combined with multivariate statistical analysis. *J. Of Pharm. And Biomed. Anal.* 161, 436–443. doi: 10.1016/j.jpba.2018.09.012

Refaay, D. A., Abdel-Hamid, M. I., Alyamani, A. A., Abdel Mougib, M., Ahmed, D. M., Negm, A., et al. (2022). Growth optimization and secondary metabolites evaluation of anabaena variabilis for acetylcholinesterase inhibition activity. *Plants* 11, 735. doi: 10.3390/plants11060735

Roselli, M., Cavalluzzi, M. M., Bruno, C., Lovece, A., Carocci, A., Franchini, C., et al. (2016). Synthesis and evaluation of berberine derivatives and analogs as potential antiacetylcholinesterase and antioxidant agents. *Phytochem. Lett.* 18, 150–156. doi: 10.1016/j.phytol.2016.10.005

Shi, Z., Liu, Z., Liu, C., Wu, M., Su, H., Ma, X., et al. (2016). Spectrum-effect relationships between chemical fingerprints and antibacterial effects of lonicerae japonicae flos and lonicerae flos base on UPLC and microcalorimetry. *Front. Pharmacol.* 7, 12. doi: 10.3389/fphar.2016.00012

Speiser, J. L., Miller, M. E., Tooze, J., and Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst. Appl.* 134, 93–101. doi: 10.1016/j.eswa.2019.05.028

Stanciu, G. D., Luca, A., Rusu, R. N., Bild, V., Beschea Chiriac, S. I., Solcan, C., et al. (2019). Alzheimer's disease pharmacotherapy in relation to cholinergic system involvement. *Biomolecules* 10, 40. doi: 10.3390/biom10010040

Taqui, R., Debnath, M., Ahmed, S., and Ghosh, A. (2022). Advances on plant extracts and phytocompounds with acetylcholinesterase inhibition activity for possible treatment of alzheimer's disease. *Phytomed. Plus* 2, 100184. doi: 10.1016/j.phyplu.2021.100184

Vezzulli, F., Rocchetti, G., Lambri, M., and Lucini, L. (2022). Metabolomics combined with sensory analysis reveals the impact of different extraction methods on coffee beverages from coffea arabica and coffea canephora var. robusta. *Foods* 11, 807. doi: 10.3390/foods11060807

Wang, W., Zhu, S., Chen, H., Wu, N., Chen, H., and Wang, D. (2022). Development and validation of ultrahigh-performance liquid chromatography coupled with triple quadrupole mass spectrometry method for quantitative determination of ten active compounds in ge-gen-jiao-tai-wan. *J. Anal. Methods Chem.* 2022, 4713799. doi: 10.1155/2022/4713799

Yang, Q., Kang, Q., Huang, Q., Cui, Z., Bai, Y., and Wei, H. (2021). "Linear correlation analysis of ammunition storage environment based on pearson correlation analysis," in *Journal of physics: Conference series*, vol. 1948. (Hangzhou, China: IOP Publishing), 012064.

Zhang, C., Zheng, X., Ni, H., Li, P., and Li, H.-J. (2018). Discovery of quality control markers from traditional chinese medicines by fingerprint-efficacy modeling: Current status and future perspectives. *J. Pharm. Biomed. Anal.* 159, 296–304. doi: 10.1016/j.jpba.2018.07.006

Zhao, H., Zhou, S., Zhang, M., Feng, J., Wang, S., Wang, D., et al. (2016). An *in vitro* ache inhibition assay combined with uf-hplc-esi-q-tof/ms approach for screening and characterizing of ache inhibitors from roots of *Coptis* chinensis franch. *J. Pharm. Biomed. Anal.* 120, 235–240. doi: 10.1016/j.jpba.2015.12.025

Zhong, F., Huang, L., Qi, L., Ma, Y., and Yan, Z. (2020). Full-length transcriptome analysis of *Coptis deltoidea* and identification of putative genes involved in benzylisoquinoline alkaloids biosynthesis based on combined sequencing platforms. *Plant Mol. Biol.* 102, 477–499. doi: 10.1007/s11103-019-00959-y

Zhong, F., Shen, C., Qi, L., and Ma, Y. (2018). A multi-level strategy based on metabolic and molecular genetic approaches for the characterization of different coptis medicines using HPLC-UV and RAD-seq techniques. *Molecules* 23, 3090. doi: 10.3390/molecules23123090

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 02 November 2022
DOI 10.3389/fpls.2022.1035121

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Hexin Tan,
Second Military Medical University, China
Yogesh Kumar,
University Medical Center Hamburg-Eppendorf, Germany

*CORRESPONDENCE
Ying Xiao
xiaoyingtcm@shutcm.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 02 September 2022
ACCEPTED 10 October 2022
PUBLISHED 02 November 2022

CITATION
Shi X, Geng J, Feng J, Yang Y, Ma X, Chen W and Xiao Y (2022) Identification and investigation of a novel NADP+-dependent secoisolariciresinol dehydrogenase from *Isatis indigotica*. *Front. Plant Sci.* 13:1035121. doi: 10.3389/fpls.2022.1035121

Identification and investigation of a novel NADP$^+$-dependent secoisolariciresinol dehydrogenase from *Isatis indigotica*

Xiaoyi Shi[1,2†], Jiaran Geng[1†], Jingxian Feng[1†], Yingbo Yang[1,3], Xueqi Ma[1], Wansheng Chen[1,4] and Ying Xiao[1*]

[1]Research and Development Center of Chinese Medicine Resources and Biotechnology, The Ministry of Education (MOE) Key Laboratory for Standardization of Chinese Medicines, Institute of Chinese Materia Medica, Shanghai University of Traditional Chinese Medicine, Shanghai, China, [2]Shanghai Foreign Language School Affiliated to Shanghai International Studies University (SISU), Shanghai, China, [3]Jiangsu Kanion Pharmaceutical Co., Ltd., Lianyungang, China, [4]Department of Pharmacy, Changzheng Hospital, Naval Medical University (Second Military Medical University), Shanghai, China

Cofactors are crucial for the biosynthesis of natural compounds, and cofactor engineering is a useful strategy for enzyme optimization due to its potential to enhance enzyme efficiency. Secoisolariciresinol dehydrogenase (SIRD) was reported to convert secoisolariciresinol into matairesinol in an NAD$^+$-dependent reaction. Here, a SIRD designated as *Ii*SIRD2 identified from *Isatis indigotica* was found to utilize NADP$^+$ as the cofactor. To explore the structural basis for this unique cofactor preference, model-based structural analysis was carried out, and it was postulated that a variation at the GXGGXG glycine-rich motif of *Ii*SIRD2 alters its cofactor preference. This study paves way for future investigations on SIRD cofactor specificity and cofactor engineering to improve SIRD's catalytic efficiency.

KEYWORDS

cofactor specificity, *Isatis indigotica*, lignan, secoisolariciresinol dehydrogenase, matairesinol, structural biology

Abbreviations: SIRD, secoisolariciresinol dehydrogenase; TPP, thiamine pyrophosphate; FAD, flavin adenine dinucleotide; 2,5-DKG, 2,5-diketo-D-gluconic acid reductase; DIR, dirigent protein; PLR, pinoresinol-lariciresinol reductase; *Ii*, *Isatis indigotica*; *Pp*, Podophyllum peltatum.

Introduction

Cofactors are obligatory adducts in the catalytic machinery of numerous enzymes. As integral components of the holoenzymes, cofactors are imperatives for enzymatic and pathway functionality. Examples of cofactors include thiamine pyrophosphate (TPP) in pyruvate decarboxylase for yeast fermentation, flavin adenine dinucleotide (FAD) in acetyl-CoA-dehydrogenase for beta-oxidation of fatty acids, NADPH in adrenodoxin reductase for steroid hormone synthesis, etc. Without the effective participation of these cofactors, enzymes are incapable of efficiently transforming substrates into products. It is a challenging situation frequently encountered in metabolic engineering, which often involves introducing animal or plant metabolic pathways culled from nature into microorganisms with very different cellular environments and cofactor supplies (Akhtar and Jones, 2014). Besides, there is also the challenge of achieving cellular redox balance to enable biosynthesis at the maximum capacity, since cofactors can alter the intracellular redox state (Chen et al., 2018). To address these problems, researchers often resort to cofactor engineering. Previous successes in cofactor engineering include the optimization of vitamin C production. The formation of a vitamin C precursor, 2-keto-L-gluconic acid, is catalyzed by the NADPH-dependent 2,5-diketo-D-gluconic acid reductase (2,5-DKG). Banta et al. constructed 2,5-DKG mutants that could utilize both NADH and NADPH as the cofactor, eventually yielding an enzyme more active than the wild-type (Banta et al., 2002). This study highlights the significance of building cofactor specificity systems and identifying enzyme mutants with different cofactor preferences for cofactor engineering and pathway optimization.

Secoisolariciresinol dehydrogenase (SIRD) is an NAD^+-dependent enzyme that catalyzes the bioconversion of secoisolariciresinol into matairesinol. It is an important catalytic module for the biosynthetic pathways of lignans in plants. In one of the major biosynthetic pathways, coniferyl alcohol is dimerized by the plant dirigent protein (DIR) and converted into pinoresinol, which is then converted into secoisolariciresinol by pinoresinol-lariciresinol reductase (PLR) and matairesinol by SIRD in a stepwise manner (Figure 1) (Satake et al., 2015). The bioconversion catalyzed by SIRD lies in a key branch of the lignan biosynthetic pathways in plants; it determines the structural backbones of downstream bioactive lignans such as the anticancer podophyllotoxin and the anti-cancer, anti-inflammatory, and antimicrobial hinokinin, and hence it contributes to the structural and biological diversity of plant lignans (Marcotullio et al., 2014; Shah et al., 2021). Therefore, identifying efficient SIRD modules and expounding their mode of catalysis is essential for lignan biosynthetic engineering. In addition, matairesinol exhibits diverse biological activities such as anti-cancer, anti-oxidative, and immunoregulatory effects (Schröder et al., 1990), underscoring the importance of investigating SIRD's catalysis for enhancing matairesinol production efficiency for future medical applications (Su et al., 2015; Wu et al., 2021).

Previous studies on SIRD focused on elucidating its catalytic functions. Shen et al. verified the catalytic functions of two *Dysosma versipellis* SIRDs (*Dv*SIRD) (Shen et al., 2016); Xia et al. and Arneaud et al. achieved the functional expression of

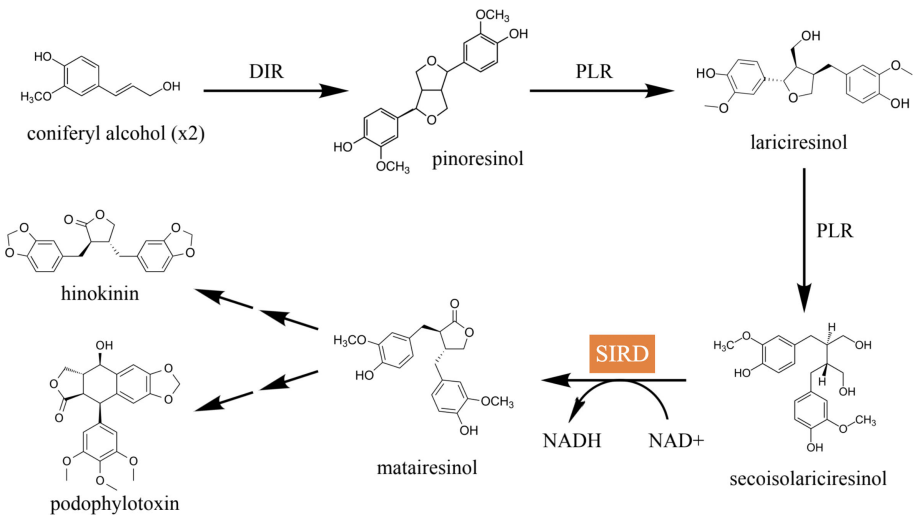


FIGURE 1
The major lignan biosynthetic pathway in plants. Coniferyl alcohol is dimerized by DIR into pinoresinol, which is then converted into lariciresinol and secoisolariciresinol in a step-wise manner catalyzed by PLR. Secoisolariciresinol is converted into matairesinol by SIRD, which is eventually converted into pharmacologically active compounds such as hinokinin and podophyllotoxin.

Podophyllum peltatum (*Pp*SIRD) in *Escherichia coli* and *Pichia pastoris*, respectively (Xia et al., 2001; Arneaud and Porter, 2015), and the structure of *Pp*SIRD was determined with X-ray crystallography (Youn et al., 2005). Additionally, Decembrino et al. assembled three plant enzymes including *Pp*SIRD in *E. coli* and achieved the production of the podophyllotoxin precursor pluviatolide, highlighting SIRD's future application in mass-producing the cancer drug precursor podophyllotoxin (Decembrino et al., 2020). All the previously reported SIRDs catalyze a strictly NAD^+-dependent bioconversion: (1) The catalytic triad of Ser^{153}, Lys^{171}, and Tyr^{167} scaffolds NAD^+ throughout the catalysis, and NAD^+'s binding to Tyr^{167} favors Tyr^{167}'s deprotonation. (2) Following substrate deprotonation and intramolecular hydride transfer, the intermediate lactol is formed and NAD^+ is reduced into NADH. (3) NADH is released from a triad and another NAD^+ binds the triad for the subsequent conversion of lactol into matairesinol in an analogous manner (Moinuddin et al., 2006).

Isatis indigotica Fort., belonging to the family *Cruciferae*, is a prevalent Chinese medicinal herb. Bioactive lignans and their corresponding derivatives have been identified as the major active ingredients of *I. indigotica*. In our previous study, notable progress has been made in understanding the biosynthetic pathway and regulatory mechanism of lignans in *I. indigotica* (Feng et al., 2021). Here, four SIRD genes were first identified in the *I. indigotica* genome. In particular, *Ii*SIRD2 was able to catalyze an $NADP^+$-dependent conversion of secoisolariciresinol into matairesinol, which presents the first report of an $NADP^+$-dependent SIRD. To fully understand the structural basis of its catalytic features, protein models were constructed using the *Pp*SIRD crystal structure as a template, revealing some unique features of *Ii*SIRD2. Based on molecular docking results, it was postulated that a variation at its GXGGXG motif enhances its affinity to $NADP^+$ as the cofactor.

Materials and methods

Materials

The chemicals used in the experiments were reagent or High-Performance Liquid Chromatography (HPLC) grade. Restriction enzymes were purchased from New England BioLabs; RNA extraction kit was purchased from TransGen Biotech; cDNA synthesis kit and one-step cloning kit were purchased from Novo Protein Scientific (Shanghai); PCR kit was purchased from Toyobo Biotech; Premix TaqTM DNA polymerase was purchased from TaKaRa Bio; Taq master mix was purchased from Shanghai Wonton Biotech. *I. indigotica* was planted at Shanghai University of Traditional Chinese Medicine (SHUTCM), and two-month-old plants were used for target gene cloning.

Identification of candidates genes

The whole genome of *I. indigotica* was used in this process (data unpublished). The TBtools program (https://github.com/CJ-Chen/TBtools) was used for sequence blasting. The "blast several sequences to a big database" function was used (outfmt: Table, NumofThreads: 2, E-value: 1e-5, NumofHits: 500, NumofAligns: 250). The protein sequence of five functional SIRDs were retrieved from GenBank, including *Ps*SIRD (GenBank ID: ALD51315.1), *Dv*SIRD (GenBank ID: ACB87357.1), *Dp*SIRD (GenBank ID: AHB18702.1), *Sh*SIRD (GenBank ID: ABN14311.1), and *Dt*SIRD (GenBank ID: ABD78859.1). Blasting these five SIRD protein sequences to the total protein database of *I. indigotica* using a tBLASTn algorithm, chromosome locations of all the hits were acquired. The "fasta extract" function in TBtools was then used to extract the protein sequences of these hits, and the acquired sequences were run against the SWISS-PROT protein database (as query sequences) in NCBI (https://www.ncbi.nlm.nih.gov/). The *I. indigotica* proteins matched with the reported functional SIRDs or the short-chain dehydrogenase family in the SWISS-PROT database were selected as SIRD candidates. Phylogenetic relationships were analyzed in MEGA 6.06 (https://www.megasoftware.net/) using the maximum likelihood method with the pairwise deletion option. Tree reliability was estimated using a bootstrap analysis of 1000 replicates.

Cloning of the *I. indigotica* candidate gene

I. indigotica leave tissues were collected and frozen by liquid nitrogen and ground to fine powders. The total RNA of *I. indigotica* was then isolated using TRIzol reagent and then reverse-transcribed into cDNA according to the manufacturers' instructions. Primers were designed as SIRD_CDS to amplify all the candidate genes in PCR using the *I. indigotica* total cDNA as the template (Table S1). The PCR products were separated in a 1% agarose gel. After cloning the amplified gene products into pMD^{TM}19-T, it was transformed into *E. coli* Top10 strain and sequenced by Sangon Biotech. Then, the candidate genes were amplified using primers SIRD_32a (Table S1) in PCR and cloned into the expression vector pET-32a at the *Not*I and *Xho*I restriction sites, thus generating 32a-*Ii*SIRD constructs. The constructs and the

pET-32a control were transformed into *E. coli* expression strain BL21 for heterologous expression.

Heterologous expression and protein purification

E. coli was grown overnight with shaking at 200 rpm in LB medium with 100 mg/L of ampicillin (LB Amp medium) and then inoculated into 500 ml of fresh LB Amp medium under the aforementioned condition. When the culture was grown to an optical density at 600 nm (OD600) of 0.6, 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce expression. The culture underwent 48 h of induction at 18°C with shaking at 80 rpm and bacteria were harvested by centrifugation at 7830 rpm for 5 min at 4°C. Cell pellets were resuspended in 10 ml of suspension buffer (50 mmol/L Tris-HCl, 20% glycerol, 10 mmol/L 2-mercaptoethanol, pH 8.0) for every 200 ml of culture. To break the cell, an ultrasonic cell crusher was employed under the following condition: power 35%, ultrasound 3 s, gap 2 s, 50 Hz, 10 min, maximum temperature 10°C. The suspensions were next centrifuged at 7830 rpm for 15 min at 4°C and impurities were removed using a 0.22 μm membrane filter to yield crude protein extracts. Protein purification was performed using a His Spin Trap column following the manufacturer's instructions (GE Healthcare). The purity of the His-tag-fused *Ii*SIRDs was examined by 12% SDS–PAGE, and the protein concentration was determined by the Bradford method (Bradford, 1976) with bovine serum albumin (BSA) as the standard.

Enzyme assay and LC-MS analysis

The assay mixture (100 μL) consisting of Tris buffer (20 mM, pH 8.8), 1 mM NADP (sodium salt), 500 μm secoisolariciresinol, and 10 μg of purified protein was incubated at 30°C with shaking at 300 rpm for 12 hours (overnight), and the pET-32a vector was used as negative control. The reaction mixtures were analyzed by LC-MS using a triple-quadrupole mass spectrometer (Model 6410, Agilent, Santa Clara, CA) according to the method previously reported (Xiao et al., 2015).

Structural modeling and molecular docking

First, the primary structure of *Ii*SIRD was analyzed based on its alignment with five other functional SIRDs and five short-chain dehydrogenases/reductases in the Jalview program (https://www.jalview.org/): *Ps*SIRD (GenBank ID: ALD51315.1), *Dv*SIRD (GenBank ID: ACB87357.1), *Dp*SIRD (GenBank ID: AHB18702.1), *Sh*SIRD (GenBank ID: ABN14311.1), *Dt*SIRD (GenBank ID: ABD78859.1), the Rv2002 gene product of *Mycobacterium tuberculosis* (PDB ID: 1NFF), 3 α, 20 β-hydroxysteroid dehydrogenase from *Streptomyces exfoliatus* (*Se*HSD; PDB ID: 2HSD), R-alcohol dehydrogenase from *Lactobacillus brevis* (*Lb*RADH; PDB ID: 1NXQ), glucose dehydrogenase from *Priestia megaterium* (*Pm*GluDH; PDB ID: 1GCO), and 3-hydroxyacyl-CoA dehydrogenase from *Rattus norvegicus* (*Rn*HAD; PDB ID: 1E6W). Then, the secondary structure of *Ii*SIRD was predicted on the PSIPRED workbench web-server (http://bioinf.cs.ucl.ac.uk/psipred/). Furthermore, protein models were constructed using the SWISS-MODEL web-server (https://swissmodel.expasy.org/) based on the previously reported crystal structures of *Podophyllum* SIRD (*Pp*SIRD) in the apoenzyme form (PDB ID: 2BGK), NAD^+-bound holoenzyme form (PDB ID: 2BGL), and the NAD^+-matairesinol-bound form (PDB ID: 2BGM) (Arneaud and Porter, 2015). It is important to note that the sequence identity between *Pp*SIRD and the modeled enzyme is calculated to be 45.91% on the SWISS-MODEL web-server, and considering that a protein sequence with more than 30% similarity to a known structure may frequently be predicted with the accuracy of a low-resolution X-ray structure (Xiang, 2006), the structural prediction is likely to be reliable. Then, protein tertiary structures were visualized in the PyMol program (https://pymol.org/2/). To examine the cofactor preference of *Ii*SIRD, molecular docking was performed using AutoDock Vina (https://vina.scripps.edu/). Moreover, the root-mean-square deviation of atomic positions (RMSD) was calculated using the SuperPose web server (http://superpose.wishartlab.com/) to determine the effect of point mutations on the enzymatic structure. Finally, to analyze the glycine-rich motif and the enzyme pocket, pocket-cavity search and volume prediction were carried out using the POCASA 1.1 web-server (http://g6altair.sci.hokudai.ac.jp/g6/service/pocasa/), and loop refinement was carried out using Modeller (https://salilab.org/modeller/).

Results and discussion

Identification of SIRD gene candidates

Four SIRD gene candidates were identified from the genome annotation pool of *I. indigotica*, and they were translated into amino acid sequences (Tables S2, 3). The upstream and downstream sequences of the four SIRD gene candidates were presented in Table S4. A phylogenic tree was then constructed together with five functional SIRD proteins previously reported (Figure S1). It was noticed that the *I. indigotica* candidates (*Ii*SIRDs) and the functional SIRDs cluster independently on the phylogenic tree, hinting that *I. indigotica* candidates might not share the same catalytic behavior with the known SIRDs.

Characterization of *Ii*SIRD activity

SDS-PAGE analysis of *Ii*SIRD1, *Ii*SIRD2, *Ii*SIRD3, and *Ii*SIRD4 expressions in *E.coli* revealed that the recombinant SIRDs have a molecular mass of 40~50 kDa (Fig. S2), which is close to the predicted molecular mass. LC-MS analysis results of enzyme assays showed that secoisolariciresinol was successfully converted into matairesinol by *Ii*SIRD2 using $NADP^+$ as the cofactor (Figure 2). The other three candidates show no enzymatic activity. Furthermore, *IiSIRD2's* gene sequence has been uploaded to GenBank (GenBank ID: OM777730).

The mechanism underlying the cofactor specificity of *Ii*SIRD2

To understand why *Ii*SIRD2 is capable of catalyzing an $NADP^+$-dependent reaction different from other functional SIRDs, protein models were constructed and structural analyses were carried out as follows.

Primary and Secondary Structure Analysis: Protein alignment (Figure 3A) revealed that similar to the five functional SIRDs, *Ii*SIRD2 is conserved at the highly conserved catalytic triad consisting of Ser^{177}, Tyr^{190}, and Lys^{194}, which was suggested to be essential for catalysis to occur (Youn et al., 2005). *Ii*SIRD2 is also conserved at Asp^{78} which accounts for SIRDs' specificity for NADH instead of NADPH (Youn et al., 2005). Pro^{220} and Val^{225}, two sites involved in cofactor stabilization are also conserved for *Ii*SIRD2 (Youn et al., 2005). Another important motif conserved for NAD^+-binding enzymes is the glycine-rich motif (GXGGXG; amino acid position 48~54), which is known to bind the pyrophosphate group of NAD^+ (Youn et al., 2005). Surprisingly, *Ii*SIRD2 exhibits a deviation (GXSGXG) from the conventional pattern (GXGGXG) at this motif (Figure 3A, Table 1), substituting serine for the glycine at site 51. Additionally, the five reported functional SIRDs contain isoleucine at the glycine-rich motif (GXGGIG), while *Ii*SIRD2 replaces isoleucine with leucine at site 53 (GXSGLG). Although it was previously reported that SIRD has specificity toward NAD^+ but not $NADP^+$, it was noticed from the enzyme assay result that *Ii*SIRD2 can utilize $NADP^+$ as its cofactor to catalyze the reaction, leading to the question of whether its unique glycine-rich motif accounts for the special cofactor preference to *Ii*SIRD2. Then, the secondary structure of *Ii*SIRD2 was predicted (Figure 3B): similar to the reported *Podophyllum peltatum* SIRD (*Pp*SIRD) (Arneaud and Porter, 2015), *Ii*SIRD2 monomer adopts an α/β domain structure, containing 7 β-strands flanked by 8 α-strands, reminiscent of the Rossmann fold typical for NAD(P)(H) binding.

Structural Modeling and Molecular Docking Studies: S51G and L53I were introduced to *Ii*SIRD2 to create models of three

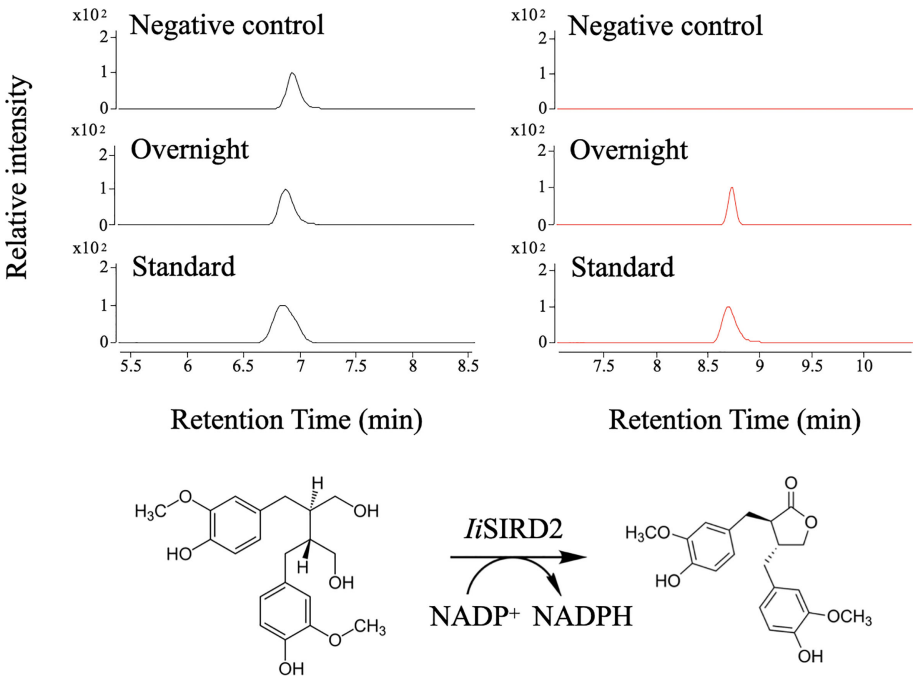


FIGURE 2
Biochemical assays for *Ii*SIRD2 function. Matairesinol was detected in the enzyme assay mixture, proving that *Ii*SIRD2 catalyzes an $NADP^+$-dependent reaction of converting secoisolariciresinol into matairesinol.

10.3389/fpls.2022.1035121

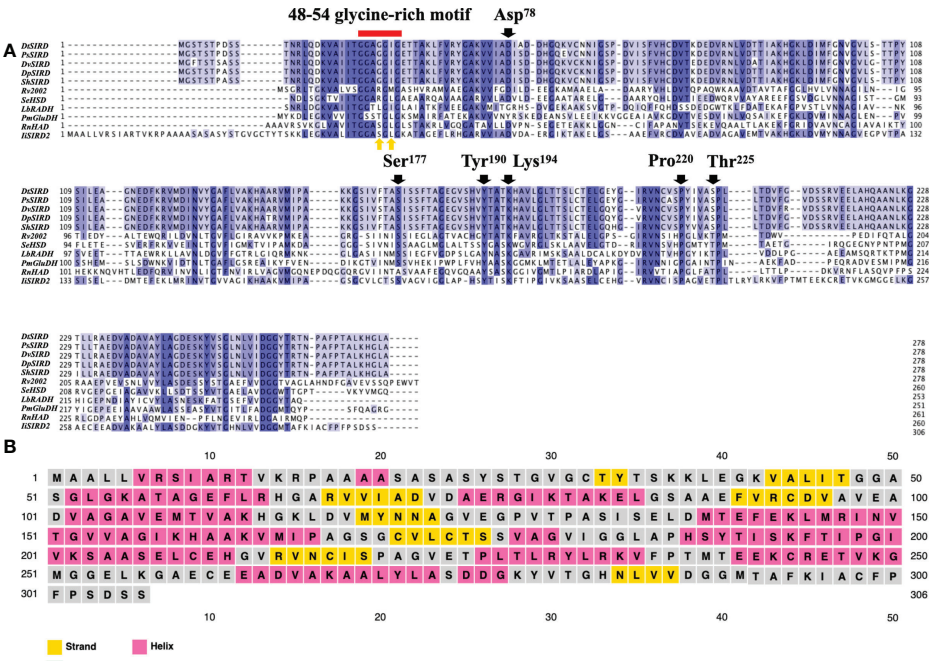


FIGURE 3
Primary and secondary structure of *Ii*SIRD2. **(A)** *Ii*SIRD2 is conserved at Ser^{177}, Tyr^{190}, Lys^{194}, Asp^{78}, Pro^{220}, and Val^{225}. However, *Ii*SIRD2 displays a glycine-rich motif (GXSGLG) different from the five functional SIRDs (GXGGIG). **(B)** Secondary structure prediction of *Ii*SIRD2.

protein mutants: *Ii*SIRD2-S51G, *Ii*SIRD2-L53I, and *Ii*SIRD2-S51G/L53I. Their binding affinity to NAD(P)$^+$ was tested to investigate the effect of these two sites on NAD(P)$^+$ binding affinities. To guarantee the reliability of the data, molecular docking was performed in triplicate and the average values were computed. The molecular docking results are shown in Table 2. In general, *Ii*SIRD2 (WT) exhibits a stronger affinity to both NADP$^+$ and NAD$^+$ compared with *Pp*SIRD. The reason for this alteration in cofactor preference perhaps lies in its variation at sites 51 and 53: introducing S51G to *Ii*SIRD2 decreases its affinity to NADP$^+$ significantly and slightly lowers its affinity to NAD$^+$, suggesting that Ser^{51} enhances *Ii*SIRD2's affinity to both coenzymes, especially NADP$^+$, thus explaining the enzyme assay result. Additionally, it is interesting that L53I only slightly alters binding affinity to both cofactors, while both S51G and S51G/L53I lower NAD$^+$ binding affinity by ~0.3 kcal/mol, indicating that Leu^{53} might play a minor role in determining cofactor preference compared with Ser^{51}. Based on molecular docking results, it is safe to postulate that site 51 within the glycine-rich motif is essential in determining NAD$^+$-binding affinity, while site 53 is comparatively less important. Furthermore, the RMSD value upon point mutation was determined using the SuperPose web server to evaluate the magnitude of structural change caused by the introduction of S51G and L53I to *Ii*SIRD2. In line with molecular docking results, S51G has a high RMSD value of 3.33, whereas L53I has an insignificant RMSD value of 0.84, showing that site 51 is more relevant in defining the structure of the enzyme. Replacing glycine with serine at site 51 might serve to enhance the enzyme's affinity to NADP$^+$ while substituting isoleucine for leucine at site 53 might counterbalance that change by slightly lowering its affinity to NADP$^+$. In brief, the unique glycine-rich motif of *Ii*SIRD2 offers a new perspective on how such a motif correlates with SIRD's cofactor preference and explains the NADP$^+$ specificity of *Ii*SIRD2.

Analysis of the Glycine-rich Motif and the Enzyme Pocket: Based on the above analysis, attention was then dedicated to the glycine-rich motif (Figures 4A, B) and the substrate-binding pocket (Figure 4C). At the glycine-rich motif, Ser^{51} pokes out from the traditional structure, leading to a

TABLE 1 Variations of the glycine-rich motif (Amino acid variations are marked in red).

Genes	The Glycine-rich Motif (48-54)
DtSIRD	GGAGGIG
PsSIRD	GGAGGIG
DvSIRD	GGAGGIG
DpSIRD	GGAGGIG
ShSIRD	GGAGGIG
IiSIRD2	GGASGLG

10.3389/fpls.2022.1035121

TABLE 2 Molecular docking results (Average values calculated from triplicate results were highlighted in bold).

		1	2	3	average
*Pp*SIRD (2BGL)	NAD+	-9	-8	-8.6	**-8.53**
	NADP+	-8.1	-8.5	-8.1	**-8.23**
*Ii*SIRD (WT)	NAD+	-10.1	-9.6	-9.1	**-9.60**
	NADP+	-9.8	-10.2	-9.5	**-9.83**
S51G	NAD+	-9.2	-9.3	-9.5	**-9.33**
	NADP+	-8.7	-9.3	-8.7	**-8.90**
L53I	NAD+	-9.6	-9.7	-9.6	**-9.63**
	NADP+	-9.8	-9.4	-10.1	**-9.77**
S51G, L53I	NAD+	-9.4	-9.1	-9.4	**-9.30**
	NADP+	-8.7	-8.7	-8.7	**-8.70**

conformational change that might alter the way NAD(P)$^+$ binds the motif due to steric hindrance (Figures 4A, B). NAD(P)$^+$ might be able to bind the motif in a more favorable position, rendering NAD(P)$^+$ the cofactor of *Ii*SIRD2. In brief, this conformational change offers some clues to the unique cofactor affinity of *Ii*SIRD2. The volume of the catalytic pocket was predicted. *Pp*SIRD (PDB ID: 2BGM) and *Ii*SIRD2 were predicted to have pocket volumes of 1340 and 1012, respectively (unit gird size of 1 Å). Interestingly, a closer look at the enzyme's tertiary structures reveals a short loop structure at *Ii*SIRD2's pocket gate (Figure 4C), and loop refinement with Modeller indicates that this structure may have a significant influence on the pocket structure, which warrants more examination in the future.

Conclusions and perspectives

This study presents the first report of an NADP$^+$-dependent SIRD. Protein model analysis and molecular docking studies revealed the role the glycine-rich motif plays in determining SIRD's cofactor preference, shedding some light on SIRD's catalytic mechanism. Since SIRD is an important entry point for downstream lignan synthesis, comprehending its catalytic behaviors is crucial for understanding the biological diversity of these health-protecting lignans.

The identification of the NADP$^+$-dependent *Ii*SIRD2 and its unique glycine-rich motif pave way for future research on cofactor engineering, a nascent strategy of interest that might be employed to evolve other SIRDs to alter cofactor specificity

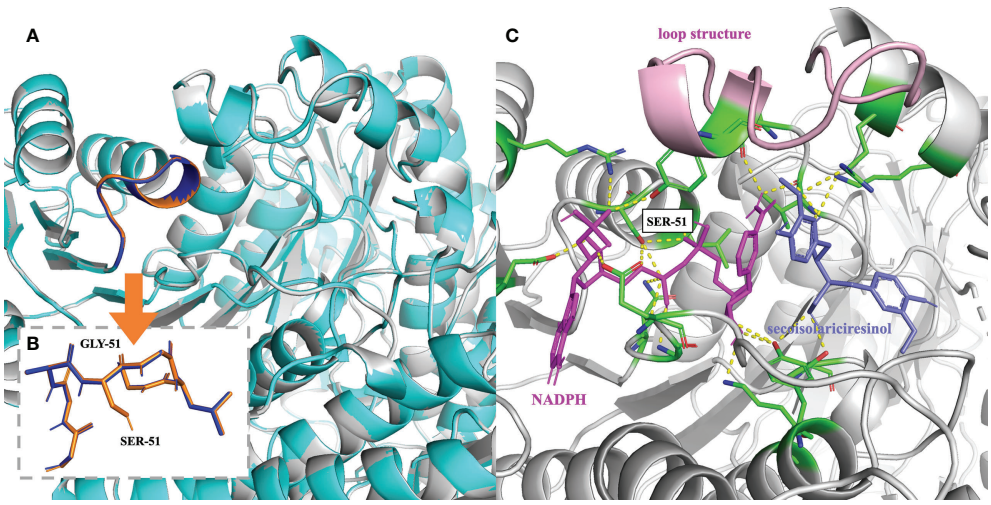


FIGURE 4
Mechanism underlying the cofactor specificity of *Ii*SIRD2. **(A)** Superimposition of *Ii*SIRD2 (gray) and *Ii*SIRD-S51G (cyan). The glycine-rich motif is highlighted (*Ii*SIRD-S51G: blue; *Ii*SIRD: orange). **(B)** Superimposition of the glycine-rich motif of *Ii*SIRD-S51G (blue) and *Ii*SIRD2 (orange). Steric hindrance caused by Ser^{51} may alter the way NAD(P)$^+$ binds the motif. **(C)** Structure of *Ii*SIRD (gray) bound with NADPH (magenta) and secoisolariciresinol (slate). Yellow-dotted lines represent potential hydrogen bonds, and key residues including Ser^{51} are marked in green. The loop structure at the pocket entry is marked in pink.

for enzyme optimization. Besides the success of enhancing the efficiency of vitamin C production by altering the cofactor preference of the relevant enzyme (Banta et al., 2002), phosphite dehydrogenase mutants with an enhanced preference for $NADP^+$ were recently demonstrated to be applicable to the establishment of an NADPH regeneration system for an NADPH-dependent reaction (Abdel-Hady et al., 2021), hinting that the SIRD cofactor specificity system enriched by the $NADP^+$-dependent *Ii*SIRD2 might provide a promising alternative cofactor regeneration machinery for practical applications to reduce production costs and increase efficiency for many NADPH-dependent reactions. Finally, this study also contributes to the full elucidation of the *I. indigotica* lignan biosynthetic pathway and thus helps to prompt the possibility of a potentially efficient enzyme-based metabolic engineering for the large-scale production of health-protecting lignans including matairesinol and its derivatives.

To conclude, this study reports the first $NADP^+$-dependent SIRD, providing theoretical support for fully expounding *Ii*SIRD2's catalytic characteristics and molecular mechanism, paving way for cofactor engineering, and hence having important implications regarding the efficient bioconversion of matairesinol.

Data availability statement

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

Author contributions

XS, JG, YY, and XM carried out the experiments and analyzed data. XS and JF performed protein modeling and molecular docking studies. XS wrote the manuscript. XS, YX, and JG made manuscript revisions. YX and WC supervised the study. All authors contributed to the article and approved the submitted version.

Funding

This work was sponsored by the National Natural Science Foundation of China (81874335, 32170402, and 31872665) and Shanghai local Science and Technology Development Fund Program guided by the Central Government (YDZX20203100002948).

Conflict of interest

Author YY was employed by Jiangsu Kanion Pharmaceutical Co., Ltd.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1035121/full#supplementary-material

References

Abdel-Hady, G. N., Ikeda, T., Ishida, T., Funabashi, H., Kuroda, A., and Hirota, R. (2021). Engineering cofactor specificity of a thermostable phosphite dehydrogenase for a highly efficient and robust nadph regeneration system. *Front. Bioeng Biotechnol.* 9. Available at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.647176 (Accessed October 15, 2022).

Akhtar, M. K., and Jones, P. R. (2014). Cofactor engineering for enhancing the flux of metabolic pathways. *Front. Bioeng Biotechnol.* 2. doi: 10.3389/fbioe.2014.00030

Arneaud, S. L., and Porter, J. R. (2015). Investigation and expression of the secoisolariciresinol dehydrogenase gene involved in podophyllotoxin biosynthesis. *Mol. Biotechnol.* 57, 961–973. doi: 10.1007/s12033-015-9888-8

Banta, S., Swanson, B. A., Wu, S., Jarnagin, A., and Anderson, S. (2002). Optimizing an artificial metabolic pathway: engineering the cofactor specificity of corynebacterium 2,5-diketo-d-gluconic acid reductase for use in vitamin c biosynthesis. *Biochemistry* 41, 6226–6236. doi: 10.1021/bi015987b

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Analytical Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Chen, X., Gao, C., Guo, L., Hu, G., Luo, Q., Liu, J, et al. (2018). Dceo biotechnology: tools to design, construct, evaluate, and optimize the metabolic pathway for biosynthesis of chemicals. *Chem. Rev.* 118, 4–72. doi: 10.1021/acs.chemrev.6b00804

Decembrino, D., Ricklefs, E., Wohlgemuth, S., Girhard, M., Schullehner, K., Jach, G., et al. (2020). Assembly of plant enzymes in e. coli for the production of the valuable (-)-podophyllotoxin precursor (-)-pluviatolide. *ACS Synth. Biol.* 9, 3091–3103. doi: 10.1021/acssynbio.0c00354

Feng, J., Huang, D., Yang, Y., Chen, J., Qiu, S., Lv, Z., et al. (2021). Isatis indigotica: from (ethno) botany, biochemistry to synthetic biology. *Mol. Horticulture* 1, 17. doi: 10.1186/s43897-021-00021-w

Marcotullio, M. C., Pelosi, A., and Curini, M. (2014). Hinokinin, an emerging bioactive lignan. *Molecules* 19, 14862–14878. doi: 10.3390/molecules190914862

Moinuddin, S. G., Youn, B., Bedgar, D. L., Costa, M. A., Helms, G. L., Kang, C., et al. (2006). Secoisolariciresinol dehydrogenase: mode of catalysis and stereospecificity of hydride transfer in podophyllum peltatum. *Org. Biomol. Chem.* 4, 808–816. doi: 10.1039/b516563f

Satake, H., Koyama, T., Bahabadi, S. E., Matsumoto, E., Ono, E., and Murata, J. (2015). Essences in metabolic engineering of lignan biosynthesis. *Metabolites* 5, 270–290. doi: 10.3390/metabo5020270

Schröder, H. C., Merz, H., Steffen, R., Müller, W. E., Sarin, P. S., Trumm, S., et al. (1990). Differential *in vitro* anti-hiv activity of natural lignans. *Z Naturforsch. C J. Biosci.* 45, 1215–1221. doi: 10.1515/znc-1990-11-1222

Shah, Z., Gohar, U. F., Jamshed, I., Mushtaq, A., Mukhtar, H., Zia-Ul-Ha, M., et al. (2021). Podophyllotoxin: history, recent advances and future prospects. *Biomolecules* 11, 603. doi: 10.3390/biom11040603

Shen, Y., Chen, R. D., Xie, K. B., Zou, J. H., and Dai, J. G. (2016). [cloning, expression and functional identification of secoisolariciresinol dehydrogenase gene from dysosma versipellis callus]. *Zhongguo Zhong Yao Za Zhi* 41, 4568–4571. doi: 10.4268/cjcmm20162414

Su, S., Cheng, X., and Wink, M. (2015). Cytotoxicity of arctigenin and matairesinol against the t-cell lymphoma cell line ccrf-cem. *J. Pharm. Pharmacol.* 67, 1316–1323. doi: 10.1111/jphp.12426

Wu, Q., Wang, Y., and Li, Q. (2021). Matairesinol exerts anti-inflammatory and antioxidant effects in sepsis-mediated brain injury by repressing the mapk and nf-κb pathways through up-regulating ampk. *Aging (Albany Ny)* 13, 23780–23795. doi: 10.18632/aging

Xia, Z. Q., Costa, M. A., Pelissier, H. C., Davin, L. B., and Lewis, N. G. (2001). Secoisolariciresinol dehydrogenase purification, cloning, and functional expression. implications for human health protection. *J. Biol. Chem.* 276, 12614–12623. doi: 10.1074/jbc.M008622200

Xiang, Z. (2006). Advances in homology protein structure modeling. *Curr. Protein Pept. Sci.* 7 (3), 217–227. doi: 10.2174/138920306777452312

Xiao, Y., Ji, Q., Gao, S., Tan, H., Chen, R., Li, Q., et al. (2015). Combined transcriptome and metabolite profiling reveals that iiplr1 plays an important role in lariciresinol accumulation in isatis indigotica. *J. Exp. Bot.* 66, 6259–6271. doi: 10.1093/jxb/erv333

Youn, B., Moinuddin, S. G., Davin, L. B., Lewis, N. G., and Kang, C. (2005). Crystal structures of apo-form and binary/ternary complexes of podophyllum secoisolariciresinol dehydrogenase, an enzyme involved in formation of health-protecting and plant defense lignans. *J. Biol. Chem.* 280, 12917–12926. doi: 10.1074/jbc.M413266200

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 14 November 2022
DOI 10.3389/fpls.2022.1049015

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Mariam Gaid,
Independent Researcher,
Braunschweig, Germany
Amr El-Demerdash,
John Innes Centre, United Kingdom

*CORRESPONDENCE
Zhenxing Zou
zouzhenxing@csu.edu.cn
Haibo Tan
tanhaibo@scbg.ac.cn

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

RECEIVED 20 September 2022
ACCEPTED 14 October 2022
PUBLISHED 14 November 2022

CITATION
Chen Y, Wang H, Ke X, Sang Z, Kuang M, Peng W, Tan J, Zheng Y, Zou Z and Tan H (2022) Five new secondary metabolites from an endophytic fungus *Phomopsis* sp. SZSJ-7B. *Front. Plant Sci.* 13:1049015. doi: 10.3389/fpls.2022.1049015

Five new secondary metabolites from an endophytic fungus *Phomopsis* sp. SZSJ-7B

Yan Chen[1,2,3], Huan Wang[3], Xin Ke[1,2,3], Zihuan Sang[1,2,3], Min Kuang[1,2], Weiwei Peng[1,2,3], Jianbing Tan[1,2], Yuting Zheng[1,2], Zhenxing Zou[1,2]* and Haibo Tan[1,2,3]*

[1]Xiangya School of Pharmaceutical Sciences, Central South University, Changsha, China, [2]Hunan Key Laboratory of Diagnostic and Therapeutic Drug Research for Chronic Diseases, Central South University, Changsha, China, [3]Key Laboratory of South China Agricultural Plant Molecular Analysis and Genetic Improvement, Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

Two previously undescribed lactones, phomolides A and B (**1** and **2**), and three new sesquiterpenoids, phomenes A–C (**3**–**5**), together with one known compound, colletotricholide A (**6**), were isolated from the endophytic fungus *Phomopsis* sp. SZSJ-7B. Their chemical structures, including the absolute configurations, were comprehensively established by extensive analyses of NMR, high-resolution electrospray ionization mass spectrometry, electronic circular dichroism powered by theoretical calculations, and X-ray diffractions. Moreover, the cytotoxic and antibacterial activities of compounds **1**–**6** were also evaluated, and the results demonstrated that compound **2** showed significant antibacterial effects towards methicillin-resistant *Staphylococcus aureus* and *S. aureus* strains with minimum inhibitory concentration as low as 6.25 μg/ml, which was comparable to that of the clinical drug vancomycin. Moreover, all compounds showed no cytotoxic activity.

KEYWORDS

endophytic fungus, *Phomopsis*, secondary metabolites, antibacterial activity, cytotoxic activity

1 Introduction

Endophytes play an important role of producing biologically meaningful natural products and can be considered as a strategically promising bio-resource of significantly economic potential for the agrochemical and pharmaceutical industries (Gouda et al., 2016). It is well documented that the genus *Phomopsis* can generate structurally diverse and pharmaceutically active secondary metabolites (Huang et al., 2008; Yu et al., 2008; Hemtasin et al., 2011), including xanthones (Ding et al., 2013), α-pyrones (Cai et al., 2017), steroids (Hu et al., 2017), sesquiterpenes (Xie et al., 2018), diterpenes

(Wei et al., 2014), triterpenes (Li et al., 2008), oblongolides (Bunyapaiboonsri et al., 2010), pyrenocines (Hussain et al., 2012), alkaloids (Chen et al., 2019), cytochalasins (Yan et al., 2016), *etc*. Moreover, these intriguing natural compounds shared various biological activities such as antitumor (Pavao et al., 2016), immunosuppressive (Wei et al., 2014), antifungal (Krohn et al., 2011), antioxidant (Chen et al., 2018), antibacterial (Jouda et al., 2016), anti-inflammatory (Xu ;et al., 2019b), and α-glucosidase inhibitory effects (Huang et al., 2018).

Our group pursued continuous research commitments towards discovering structurally fascinating and biologically significant natural products from endophytic fungi in recent years, and a series of metabolites with excellent antibacterial and antitumor activities from endophytic fungi of the genus *Phomopsis* have been reported (Xu et al., 2019a; Chen et al., 2020; Liu et al., 2021). In continuation of our ongoing endeavors, a strain of *Phomopsis* sp. SZSJ-7B isolated from the fresh leaves of *Alpinia shengzhen*, which is a beautiful horticultural plant (Zingiberaceae), was chosen as the appealing target for the chemical constituent investigation. Preliminary thin-layer chromatography and high-performance liquid chromatography (HPLC) screenings of the strain SZSJ-7B were conducted, and the experimental data showed that its ethyl acetate (EtOAc) extracts exhibited a remarkable diversity of secondary metabolites. A further systematical chemical study of the strain led to the isolation of five previously undescribed metabolites including two lactones, phomolides A and B, and three sesquiterpenoids, phomenes A–C. Herein the details of the extraction, purification, structure elucidation, and their biological evaluation are described.

2 Results and discussion

Compound **1** was isolated as a white solid. Its molecular formula $C_{11}H_{12}O_4$ was deduced from high-resolution electrospray ionization mass spectrometry (HRESIMS) *m/z* 209.0815 $[M + H]^+$ [calculated (calcd) for $C_{11}H_{13}O_4$, 209.0808], indicating six degrees of hydrogen deficiency. The infrared (IR) spectrum of **1** showed obvious absorption bands at 3,325, 1,699, and 1,022 cm^{-1} and revealed the presence of corresponding hydroxyl and carbonyl functionalities together with the ether bonds. The ^1H nuclear NMR data (Table 1) of **1** showed two singlet methyl groups (δ_H 2.12 and 2.57) and an upfield doublet methyl group (δ_H 1.65, d, J = 5.2 Hz). Its ^{13}C NMR (Table 1) and heteronuclear single quantum coherence (HSQC) spectra showed 11 carbon signals including six quaternary carbons (δ_C 163.2, 161.9, 159.1, 143.0, 120.6, and 104.5), two oxymethine (δ_C 98.7 and 97.7), and three methyl groups (δ_C 18.7, 16.1, and 10.0).

Moreover, the significant heteronuclear multiple bond correlation (HMBC) correlations (Figure 1) from H_3-9 to C-3 (δ_C 104.5), C-7 (δ_C 120.6), and C-8 (δ_C 143.0), H_3-10 to C-6 (δ_C 161.9), C-7, and C-8, and H-5 to C-3 and C-7 strongly suggested the existence of a penta-substituted benzene ring. In addition, on the basis of the ^1H–^1H correlation spectroscopy (COSY) fragment of C-1/C-11, the obvious HMBC correlations from H_3-11 to C-1 (δ_C 97.7) coupling with H-1 to C-2 (δ_C 163.2) and C-4 (δ_C 159.1) conclusively confirmed the planar structure of **1** as shown in Figure 2. In order to further clarify the absolute configuration, electronic circular dichroism (ECD) calculation of **1** was performed on mPW1PW91/SVP level of theory. As a

TABLE 1 ^1H (500 MHz) and ^{13}C (125 MHz) NMR data of 1 and 2 in CD_3OD (δ in ppm, J in Hz).

1			**2**					
No.	**δ_H (J in Hz)**	**δ_C**	**No.**	**δ_H (J in Hz)**	**δ_C**	**No.**	**δ_H (J in Hz)**	**δ_C**
1	5.64, q, (5.2)	97.7, CH	1	2.24, m 1.93, m	31.8, CH_2	13	2.95, s	46.5, CH_2
2		163.2, C	2	1.31, m	29.1, CH_2	14	0.95, s	19.9, CH_3
3		104.5, C	3	1.55, m 1.47, m	31.3, CH_2	15	0.84, d, (6.5)	14.7, CH_3
4		159.1, C	4	1.72, dt, (6.5, 10.6)	38.5, CH	1’		164.1, C
5	6.32, s	98.7, CH	5		39.0, C	2’		104.1, C
6		161.9, C	6α 6β	2.19, m 1.03, m	34.8, CH_2	3’		160.9, C
7		120.6, C	7	2.44, m	28.3, CH	4’	6.35, d, (2.0)	99.8, CH
8		143.0, C	8α 8β	1.83, m 2.12, m	28.5, CH_2	5’		161.6, C
9	2.57, s	16.1, CH_3	9	5.34, m	116.8, CH	6’	6.50, d, (2.0)	114.4, CH
10	2.12, s	10.0, CH_3	10		146.7, C	7’		145.9, C
11	1.65, d, (5.2)	18.7, CH_3	11		59.5, C	8’	2.57, s	20.8, CH_3
			12	5.46, s	100.8, CH			

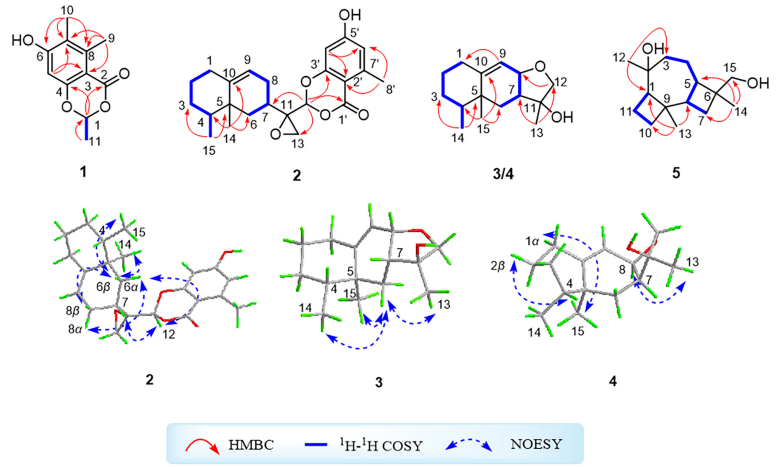


FIGURE 1
^{1}H–^{1}H correlation spectroscopy and key heteronuclear multiple bond correlation correlations of compounds **1**–**5** and key nuclear Overhauser effect spectroscopy correlations of compounds **2**–**4**.

result, the experimental ECD spectrum perfectly matched with the calculated ECD spectrum of 1*S* configuration for **1**, showing the same clear Cotton effect at 205 nm. Thus, the absolute configuration of **1** was determined to be 1*S* (Figure 3), and it was revealed as a new, natural, rarely occurring acetal derivative, which was given the trivial name phomolide A.

Compound **2** was isolated as a colorless oil. Its molecular formula $C_{23}H_{28}O_5$ was deduced from its HRESIMS spectrum with a molecular ion peak observed at *m/z* 385.2007 $[M + H]^+$ (calcd for $C_{23}H_{29}O_5$, 385.2010), which chemologically implied 10 degrees of hydrogen deficiency. Moreover, the IR spectrum of compound **2** showed a series of characteristic absorption bands at 3,357, 1,714, and 1,020 cm^{-1}, which were attributed to hydroxyl and carbonyl functional moieties as well as ether bonds. The ^{1}H NMR data (Table 1) of **2** showed typical proton resonances for three methyl groups at δ_H 0.95 (s, H_3-14), 0.84 (d, J = 6.5 Hz, H_3-15), and 2.57 (s, H-8'), an oxygenated methine moiety at δ_H 5.46 (s, H-12), and two aromatic protons at δ_H 6.35 (d, J = 2.0 Hz, H-4′) and δ_H 6.50 (d, J = 2.0 Hz, H-6′) together with an olefinic proton at δ_H 5.34 (m, H-9). The ^{13}C NMR (Table 1) and HSQC spectra exhibited 24 carbon signals comprising three methyls, six methylenes, six methines, and nine quaternary carbons. The ^{1}H–^{1}H COSY spectrum of **2** revealed the existence of two independent spin systems of H_2-1/H_2-2/H_2-3/H-4/H_3-15 and H_2-6/H-7/H_2-8/H-9.

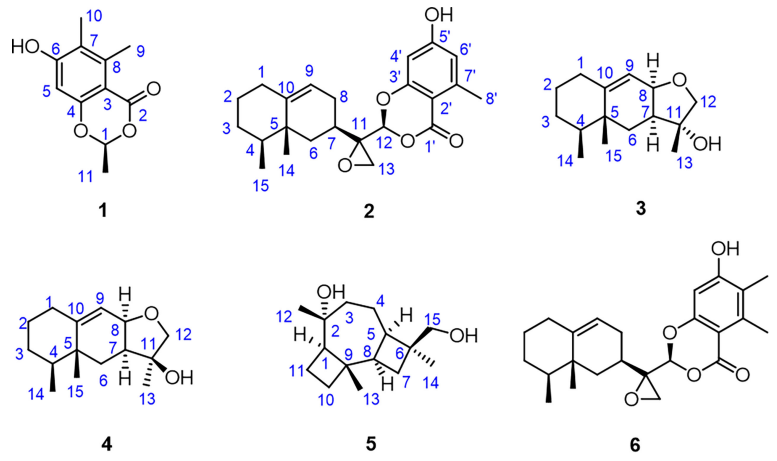


FIGURE 2
Structures of compounds **1**–**6**.

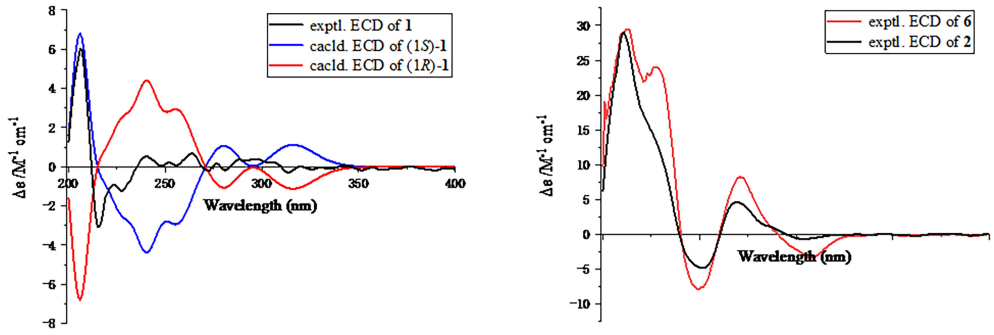


FIGURE 3
Experimental and calculated electronic circular dichroism spectra of compounds **1**, **2**, and **6**.

Compound **2** was further conclusively revealed as a novel meroterpenoid with eremophilan and acetophenone units conjugating as an acetal skeleton after the careful comparison of 1D NMR data of **2** with those of the known compound colletotricholide A **(6)** (Zhao et al., 2020), which was also co-isolated from this strain. The main difference in NMR spectra between **2** and colletotricholide A **(6)** was attributed to the lack of a methyl group in **2** at C-6' position. This speculation could be further verified by the ^{1}H and ^{13}C NMR signals for H-6' (δ_H 6.50) and C-6' (δ_C 114.4) in **2** and the key HMBC correlations from H-6' to C-1' (δ_C 164.1) and C-4' (δ_C 99.8). Therefore, the planar structure of **2** was identified as shown in Figure 2.

The relative configuration of **2** was determined by the nuclear Overhauser effect spectroscopy (NOESY) experiment (Figure 1). As shown in Figure 1, the NOESY correlations of H-6β with H-8β, H_3-14, and H_3-15 revealed that these protons were on the same face and assumed as β-oriented, while the correlations of H-6α with H-8α together with H-12 with H-7 and H-8α indicated that H-7 and H-12 were α-oriented. The CD spectrum of **2** showed positive Cotton effects at 211 and 268 nm and negative Cotton effect at 252 nm, which were very similar with those in the CD spectrum of the known compound **6**. By comparing the CD curves of compounds **2** and **6** (Figure 3), it could be determined that compounds **2** and **6** ought to share the same absolute configuration. Therefore, the absolute configuration of compound **2** was designed as 4*S*,5*R*,7*R*,11*R*,12*S* and given the trivial name phomolide B.

Compound **3** was isolated as a yellow oil. The molecular formula of **3** was determined to be $C_{15}H_{24}O_2$ by the HRESIMS analysis, indicating four degrees of hydrogen deficiency. Compound **3** exhibited obvious absorption bands at 3,363 and 1,024 cm^{-1} in the IR spectrum, which indicated the presence of hydroxyl group and ether bond. The ^{1}H NMR data of **3**, as shown in Table 3, illustrated two singlet methyl functional groups (δ_H 0.95 and 1.33) and a doublet methyl group (δ_H 0.82, d, J = 6.6 Hz). The ^{13}C NMR (Table 3), supported with the HSQC of **3**, indicated the presence of 15 carbon atoms attributed to three methyl groups (δ_C 15.7, 20.3, and 20.8), five methylene groups (δ_C 29.8, 30.9, 32.1, 32.2, and 78.6), four methine groups (δ_C 38.0, 45.1, 75.1, and 115.6), and three quaternary carbons (δ_C 39.3, 81.8, and 153.7). All the aforementioned conclusive information collectively indicated that compound **3** is a sesquiterpene derivative.

The ^{1}H–^{1}H COSY spectrum of **3** which displayed two consecutive correlations of H_2-1/H_2-2/H_2-3/H-4/H_3-14 and H_2-6/H-7/H-8/H-9 successfully suggested the presence of two independent substructures **a** (C-1/C-2/C-3/C-4/C-14) and **b** (C-6/C-7/C-8/C-9). The further comparison of the 1D NMR spectroscopic data (Table 2) with those of the known compound cyclodebneyol (Burden et al., 1986) tentatively revealed that compound **3** shared the same planar structure as that of the previously reported natural product cyclodebneyol. Moreover, the key HMBC correlations from H_3-14 to C-3 (δ_C 30.9) and C-5 (δ_C 39.3); H_3-15 to C-6 (δ_C 32.2), C-4 (δ_C 38.0), and C-10 (δ_C 153.7); and H_3-13 to C-7 (δ_C 45.1) and C-12 (δ_C 78.6), combined with the COSY fragments **a** and **b**, further confirmed the aforementioned conclusion (Figure 1).

The obvious differences in the chemical shifts of both ^{1}H and ^{13}C NMR data between **3** and cyclodebneyol suggested that these two compounds ought to be a pair of closely related stereoisomers. Moreover, the cross-peaks of H_3-14/H-6β, H_3-13/H-6β, and H_3-15/H-6β in the NOESY spectrum were clearly distinguished; thus, it could be readily speculated that the three methyls H_3-13, H_3-14, and H_3-15 in **3** directed on the same side in its 6/6/5 fused ring skeleton and assumed as β-oriented. However, the proton chemical shift of H-7 was heavily overlapped with H-2β, so the NOESY correlations of H-7/H-4 could not conclusively determine the orientation of protons H-7 and H-4 to further completely confirm the final relative configuration of **3**.

In order to absolutely determine the relative configuration of C-7 for **3**, the gauge-independent atomic orbital (GIAO) density functional theory (DFT) ^{13}C NMR calculations (McWeeny, 1961; Ditchfield, 1972) towards the structures **3a** and **3b** were

TABLE 2 Calculated ^{13}C chemical shifts ($CDCl_3$) fitting with the experimental data of compounds 3a, 3b, 4a, and 4b following the STS protocol.

Exptl.	3				Exptl.	4			
	3a	Dev	3b	Dev		4a	Dev	4b	Dev
32.1	32.07	0.03	32.87	0.67	32.4	32.61	0.21	32.43	0.03
31.43	28.74	1.06	27.08	4.22	29.9	26.00	3.90	28.66	1.24
25.58	29.32	1.58	29.77	2.61	30.9	29.53	1.37	29.54	1.36
28.29	37.68	0.32	43.67	4.34	38.5	42.78	4.28	39.90	1.40
42.34	40.00	0.70	42.28	1.64	39.1	38.58	0.52	39.19	0.09
40.94	32.42	0.22	35.00	1.37	30.4	35.41	5.01	31.99	1.59
33.57	47.18	2.08	49.33	2.96	43.4	44.00	0.60	44.03	0.63
48.06	74.93	0.17	78.33	2.26	75.5	75.38	0.12	75.44	0.06
77.36	118.92	3.32	123.75	7.66	115.6	121.20	5.60	119.66	4.06
123.26	152.23	1.47	146.34	7.61	154.6	150.01	4.59	152.02	2.58
146.09	81.55	0.25	78.38	4.39	79.4	79.36	0.04	79.00	0.40
77.41	75.86	2.74	81.08	1.54	78.6	76.83	1.77	77.44	1.16
80.14	21.44	0.64	24.16	1.82	27.7	26.34	1.36	26.15	1.55
22.62	15.88	0.18	15.93	1.39	15.8	16.64	0.54	16.11	0.31
14.31	20.77	0.47	19.19	2.69	20.3	17.74	2.56	20.56	0.26
17.61	MAE[a]	1.02	MAE[a]	3.14		MAE[a]	2.16	MAE[a]	1.11
	RMS[b]	1.42	RMS[b]	3.77		RMS[b]	2.89	RMS[b]	1.54
	P_{mean}	32.34%	P_{mean}	0.58%		P_{mean}	3.12%	P_{mean}	31.18%
	P_{rel}	100.00%	P_{rel}	0.00%		P_{rel}	0.00%	P_{rel}	100.00%

[a]Absolute error.
[b]Root mean square.

performed at the ωB97x-D/6-31G* (Chai and Head-Gordon, 2008) (IEFPCM, $CDCl_3$) level, and the calculation data were then compared with their experimental ^{13}C NMR data following the reported sorted training set (STS) protocol (Li et al., 2020). According to the linear regression analysis of ^{13}C NMR chemical shifts, the values of the correlation coefficient (R^2) were 0.9989 for **3a** and 0.9800 for **3b** (Figure 4). Moreover, the resulting P_{rel} value of **3a** is 100%, and the mean absolute error (MAE), root mean square error (RMSE), and P_{mean} values of **3a** showed that the calculated ^{13}C NMR data match the experimental data very

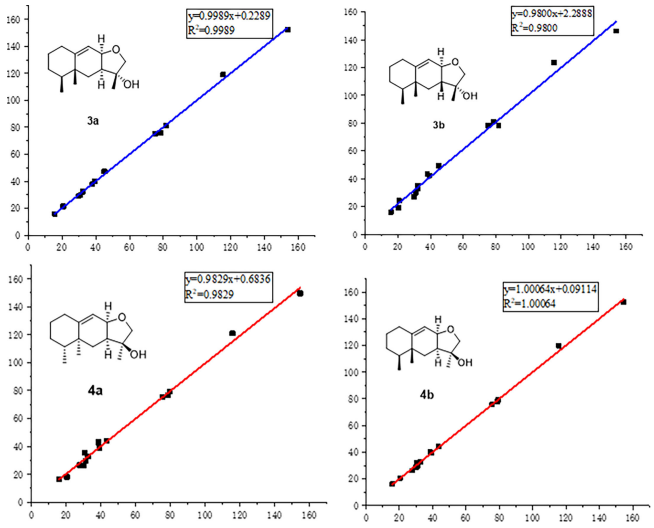


FIGURE 4
Regression analyses of experimental and calculated ^{13}C NMR chemical shifts for **3a**, **3b**, **4a**, and **4b**.

well, indicating that **3a** or its enantiomer is the correct structure for **3** (Table 3). With the aforementioned informative results, the relative structure of **3** was thus unambiguously established as shown in Figure 2.

Moreover, the absolute configuration of **3** was also determined by the time-dependent density-functional theory (TDDFT) calculated circular dichroism (CD) spectrum at the mPW1PW91/SVP level. As shown in Figure 5, the calculated ECD curve of 4*S*,5*R*,7*S*,8*R*,11*R*-**3** perfectly matched with the experimental ECD curve, which strongly suggested that compound 3 shared an absolute configuration of 4*S*,5*R*,7*S*,8*R*,11*R*. Therefore, the structure of compound **3** was completely established and given the trivial name phomene A.

Compound **4** was obtained as a yellow oil with the same molecular formula as **3,** which was determined by HRESIMS ion peak at *m*/*z* 237.1854. Obviously, the ^{13}C NMR spectroscopic data (Table 2) and HSQC spectrum of **4** collectively suggested 15 carbon signals, and all of them showed very similar chemical shifts to those of **3**. The little differences between the chemical shifts of **3** and **4** in NMR spectra strongly implied that they should be a pair of diastereoisomers sharing the same planar structure. The further careful analysis of the NOESY spectrum of **4** could efficiently establish its relative configuration. Compared with the NOESY spectrum of **3,** a clear NOESY correlation of H-8/H_3-13 could be readily found, illustrating that the critical protons H-8 and H_3-13 were on the same side in the fused ring system and assumed as α-oriented. In addition, the cross-peaks of H-2α/H-4 and H-1β/H_3-14 in the NOESY spectrum clearly demonstrated that the two methyls H_3-14 and H_3-15 were β-oriented. Therefore, the relative structure of **4** was tentatively assigned as a C-11 epimer of **3,** as shown in Figure 2, although the same stereochemical issue clouded the H-7 chirality as that of **3**.

In order to further confirm the relative configuration of **4,** we also carried out a ^{13}C NMR calculation for **3.** As a result, the P_{mean} and P_{rel} parameters as well as MAE and RMS values further showed that **4b** or its enantiomer should be the correct structure for **4**, as shown in Table 3. Moreover, the absolute configuration of **4** was determined to be 4*S*,5*R*,7*S*,8*R*,11*S* based on the experimental ECD spectrum, which was highly similar to the calculated ECD spectrum (Figure 5). Thus, the absolute configuration of compound **4** was fully confirmed and given the trivial name phomene B.

Compound **5** was isolated as yellow crystals. Its molecular formula of $C_{15}H_{26}O_2$ was deduced by the HRESIMS spectrum with a protonated ion peak discovered at *m*/*z* 239.2004 $[M + H]^+$ (calcd for $C_{15}H_{27}O_2$, 239.2006), indicating three degrees of hydrogen deficiency. The IR spectrum of **5** revealed an obvious absorption band at 3,315 cm^{-1}, indicating the presence of a series of free hydroxyl functionalities. The ^{1}H NMR data

TABLE 3 ^{1}H (500 MHz) and ^{13}C NMR (125 MHz) data of 3–5 in $CDCl_3$ (δ in ppm, *J* in Hz).

No.	3		4		5	
	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C
1α	2.00, d, (12.1)	32.1, CH_2	2.01, d, (11.9)	32.4, CH_2	1.99, m	53.1, CH
1β	2.29, td, (4.3, 12.1)		2.31, td, (4.4, 11.9)			
2α	1.25, m	29.8, CH_2	1.26, m	29.9, CH_2		74.0, C
2β	1.85, m		1.88, m			
3α	1.48, m	30.9, CH_2	1.49, m	30.9, CH_2	1.31, m	46.3, CH_2
3β					1.94, m	
4	1.54, m	38.0, CH	1.53, m	38.5, CH	1.56, m	21.5, CH_2
5		39.3, C		39.1, C	2.14, ddd, (2.4, 9.4, 15.9)	49.6, CH
6α	0.94, m	32.2, CH_2	1.91, m	30.4, CH_2		36.7, C
6β	1.76, m					
7α	1.89, m	45.1, CH	1.88, m	43.4, CH	1.43, m	29.2, CH_2
7β					1.74, m	
8	4.50, t, (5.3)	75.1, CH	4.18, t, (5.2)	75.5, CH	2.61, dd, (9.4, 15.9)	41.9, CH
9	5.55, d, (5.3)	115.6, CH	5.50, d, (5.2)	115.6, CH		42.4, C
10α		153.7, C		154.6, C	1.54, m	37.6, CH_2
10β					1. 49, m	
11α		81.8, C		79.4, C	1.79, m	18.6, CH_2
11β					1.73, m	
12α	3.73, d, (9.6)	78.6, CH_2	3.66, d, (8.9)	78.6, CH_2	1.24, s	23.4, CH_3
12β	3.87, d, (9.6)		3.84, d, (8.9)			
13	1.33, s	20.8, CH_3	1.46, s	27.7, CH_3	1.22, s	22.4, CH_3
14	0.82, d (6.6)	15.7, CH_3	0.84, d, (6.4)	15.8, CH_3	1.22, s	25.4, CH_3
15a	0.95, s	20.3, CH_3	1.01, s	20.3, CH_3	3.37, d, (10.8)	68.2, CH_2
15b					3.54, d, (10.8)	

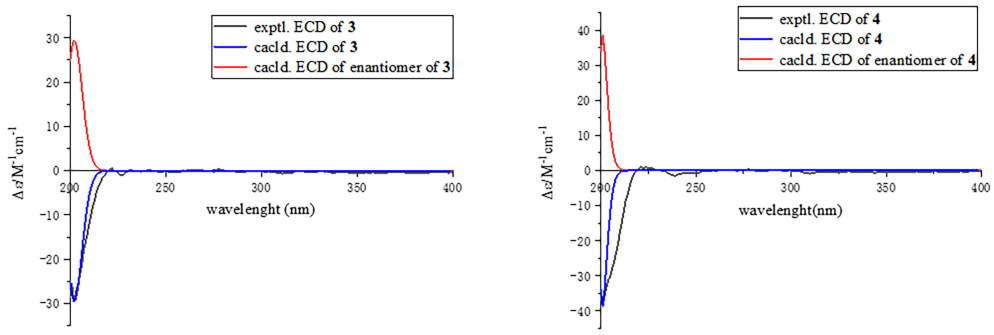


FIGURE 5
Experimental and calculated electronic circular dichroism spectra of compounds **3** and **4**.

(Table 3) of 5 showed three singlet methyl groups (δ_H 1.24, 1.24, and 1.26), a hydroxymethyl moiety (δ_H 3.37, 3.54), and various kinds of saturated aliphatic protons ranging from δ_H 1.22 to 2.61. According to the ^{13}C NMR (Table 3) and HSQC data of compound **5,** 15 carbon signals were resolved, namely, three methyl moieties (δ_C 22.2, 23.4, and 25.4), six methylene groups (δ_C 18.6, 21.3, 29.2, 37.6, 46.3, and 68.2), and three methine functionalities (δ_C 41.8, 49.3, and 53.1) together with three quaternary carbons (δ_C 36.6, 42.3, and 74.0). With careful consideration of the molecular formula of **5**, this informative data strongly indicated that compound **5** might be a tricyclic sesquiterpenoid derivative.

In the ^{1}H–^{1}H COSY spectrum, the obvious correlations of H-1/H_2-11/H_2-10 and H_2-3/H_2-4/H-5/H-8/H_2-7 suggested the presence of two independent spin fragments **a** (C-1/C-11/C-10) and **b** (C-3/C-4/C-5/C-8/C-7). With reference to fragment **a**, the HMBC correlations from H_3-13 to C-1 (δ_C 53.1), C-9 (δ_C 42.4), and C-10 (δ_C 37.6) evidently confirmed the presence of a cyclobutane ring (ring A). Meanwhile, the HMBC correlations from H_3-14 to C-5 (δ_C 49.6), C-6 (δ_C 36.7), C-7 (δ_C 29.2), and C-15 (δ_C 68.2), H_2-15 to C-5, C-6, and C-7 coupling with the COSY fragment C-5/C-6/C-7 further established the other cyclobutane ring (ring C) with a hydroxymethyl functionality attached at the C-6 position. The seven-membered ring B could be conveniently constructed by the HMBC correlations from H-1 to C-3 (δ_C 21.5) and C-8 (δ_C 41.9), H-3 to C-5, H-5 to C-9, H_3-12 to C-1, C-2 (δ_C 74.0), and C-3 on the basis of the COSY fragment C-3/C-4/C-5/C-8. Lastly, the key HMBC correlations of H_2-11 to C-2, H-7 to C-9, and H_2-4 to C-6 could further establish the connection of the 4/7/4 fused ring system (rings A–C). Therefore, the planar structure of **5** was then determined. The relative and absolute configurations of compound **5** were unambiguously confirmed by the X-ray single crystal diffraction on the CuKα with a Flack parameter of -0.05 (12). Finally, the absolute configuration was thus designated as 1*R*,2*R*,5*R*,6*S*,8*R*,9*R*, as shown in Figure 6, and given the trivial name phomene C.

At this stage, compounds **1–6** were evaluated for antimicrobial activities against the bacteria *Escherichia coli*, *S. aureus*, and methicillin-resistant *S. aureus* (MRSA) (Table 4). The biological screening results showed that the new compound **2** showed very potent antimicrobial activities to Gram-positive stain *S. aureus* and MRSA with a minimum inhibitory concentration (MIC) value at 6.25 μg/ml for both. In addition, the known compound **6** also exhibited significant antimicrobial activities, with a MIC value of 1.56 μg/ml towards *S. aureus* and 0.78 μg/ml towards MRSA, respectively, which were very close to those of the positive control vancomycin (0.78 μg/ml for *S. aureus* and 0.78 μg/ml for MRSA). However, we have only tested the bacteriostatic potential of the compounds, and their bactericidal activity is still unknown and deserves further extensive exploration.

Furthermore, compounds **1–6** were also tested for their cytotoxicity against a panel of human cancer cell lines including SF-268, MCF-7, HepG-2, and A549 and normal cell line LX-2 (Table 5). However, all the tested compounds were found to be devoid of antitumor activity even at the concentration of 100 μM. The aforementioned biological screening results collectively pointed that the acetal meroterpenoids **2** and **6** might be severed as promising lead compounds towards antibacterial innovative drug development.

Conclusively, numerous excellent efforts towards the secondary metabolites of the plant endophytic fungi have successfully clarified that the endophytic fungi shared the outstanding ability to produce pharmaceutically meaningful natural products (Debbab et al., 2013; Mousa and Raizada, 2013; Brader et al., 2014) or similar bioactive metabolites as their hosts (Stierle et al., 1993; Kusari et al., 2011; Kusari et al., 2012). In recent years, the extraction and the isolation of bioactive leading natural products, with aim of discovering innovative drugs from the plant endophytic fungi, are emerging as a hot research topic for both natural product and medicinal chemists (Praptiwi et al., 2018; Tanapichatsakul et al., 2018; Adeleke and Babalola, 2020). However, the endophytic

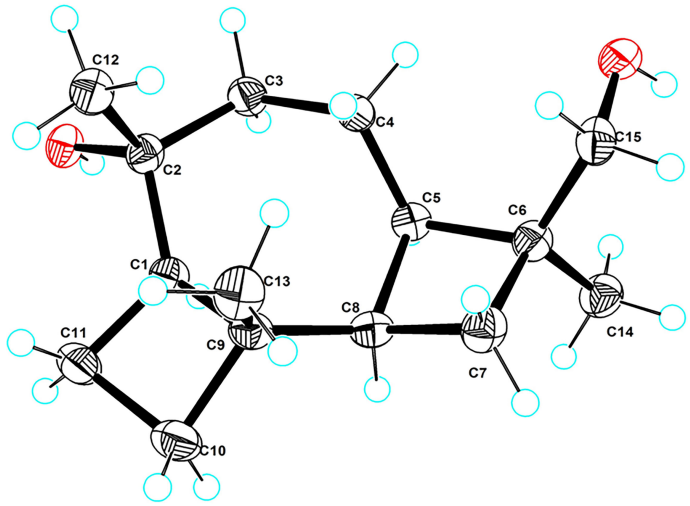


FIGURE 6
ORTEP drawing of the X-ray structure of **5**.

fungi of the *Alpinia shengzhen* plant have not been reported. This study reported the isolation and identification of the endophytic fungi of *Alpinia shengzhen* for the first time (Hu et al., 2011). As a result, six strains of endophytic fungi, namely, *Cladorrhinum* sp. SZSJ-2 (Madrid et al., 2011), *Phyllosticta capitalensis* SZSJ-3 (Arafat, 2018), *Nigrospora oryzae* SZSJ-5 (Hudson, 1963), and two strains of *Phomopsis* sp. SZSJ-7B and 7C (Zhang et al., 2000) together with *Annulohypoxylon stygium* SZSJ-7A (Hsieh et al., 2005), were isolated from the plant tissues of *Alpinia shengzhen* by section culture because of their abundant secondary metabolites and notable biological activities. These results informatively suggested that *Alpinia shengzhen* could be applied as a promising bioresource for the discovery of medicinal fungi.

In this study, a chemical investigation of the endophytic fungus *Phomopsis* sp. SZSJ-7B of *Alpinia shengzhen* was performed for the first time to discover novel lead drug molecules with chemically diverse structures and biologically significant activities. Although previous studies revealed a huge number of intriguing novel natural products isolated from the genus *Phomopsis* (Silva et al., 2006; Yang et al., 2013; Fan et al., 2020; Gong et al., 2020; Yang et al., 2020), five novel secondary metabolites and one known natural product were also successfully isolated from this genus in this time. The six natural products shared four different structure types, all of which were isolated from this genus for the first time, thus greatly enriching the structural types of natural compounds from *Phomopsis* sp. (Udayanga et al., 2011; Xu et al., 2021). Moreover, this chemical research effort can also strongly provide chemo-logical reference and experimental guidance for future studies towards the endophytic fungi and the secondary metabolites of other ginger plants.

TABLE 4 Antibacterial activity of compounds 1–6 (minimum inhibitory concentration, μg/ml).

Compounds	*Escherichia coli*	*Staphylococcus aureus*	Methicillin-resistant *S. aureus*
Vancomycin/ kanamycin	3.12	0.78	0.78
1	>100	100	100
2	>100	6.25	6.25
3	>100	100	100
4	>100	100	100
5	>100	100	100
6	>100	1.56	0.78

The acetal skeleton represents a ubiquitous and intriguing family of structurally special architecture, which is prevalent in numerous biologically meaningful natural products (Pettit et al., 2015) and many pharmaceutically significant clinical drugs, such as aquamox for hypertension (Verdel et al., 2006) and cevimeline for parasympathetic nerves (Mavragani and Moutsopoulos, 2007). In this study, phomolides A and B were characterized with a natural, rarely occurring ester-acetal skeleton, which is constructed by a carboxylic acid and a phenol hydroxyl functionality (Chen et al., 2021) together with an aldehyde fragment (Ahmed Laskar and Younus, 2019), giving rise to a distinctive difference with the common acetal moiety formatted with two alcoholic hydroxyl groups and an aldehyde. To our knowledge, there were only four examples of natural products sharing this skeleton reported in previous literatures (Zhang et al., 2020; Choi et al., 2021). The discovery of phomolides A and B further enriches the

TABLE 5 Cytotoxicity activity of compounds 1–6 (IC_{50}, μM).

Compounds	SF-268	MCF-7	HepG-2	A549	LX-2
Adriamycin	**1.24 ± 0.01**	**1.08 ± 0.04**	**1.06 ± 0.05**	**1.10 ± 0.07**	**1.22 ± 0.03**
1	115.93 ± 2.55	116.09 ± 1.98	120.72 ± 1.59	>128	91.45± 3.19
2	31.82 ± 1.72	57.22 ± 1.03	65.55 ± 3.11	42.34 ± 2.07	27.12 ± 2.53
3	118.08 ± 1.49	121.05 ± 2.16	>128	>128	93.27 ± 1.31
4	>128	>128	>128	>128	>128
5	>128	>128	>128	>128	> 128
6	46.48 ± 3.03	61.52 ± 1.79	39.53 ± 4.00	50.44 ± 6.20	21.08 ± 0.74

structural types and the members of these ester-acetal compounds. Moreover, the biological evaluations showed that phomolide B exhibited a significant inhibitory activity against *S. aureus* and MRSA, which suggested that these ester-acetal compounds might continuously serve as potent innovative impetus for the further extensive research and medicinal exploration towards the development of novel antibacterial drugs (Rossiter et al., 2017; Wu et al., 2019; Dandawate et al., 2019; Dai et al., 2020).

Phomenes A and B possess a typical eremophilane-type sesquiterpene skeleton with a fascinating 6/6/5 fused ring system. The first example of this eremophilane sesquiterpene, named cyclodebney, was reported in 1986, which was successfully isolated from tobacco necrosis virus (TNV) *Niwtiutua dcbneyi* and then illustrated to show a potent antifungal activity. Furthermore, Le isolated three other new eremophilane-type sesquiterpenes from *Sarcographa tricosa* (Le et al., 2013) as characteristic secondary metabolites. Until now, there were about seven eremophilane sesquiterpenes with this 6/6/5 fused ring system isolated from medicinal plants (Wang et al., 2016; Shao et al., 2016) and fungi (Chang et al., 2017). However, phomenes A and B are the first two examples of eremophilane-type sesquiterpenes isolated from *Phomopsis* sp., and they further increase the structural diversity of secondary metabolites in this genus.

Phomene C is a tricyclic sesquiterpenoid derivative with a very intriguing 4/7/4 fused ring scaffold, which is a fascinating type of sesquiterpenoid skeleton rather rarely occurring in nature. Up to now, only two examples of this sesquiterpenoid, named koraiol and frabenol, have been previously isolated from the oleoresin of *Pinus koraiensis* and *Fimetariella rabenhorstii*, respectively (Khan et al., 1979; Tao et al., 2011). Moreover, the sesquiterpene alcohol 5,8-cyclocaryophyllan-4-ol, which was detected in Cangerana oil (Weyerstahl et al., 1996), also shared a closely similar carbon skeleton as that of koraiol and frabenol with a 4/7/4 fused ring system. In this study, the discovery of the new sesquiterpenoid phomene C further broadened the structural diversity and enriched the family members of this type of sesquiterpenoid.

3 Conclusion

In conclusion, this study firstly conducted a systematic chemical investigation on the secondary metabolites of the endophytic fungi *Phomopsis* sp. SZSJ-7B from *Alpinia shengzhen* and successfully resulted in the isolation of two novel acetal lactones (phomolides A and B), three undescribed sesquiterpenes (phomenes A–C), and a known lactone (colletotricholide A). All of these types of compounds were also isolated from *Phomopsis* sp. for the first time, which greatly enriched the structural diversity of the secondary metabolites of the genus. The structures of the new compounds were fully characterized by a combination of spectroscopic methods, X-ray diffraction, and quantum chemistry calculations. The biological activity screening clarified that both compounds **2** and **6** exhibited significant antibacterial activities towards MRSA and *S. aureus* strains with MIC values as low as 6.25 μg/ml, which were comparable to those of the positive control vancomycin without any notable cytotoxicity, thus illustrating its significant potential in the development of innovative antibacterial drugs. Moreover, further investigations on structure–activity relationship and antibacterial mechanism directed toward this goal are currently underway and will be reported in due course.

4 Experimental

4.1 General experimental procedures

IR data were measured on a Shimadzu IR Affinity-1 spectrometer (Shimadzu, Kyoto, Japan). UV and optical rotation data were obtained by a Shimadzu UV-2600 spectrophotometer (Shimadzu, Kyoto, Japan) and an Anton Paar MCP-500 spectropolarimeter (Anton Paar, Graz, Austria). The ECD spectra were measured with Applied Photophysis Chirascan. The NMR spectra (1D and 2D) data were collected on a Bruker Avance-500 spectrometer with tetramethylsilane as an internal standard (Bruker, Fällanden,

Switzerland). The HRESIMS spectra were acquired with a Thermo MAT95XP high-resolution mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). The single crystal data were collected on an Agilent Xcalibur Novasingle-crystal diffractometer equipped with CuKα radiation. A Hitachi Primaide [Hitachi Instruments (Dalian) Co., Ltd.] equipped with a diode array detector using a preparative YMC ODS C_{18} column (20 × 250 mm, 5 μm) was used for preparative HPLC separation. Sephadex LH-20 (GE Healthcare, Uppsala, Sweden), silica gel (200–300 and 60–100 mesh, Puke., Qingdao, China), and C_{18} reversed-phase silica gel (40-75 μm, Fuji, Kasugai, Japan) were used for column chromatography. All solvents were of analytical grade (Guangzhou Chemical Regents Company, Ltd., Guangzhou, China).

4.2 Fungal material

The fungal strain *Phomopsis* sp. SZSJ-7B was isolated from the fresh leaves of *Alpinia shengzhen* collected in the South China Botanical Garden in Guangzhou City, Guangdong Province of China in September 2020. Using BLAST to search the GenBank database, SZSJ-7B (GenBank accession number: OP623444.1) has 100% similarity with *Phomopsis* sp. MJ53 (GenBank accession number: KM203620.1). The strain is preserved at the Key Laboratory of South China Agricultural Plant Molecular Analysis and Genetic Improvement, South China Botanical Garden in Guangzhou City.

4.3 Fermentation, extraction, and isolation

The prepared fresh mycelium of the strain was inoculated into each of five 500-ml Erlenmeyer flasks containing 200 ml PDB medium (200 g potato, 20 g dextrose, 3 g KH_2PO_4, 1.5 g $MgSO_4$, and 10 mg vitamin B in 1 L H_2O) and then incubated at 28°C on a rotary shaker at 180 rpm for 5 days to obtain the seed culture. Fermentation was performed in 30 3-L Fernbach flasks, each containing 1.5 L PDB medium. After having been disinfected at 121°C for 30 min in an autoclave and cooled to room temperature, each flask was inoculated with 30 ml of the seed cultures and incubated at 28°C for 30 days. After cultivation, the mycelia were extracted with EtOAc for three times, and the crude extract (10 g) was obtained. The crude extract was subjected to silica gel using gradient elution with petroleum ether– EtOAc–methanol (MeOH) (v/v/v, 50:1:0→0:10:1) to afford six main fractions (Fr.1–Fr.6).

Fr.1 (880 mg) was isolated on silica gel and eluted with ether–EtOAc gradient (v/v, 100:0→2:1) to obtain six sub-fractions (Fr.1-1 to Fr.1-6). Fr.1-5 (154 mg) was eluted isocratically with ether–EtOAc (10:1) to afford compound **6** (5 mg).

Fr.2 (697 mg) was isolated on silica gel and eluted with ether–EtOAc gradient (v/v, 100:0→1:1) to obtain seven sub-fractions (Fr.2-1 to Fr.2-7). Fr.2-6 (217 mg) was isolated on silica gel and eluted with ether–EtOAc gradient (v/v, 20:1→2:1) to obtain five sub-fractions (Fr.2-6-1 to Fr.2-6-5). Fr.2-6-3 (111 mg) was isolated on silica gel and eluted with ether–CH_3Cl gradient (v/v, 5:1→2:1) to obtain four sub-fractions (Fr.2-6-3-1 to Fr.2-6-3-4). Fr.2-6-3-4 (20 mg) was isolated on silica gel and eluted with ether–EtOAc gradient (v/v, 15:1→5:1) to obtain two sub-fractions (Fr.2-6-3-4-1 to Fr.2-6-3-4-2). Fr.2-6-3-4-2 (10 mg) was further purified by the preparative HPLC system with CH_3CN–H_2O (80:20) as eluent to afford compound **1** (3.8 mg, t_R = 7.0 min) and compound **2** (2.8 mg, t_R = 8.0 min).

Fr.3 (1.8 g) was separated by Sephadex LH-20 CC eluted with $CHCl_3$–MeOH (v/v, 1:3) to afford three sub-fractions (Fr.3-1 to Fr.3-3). Fr.3-2 (205 mg) was isolated on silica gel and eluted with ether–EtOAc gradient (v/v, 100:0→1:1) to obtain seven sub-fractions (Fr.3-2-1 to Fr.3-2-7). Fr.3-2-6 (28.4 mg) was isolated on silica gel and eluted with ether–$CHCl_3$ gradient (v/v, 10:1→1:1) to obtain compound **4** (3.6 mg).

Fr.4 (857 mg) was separated into nine subfractions (Fr.4-1 to Fr.4-9) on octadecyl-silylated silica gel column chromatography (ODS CC) with MeOH–H_2O (v/v, 50:50→100:0). Fr.4-4 (109 mg) was isolated on silica gel and eluted with ether–EtOAc gradient (v/v, 100:0→1:1) to obtain compound **3** (3.4 mg).

Fr.5 (322 mg) was separated into five subfractions (Fr.5-1 to Fr.5-5) on ODS CC with MeOH–H_2O (v/v, 30:70→100:0). Fr.5-3 (13.8 mg) was isolated on silica gel and eluted with ether–EtOAc gradient (v/v, 5:1→0:1) to obtain three sub-fractions (Fr.5-3-1 to Fr.5-3-3). Fr.5-3-2 (10.0 mg) was eluted isocratically with $CHCl_3$–MeOH (50:1) to afford compound **5** (4.8 mg).

Phomolide A: white amorphous powder; $[\alpha]_D^{25}$ + 0.07 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ϵ): 214 (3.38), 242 (2.56), 243 (2.73), and 282 (2.97) nm; IR (KBr): 3,325, 2,943, 2,833, 1,662, 1,448, 1,022, 970, and 667 cm^{-1}; HRESIMS: *m/z* 209.0815 $[M + H]^+$ (calcd for $C_{11}H_{13}O_4$, 209.0808); ^1H (500 MHz) and ^{13}C (125 MHz) NMR data (see Table 1).

Phomolide B: white amorphous powder; $[\alpha]_D^{25}$ + 4.16 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ϵ): 240 (2.78) and 266 (3.30) nm; IR (KBr): 3,360, 2,933, 2,833, 1,714, 1,680, 1,585, 1,456, 1,166, 1,020, and 669 cm^{-1}; HRESIMS: *m/z* 385.2007 $[M + H]^+$ (calcd for $C_{23}H_{29}O_5$, 385.2010); ^1H (500 MHz) and ^{13}C (125 MHz) NMR data (see Table 1).

Phomene A: yellow oil; $[\alpha]_D^{25}$ – 2.17 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ϵ): 200 (3.30) nm; IR (KBr): 3,363, 2,927, 2,858, 1,739, 1,653, 1,516, 1,454, 1,377, 1,251, 1,024, 923, 675, and 597 cm^{-1}; HRESIMS: *m/z* 237.1851 $[M + H]^+$ (calcd for $C_{15}H_{25}O_2$, 237.1849); ^1H (500 MHz) and ^{13}C (125 MHz) NMR data (see Table 3).

Phomene B: yellow oil; $[\alpha]_D^{25}$ – 4.21 (*c* 0.1, MeOH); UV (MeOH): λ_{max} (log ϵ): 200 (3.25) nm; IR (KBr): 3,361, 2,929, 2,858, 1,739, 1,653, 1,541, 1,516, 1,454, 1,379, 1,024, 952, 667, and 597 cm^{-1}; HRESIMS: *m/z* 237.1854 $[M + H]^+$ (calcd for $C_{15}H_{25}O_2$, 237.1849); ^1H (500 MHz) and ^{13}C (125 MHz) NMR data (see Table 3).

Phomene C: yellow crystal; $[\alpha]_D^{25}$ + 0.063 (*c* 0.1, MeOH); UV (MeOH): λ_{max} (log ϵ): 223 (1.60) and 237 (1.67) nm; IR (KBr): 3,315, 2,947, 2,858, 1,651, 1,375, 1,112, 1,029, 912, 665, and 603 cm^{-1}; HRESIMS: *m/z* 239.2004 $[M + H]^+$ (calcd for $C_{15}H_{27}O_2$, 239.2006); ^1H (500 MHz) and ^{13}C (125 MHz) NMR data (see Table 3).

4.4 Quantum chemistry calculations

Conformational search of structures was performed by Crest (Pracht et al., 2020), with 4 kcal/mol energy window. Optimization and frequency calculation of the obtained conformer were performed on B3LYP/TZVP (Grimme et al., 2011; Tsuzuki and Uchimaru, 2020) (IEFPCM, $CDCl_3$ and MeOH) level of theory. DFT GIAO ^{13}C NMR calculation was calculated on the ωB97xD/6-31G* (IEFPCM, $CDCl_3$) level, and the data processing followed the reported STS protocol. The calculated shielding tensors of conformers were Boltzmann-averaged based on Gibbs free energy. Theoretical ECD (TDDFT) calculation was calculated on mPW1PW91/TZVP (IEFPCM, MeOH) level. SpecDis v1.71 was used to simulate the ECD curve with sigma/gamma value of 0.35 eV (Bruhn et al., 2013). The calculated ECD curve of each conformer was Boltzmann-averaged based on their Gibbs free energy. The average calculated ECD curve of **1** was adjusted by blue shifting for 20 nm. All DFT calculations were performed by Gaussian 16 software package (Frisch et al., 2016).

4.5 X-ray crystallographic data

Crystal data for **6** $C_{15}H_{26}O_2$ (*M* = 238.36 g/mol): trigonal, space group $P3_2$ (no. 145), *a* = 13.6894 (2) Å, *c* = 6.60290 (10) Å, *V* = 1,071.60 (4) $Å^3$, *Z* = 3, *T* = 100.00 (10) K, μ(CuKα) = 0.552 mm^{-1}, Dcalc = 1.108 g/cm^3; 7,072 reflections measured (7.456 ≤ 2Θ ≤ 148.49) and 2,779 unique (R_{int} = 0.0298, R_{sigma} = 0.0367), which were used in all calculations. The final R_1 was 0.0422 [I > 2σ(I)] and wR_2 was 0.1085 (all data). Flack parameter = -0.05 (12). The crystallographic data for **5** reported in this paper has been deposited in the Cambridge Crystallographic Data Centre (deposition number: CCDC 2192981). Copies of these data can be obtained free of charge *via* https://www.ccdc.cam.ac.uk.

4.6 Cytotoxicity and antimicrobial assays

4.6.1 Cytotoxicity assays

Cytotoxicity was evaluated by the sulforhodamine B assay (Skehan et al., 1990) against five human cancer cell lines (SF-268, MCF-7, HepG2, A549, and LX-2). As a result, none of the compounds showed good cytotoxic activity.

4.6.2 Antimicrobial assays

The antibacterial activities for compounds **1**–**6** were evaluated against three bacteria embodying *S. aureus* (CMCC 26003), methicillin-resistant *S. aureus* (JCSC 3063), and *E. coli* (ATCC 8739). All of the bacteria were purchased from Guangdong Institute of Microbiology (Guangzhou, China). The MICs were determined by the broth microdilution method in 96-well plates as described in previous literature (NCCLS, 1999; CLSI, 2012; Li et al., 2014). Positive control was vancomycin or kanamycin. All test samples were dissolved in dimethyl sulfoxide and diluted with culture medium.

Data availability statement

The data presented in the study are deposited in the Cambridge Crystallographic Data Centre, accession number: CCDC 2192981.

Author contributions

YC, HW, and XK conducted the experiment and collected the experimental data. YC performed the experiments of compound isolation. ZS and ZZ carried out the ECD calculations. YC, ZZ, and HT finished the structure identification of the isolated compounds. MK, WP, JT, and YZ evaluated the activities of all the isolates. YC and HT interpreted the data and wrote the paper. YC, ZZ, and HT revised the manuscript. ZZ and HT conceived and designed the experiments. All authors contributed to the article and approved the submitted version.

Funding

Financial support for this research was provided by the National Natural Science Foundation of China (82173711), Youth Innovation Promotion Association of CAS (2020342), Natural Science Foundation of Guangdong Province (2019A1515011694), Natural Science Foundation of Hunan Province (no. 2021JJ30917), High-tech Industry Science and Technology Innovation Project of Hunan Province (2020GK4083), Postgraduate Research and Innovation Project of Hunan Province (CX20210341), Postgraduates Innovation Program of Central South University (Nos. 2021zzts0978, 2021zzts0994, and 2022zzts0899), and the Open Sharing Fund for the Large-Scale Instruments and Equipment of Central South University.

Acknowledgments

We sincerely thank Ms. Xuan Ma of South China Sea Institute of Oceanology for the X-ray measurements.

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.2022.1049015/full#supplementary-material

References

Adeleke, B. S., and Babalola, O. O. (2020). Oilseed crop sunflower (*Helianthus annuus*) as a source of food: nutritional and health benefits. *Food. Sci. Nutr.* 8, 4666–4684. doi: 10.1002/fsn3.1783

Ahmed Laskar, A., and Younus, H. (2019). Aldehyde toxicity and metabolism: the role of aldehyde dehydrogenases in detoxification, drug resistance and carcinogenesis. *Drug Metab. Rev.* 51, 42–64. doi: 10.1080/03602532.2018.1555587

Arafat, K. (2018). A novel isolate of *Phyllosticta capitalensis* causes black spot disease on guava fruit in Egypt. *Asian. J. Plant Pathol.* 12, 27–37. doi: 10.3923/ajppaj.2018.27.37

Brader, G., Compant, S., Mitter, B., Trognitz, F., and Sessitsch, A. (2014). Metabolic potential of endophytic bacteria. *Curr. Opin. Biotechnol.* 27, 30–37. doi: 10.1016/j.copbio.2013.09.012

Bruhn, T., Schaumlöffel, A., Hemberger, Y., and Bringmann., G. (2013). SpecDis: quantifying the comparison of calculated and experimental electronic circular dichroism spectra. *Chirality* 25, 243–249. doi: 10.1002/chir.22138

Bunyapaiboonsri, T., Yoiprommarat, S., Srikitikulchai, P., Srichomthong, K., and Lumyong, S. (2010). Oblongolides from the endophytic fungus phomopsis sp. BCC 9789. *J. Nat. Prod.* 73, 55–59. doi: 10.1021/np900650c

Burden, R. S., Loeffler, R. S. T., Rowell, P. M., Bailey, J. A., and Kemp, M. S. (1986). Cyclodebneyol, a fungi toxic sesquiterpene from TNV infected *Nicotiana debneyi*. *Phytochem* 25, 1607–1608. doi: 10.1016/s0031-9422(00)81217-0

Cai, R. L., Chen, S. H., Liu, Z. M., and Tan, C. B. (2017). A new alpha-pyrone from the mangrove endophytic fungus phomopsis sp. HNY29-2B. *Nat. Prod. Res.* 31, 124–130. doi: 10.1080/14786419.2016.1214833

Chai, J. D., and Head-Gordon, M. (2008). Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* 10, 6615–6620. doi: 10.1039/B810189B

Chang, J. C., Hsiao, G., Lin, R. K., Kuo, Y. H., Ju, Y. M., and Lee, T. H. (2017). Bioactive constituents from the termite nest-derived medicinal fungus *Xylaria nigripes*. *J. Nat. Prod.* 80, 38–44. doi: 10.1021/acs.jnatprod.6b00249

Chen, H. P., Huang, M. X., Li, X. W., Chen, B., Wang, J., Lin, Y. C., et al. (2018). Phochrodines a–d, first naturally occurring new chromenopyridines from mangrove entophytic fungus phomopsis sp. 33. *Fitoterapia* 124, 103–107. doi: 10.1016/j.fitote.2017.10.013

Chen, S. C., Liu, Z. M., Tan, H. B., Chen, Y. C., Liu, H. X., Zhang, W. M., et al. (2019). Tersone a-G, new pyridone alkaloids from the deep-sea fungus *Phomopsis tersa*. *Mar. Drugs* 17, 394–407. doi: 10.3390/md17070394

Chen, S. C., Liu, Z. M., Tan, H. B., Zhu, S., Liu, H. X., Zhang, W. M., et al. (2020). Photeroids a and b, unique phenol-sesquiterpene meroterpenoids from the deep-sea-derived fungus phomopsis tersa. *Org. Biomol. Chem.* 18, 642–645. doi: 10.1039/c9ob02625h

Chen, X., Li, D., Zhang, H., Duan, Y., and Huang, Y. (2021). Sinomenine-phenolic acid coamorphous drug systems: solubilization, sustained release, and improved physical stability. *Int. J. Pharm.* 598, 120389. doi: 10.1016/j.ijpharm.2021.120389

Choi, B. K., Cho, D. Y., Choi, D. K., and Shin, H. J. (2021). Miharadienes a-d with unique cyclic skeletons from a marine-derived streptomyces miharaensis. *Org. Chem. Front.* 8, 4845–4852. doi: 10.1039/D1QO00773D

CLSI (2012). *Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically. 7th edition* (Wayne, PA: Clinical and Laboratory Standards Institute).

Dai, J., Han, R., Xu, Y., Li, N., Wang, J., and Dan, W. (2020). Recent progress of antibacterial natural products: future antibiotics candidates. *Bioorg. Chem.* 101, 103922. doi: 10.1016/j.bioorg.2020.103922

Dandawate, P., Padhye, S., Schobert, R., and Biersack, B. (2019). Discovery of natural products with metal-binding properties as promising antibacterial agents. *Expert. Opin. Drug Discovery* 14, 563–576. doi: 10.1080/17460441.2019.1593367

Debbab, A., Aly, A. H., and Proksch, P. (2013). Mangrove derived fungal endophytes – a chemical and biological perception. *Fungal. Divers.* 61, 1–27. doi: 10.1007/s13225-013-0243-8

Ding, B., Yuan, J., Huang, X. S., He, L., Tan, H. M., She, Z. G., et al. (2013). New dimeric members of the phomoxanthone family: phomolactonexanthones a, b and deacetylphomoxanthone c isolated from the fungus phomopsis sp. *Mar. Drugs* 11, 4961–4972. doi: 10.3390/md11124961

Ditchfield, R. (1972). Molecular orbital theory of magnetic shielding and magnetic susceptibility. *J. Chem. Phys.* 56, 5688. doi: 10.1063/1.1677088

Fan, M. M., Xiang, G., Chen, J. W., Zhou, L., Jiao, R. H., Shen, Y., et al. (2020). Libertellenone m, a diterpene derived from an endophytic fungus phomopsis sp. S12, protects against DSS-induced colitis *via* inhibiting both nuclear translocation of NF-κB and NLRP3 inflammasome activation. *Int. Immunopharmacol.* 80, 106144. doi: 10.1016/j.intimp.2019.106144

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., and Cheeseman, J. R. (2016). *Gaussian 16, revision C.01* (Wallingford CT: Gaussian, Inc.).

Gong, J. L., Lu, Y., Wu, W. H., Xi, J. G., Tang, S. B., Yi, K. X., et al. (2020). First report of *Phomopsis heveicola* (anamorph of *Diaporthe tulliensis*) causing leaf blight of *Coffee* (*Coffea arabica*) in China. *Plant Dis.* 104, 570–571. doi: 10.1094/PDIS-09-19-1833-PDN

Gouda, S., Das, G., Sen, S. K., Shin, H. S., and Patra, J. K. (2016). Endophytes: a treasure house of bioactive compounds of medicinal importance. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.01538

Grimme, S., Ehrlich, S., and Goerigk, L. (2011). Effect of the damping function in dispersion corrected density functional theory. *J. Energy Chem.* 32, 1456–1465. doi: 10.1002/jcc.21759

Hemtasin, C., Kanokmedhakul, S., Kanokmedhakul, K., Soytong, K., Prabpai, S., Kongsaeree, P., et al. (2011). Cytotoxic pentacyclic and tetracyclic aromatic sesquiterpenes from *Phomopsis archeri*. *J. Nat. Prod.* 74, 609–613. doi: 10.1021/np100632g

Hsieh, H. M., Ju, Y. M., and Rogers, J. D. (2005). Molecular phylogeny of hypoxylon and closely related genera. *Mycologia* 97, 844–865. doi: 10.3852/mycologia.97.4.844

Huang, Z. J., Cai, X. L., Shao, C. S., Yang, J. X., Zhou, S. N., Lin, Y. C., et al. (2008). Chemistry and weak antimicrobial activities of phomopsins produced by mangrove endophytic fungus phomopsis sp. ZSU-H76. *Phytochem* 69, 1604–1608. doi: 10.1016/j.phytochem.2008.02.002

Huang, R., Jiang, B. G., Li, X. N., Zheng, K. X., He, J., Wu, S. H., et al. (2018). Polyoxygenated cyclohexe-noids with promising alpha-glycosidase inhibitory activity produced by phomopsis sp. YE3250, an endophytic fungus derived from *Paeonia delavayi*. *J. Agr. Food. Chem.* 66, 1140–1146. doi: 10.1021/acs.jafc.7b04998

Hudson, H. J. (1963). The perfect state of *Nigrospora oryzae*. *British. Mycological. Soc.* 46, 355–360. doi: 10.1016/S0007-1536(63)80027-3

Hussain, H., Krohn, K., Ahmed, I., Schulz, B., Di Pietro, S., Pescitelli, G., et al. (2012). Phomopsinones a–d: four new pyrenocines from endophytic fungus phomopsis sp. *Eur. J. Org. Chem.* 2012, 1783–1789. doi: 10.1002/ejoc.201101788

Hu, Z. X., Wu, Y., Xie, S. S., Luo, Z. W., Xue, Y. B., Zhang, Y. H., et al. (2017). Phomopsterones a and b, two functionalized ergostane-type steroids from the endophytic fungus phomopsis sp. TJ507A. *Org. Lett.* 19, 258–261. doi: 10.1021/acs.orglett.6b03557

Hu, J. Z., Ye, Y. S., Zou, P., and Liao, J. P. (2011). Studies on the hybrid breeding and biological characteristics of zingiberaceous plant (*Alpinia hainanensis* 'Shengzhen'). *J. Trop. Subtrop. Bot.* 19, 279–282. doi: 10.969/j.issn.1005-3395.2011.03.014

Jouda, J. B., Tamokou, J. D., Mbazoa, C. D., Douala-Meli, C., Sarkar, P., Bag, P. K., et al. (2016). Antibacterial and cytotoxic cytochalasins from the endophytic fungus phomopsis sp. harbored in garcinia kola (Heckel) nut. *BMC. Complement. Altern. Med.* 16, 462–462. doi: 10.1186/s12906-016-1454-9

Khan, V. A., Gatilov, Y. V., Dubovenko, Z. V., and Pentegova, V. A. (1979). Crystal structure of koraiol-a sesquiterpene alcohol with a new type of carbon skeleton from the oleoresin of *Pinus koraiensis*. *Chem. Nat. Compd.* 15, 572–576. doi: 10.1007/bf00565927

Krohn, K., Farooq, U., Hussain, H., Draeger, S., Schulz, B., van Ree, T., et al. (2011). Phomosines h-J, novel highly substituted biaryl ethers, isolated from the endophytic fungus phomopsis sp. from *Ligustrum vulgare*. *Nat. Prod. Commun.* 6, 1907–1912. doi: 10.1177/1934578X1100601229

Kusari, S., Hertweck, C., and Spiteller, M. (2012). Chemical ecology of endophytic fungi: origins of secondary metabolites. *Chem. Biol.* 19, 792–798. doi: 10.1016/j.chembiol.2012.06.004

Kusari, S., Zühlke, S., and Spiteller, M. (2011). Effect of artificial reconstitution of the interaction between the plant *Camptotheca acuminata* and the fungal endophyte *Fusarium solani* on camptothecin biosynthesis. *J. Nat. Prod.* 74, 764–775. doi: 10.1021/np1008398

Le, D. H., Takenaka, Y., Hamada, N., and Tanahashi, T. (2013). Eremophilane-type sesquiterpenes from cultured lichen mycobionts of *Sarcographa tricosa*. *Phytochem* 91, 242–248. doi: 10.1016/j.phytochem.2012.01.009

Li, J., Liu, J. K., and Wang, W. X. (2020). GIAO ^{13}C NMR calculation with sorted training sets improves accuracy and reliability for structural assignation. *J. Org. Chem.* 85, 11350–11358. doi: 10.1021/acs.joc.0c01451

Li, L. Y., Sattler, I., Deng, Z. W., Peschel, G., Grabley, S., Lin, W. H., et al. (2008). A-seco-oleane-type triterpenes from phomopsis sp. (strain HKI0458) isolated from the mangrove plant *Hibiscus tiliaceus*. *Phytochem* 69, 511–517. doi: 10.1016/j.phytochem.2007.08.010

Liu, H. B., Liu, Z. M., Chen, Y. C., Li, D. L., Liu, H. X., Zhang, W. M., et al. (2021). Cytotoxic diaporindene and tenellone derivatives from the fungus *Phomopsis lithocarpus*. *Chin. J. Nat. Med.* 19, 874–880. doi: 10.1016/S1875-5364(21)60095-X

Li, C. R., Zhai, Q. Q., Wang, X. K., Li, G. Q., Zhang, W. X., You, X. F., et al. (2014). *In vivo* antibacterial activity of MRX-I, a new oxazolidinone. *Antimicrob. Agents. Chemother.* 58, 2418–2421. doi: 10.1128/aac.01526-13

Madrid, H., Cano, J., Gené, J., and Guarro, J. (2011). Two new species of *Cladorrhinum*. *Mycologia* 103, 795–805. doi: 10.3852/10-150

Mavragani, C. P., and Moutsopoulos, H. M. (2007). Conventional therapy of sjogren's syndrome. *Clin. Rev. Allergy Immunol.* 32, 284–291. doi: 10.1007/s12016-007-8008-3

McWeeny, R. (1961). Perturbation theory for the fock-dirac density matrix. *Phys. Rev.* 126, 1028–1034. doi: 10.1103/PhysRev.126.1028

Mousa, W. K., and Raizada, M. N. (2013). The diversity of anti-microbial secondary metabolites produced by fungal endophytes: an interdisciplinary perspective. *Front. Microbiol.* 4. doi: 10.3389/fmicb.2013.00065

NCCLS (1999). *Methods for determining bactericidal activity of antimicrobial agents* (Wayne, PA: Approved guideline M26-A National Committee for Clinical Laboratory Standards).

Pavao, G. B., Vinícius, V. P., de Oliveira, A. L., Mara, R. A., Lusânia, M. G. A., Hosana, M. D., et al. (2016). Differential genotoxicity and cytotoxicity of phomoxanthone a isolated from the fungus *Phomopsis longicolla* in HL60 cells and peripheral blood lymphocytes. *Toxicol. In Vitro.* 37, 211–217. doi: 10.1016/j.tiv.2016.08.010

Pettit, G. R., Xu, J. P., Chapuis, J. C., and Melody, N. (2015). The cephalostatins. 24. isolation, structure, and cancer cell growth inhibition of cephalostatin 20. *J. Nat. Prod.* 78, 1446–1450. doi: 10.1021/acs.jnatprod.5b00129

Pracht, P., Bohle, F., and Grimme, S. (2020). Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Phys. Chem. Chem. Phys.* 22, 7169–7192. doi: 10.1039/C9CP06869D

Praptiwi, M. R., Wulansari, D., Fathoni, A., and Agusta, A. (2018). Antibacterial and antioxidant activities of endophytic fungi extracts of medicinal plants from *Central sulawesi*. *J. Appl. Pharm. Sci.* 8, 069–074. doi: 10.7324/JAPS.2018.8811

Rossiter, S. E., Fletcher, M. H., and Wuest, W. M. (2017). Natural products as platforms to overcome antibiotic resistance. *Chem. Rev.* 117, 12415–12474. doi: 10.1021/acs.chemrev.7b00283

Shao, H., Mei, W. L., Kong, F. D., Li, W., Zhu, G. P., Dai, H. F., et al. (2016). Sesquiterpenes of agarwood from *Gyrinops salicifolia*. *Fitoterapia* 113, 182–187. doi: 10.1016/j.fitote.2016.07.015

Silva, G. H., Teles, H. L., Zanardi, L. M., Costa-Neto, C. M., Castro Gamboa, I., Bolzani, V., et al. (2006). Cadinane sesquiterpenoids of *Phomopsis cassiae*, an endophytic fungus associated with *Cassia spectabilis* (Leguminosae). *Phytochem* 67, 1964–1969. doi: 10.1016/j.phytochem.2006.06.004

Skehan, P., Storeng, R., Scudiero, D., Monks, A., Mcmahon, J., Vistica, D., et al. (1990). New colorimetric cytotoxicity assay for anticancer-drug screening. *J. Natl. Cancer. Inst.* 82, 1107–1112. doi: 10.1093/jnci/82.13.1107

Stierle, A., Strobel, G., and Stierle, D. (1993). Taxol and taxane production by taxomyces andreanae, an endophytic fungus of pacific yew. *Science* 260, 214–216. doi: 10.1126/science.8097061

Tanapichatsakul, C., Monggoot, S., Gentekaki, E., and Pripdeevech, P. (2018). Antibacterial and antioxidant metabolites of diaporthe spp. isolated from flowers of *Melodorum fruticosum*. *Curr. Microbiol.* 7, 476–483. doi: 10.1007/s00284-017-1405-9

Tao, M. H., Yan, J., Wei, X. Y., Li, D. L., Zhang, W. M., and Tan, J. W. (2011). A novel sesquiterpene alcohol from *Fimetariella rabenhorstii*, an endophytic fungus of *Aquilaria sinensis*. *Nat. Prod. Commun.* 6, 763–766. doi: 10.1002/mnfr.201100206

Tsuzuki, S., and Uchimaru, T. (2020). Accuracy of intermolecular interaction energies, particularly those of hetero-atom containing molecules obtained by DFT calculations with grimme's D2, D3 and D3BJ dispersion corrections. *Phys. Chem. Chem. Phys.* 22, 22508–22519. doi: 10.1039/D0CP03679J

Udayanga, D., Liu, X., McKenzie, E. H. C., Chukeatirote, E., Bahkali, A. H. A., and Hyde, K. D. (2011). The genus *Phomopsis*: Biology, applications, species concepts and names of common phytopathogens. *Fungal. Divers.* 50, 189–225. doi: 10.1007/s13225-011-0126-9

Verdel, B. M., Souverein, P. C., Egberts, A. C., and Leufkens, H. G. (2006). Difference in risks of allergic reaction to sulfonamide drugs based on chemical structure. *Ann. Pharmacother.* 40, 1040–1046. doi: 10.1016/j.cyto.2007.04.004

Wang, H. N., Dong, W. H., Huang, S. Z., Wang, J., Mei, W. L., Dai, H. F., et al. (2016). Three new sesquiterpenoids from agarwood of *Aquilaria crassna*. *Fitoterapia* 114, 7–11. doi: 10.1016/j.fitote.2016.07.014

Wei, W., Gao, J., Shen, Y., Chu, Y. L., Xu, Q., and Tan, R. X. (2014). Immunosuppressive diterpenes from phomopsis sp. S12. *Eur. J. Org. Chem.* 2014, 5728–5734. doi: 10.1002/ejoc.201402491

Weyerstahl, P., Schneider, S., and Marschall, H. (1996). Constituents of the Brazilian cangerana oil. *Flavour. Fragr. J.* 11, 81–94. doi: 10.1002/(SICI)1099-1026

Wu, S. C., Liu, F., Zhu, K., and Shen, J. Z. (2019). Natural products that target virulence factors in antibiotic-resistant *Staphylococcus aureus*. *J. Agric. Food. Chem.* 67, 13195–13211. doi: 10.1021/acs.jafc.9b05595

Xie, S. S., Wu, Y., Qiao, Y. B., Guo, Y., Wang, J. P., Zhang, Y. H., et al. (2018). Protoilludane, illudalane, and botryane sesquiterpenoids from the endophytic fungus phomopsis sp. TJ507A. *J. Nat. Prod.* 81, 1311–1320. doi: 10.1021/acs.jnatprod.7b00889

Xu, J. L., Liu, Z. M., Chen, Y. C., Tan, H. B., Liu, H. X., Zhang, W. M., et al. (2019a). Lithocarols a-f, six tenellone derivatives from the deep-sea derived fungus *Phomopsis lithocarpus* FS508. *Bioorg. Chem* 87, 728–735. doi: 10.1016/j.bioorg.2019.03.078

Xu, T. C., Lu, Y. H., Wang, J. F., Liu, S. S., Liu, C. S., Wu, S. H., et al. (2021). Bioactive secondary metabolites of the genus *Diaporthe* and anamorph *Phomopsis* from terrestrial and marine habitats and endophytes: 2010–2019. *Microorganisms* 9, 217. doi: 10.3390/microorganisms9020217

Xu, K., Zhang, X., Chen, J. W., Tan, R. X., Jiao, R. H., Ge, H. M., et al. (2019b). Anti-inflammatory diterpenoids from an endophytic fungus phomopsis sp. S12. *Tetrahedron. Lett.* 60, 151045–151045. doi: 10.1016/j.tetlet.2019.151045

Yang, H. Y., Gao, Y. H., Niu, D. Y., Gao, X. M., Du, G., Hu, Q. F., et al. (2013). Xanthone derivatives from the fermentation products of an endophytic fungus phomopsis sp. *Fitoterapia* 91, 189–193. doi: 10.1016/j.fitote.2013.09.004

Yang, Z. J., Zhang, Y. F., Wu, K., Jiang, Z. T., Ge, M., Shao, L., et al. (2020). New azaphilones, phomopsones a-c with biological activities from an endophytic fungus phomopsis sp. CGMCC No.5416. *Fitoterapia* 145, 104573. doi: 10.1016/j.fitote.2020.104573

Yan, B. C., Wang, W. G., Hu, D. B., Sun, X., Kong, L. M., Pu, J. X., et al. (2016). Phomopchalasins a and b, two cytochalasans with polycyclic-fused skeletons from

the endophytic fungus phomopsis sp. shj2. *Org. Lett.* 18, 1108–1111. doi: 10.1021/acs.orglett.6b00214

Yu, B. Z., Zhang, G. H., Du, Z. Z., Zheng, Y. T., Xu, J. C., and Luo, X. D. (2008). Phomoeuphorbins a-d, azaphilones from the fungus *Phomopsis euphorbiae*. *Phytochem* 69, 2523–2526. doi: 10.1016/j.phytochem.2008.07.013

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214. doi: 10.1089/10665270050081478

Zhang, W. G., Wang, M. M., Zhang, S., Xu, K. P., and Tan, H. B. (2020). Eutyscoparols a-G, polyketide derivatives from endophytic fungus *Eutypella scoparia* SCBG-8. *Fitoterapia* 146, 104681. doi: 10.1016/j.fitote.2020.104681

Zhao, W. T., Liu, Q. P., Chen, H. Y., Zhao, W., Gao, Y., and Yang, X. L. (2020). Two novel eremophylane acetophenone conjugates from *Colletotrichum gloeosporioides*, an endophytic fungus in *Salvia miltiorrhiza*. *Fitoterapia* 141, 104474. doi: 10.1016/j.fitote.2020.104474

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 November 2022
DOI 10.3389/fpls.2022.998755

OPEN ACCESS

EDITED BY
Ramesha Thimmappa,
Amity University, India

REVIEWED BY
Teodora Basile,
Research Centre of Viticulture and Oenology (CREA), Italy
Alam Zeb,
University of Malakand, Pakistan
Konstantinos Dimas,
University of Thessaly, Greece

*CORRESPONDENCE
Mihalis I. Panayiotidis
mihalisp@cing.ac.cy

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

RECEIVED 20 July 2022
ACCEPTED 26 October 2022
PUBLISHED 15 November 2022

CITATION
Kyriakou S, Michailidou K, Amery T, Stewart K, Winyard PG, Trafalis DT, Franco R, Pappa A and Panayiotidis MI (2022) Polyphenolics, glucosinolates and isothiocyanates profiling of aerial parts of *Nasturtium officinale* (Watercress). *Front. Plant Sci.* 13:998755. doi: 10.3389/fpls.2022.998755

Polyphenolics, glucosinolates and isothiocyanates profiling of aerial parts of *Nasturtium officinale* (Watercress)

Sotiris Kyriakou[1], Kyriaki Michailidou[2], Tom Amery[3], Kyle Stewart[4], Paul G. Winyard[4], Dimitrios T. Trafalis[5], Rodrigo Franco[6,7], Aglaia Pappa[8] and Mihalis I. Panayiotidis[1]*

[1]Department of Cancer Genetics, Therapeutics & Ultrastructural Pathology, The Cyprus Institute of Neurology & Genetics, Nicosia, Cyprus, [2]Biostatistics Unit, The Cyprus Institute of Neurology & Genetics, Nicosia, Cyprus, [3]The Watercress Company, Dorchester, United Kingdom, [4]Watercress Research Limited, Devon, United Kingdom, [5]Laboratory of Pharmacology, Medical School, National & Kapodistrian University of Athens, Athens, Greece, [6]Redox Biology Centre, University of Nebraska-Lincoln, Lincoln, NE, United States, [7]Department of Veterinary Medicine & Biomedical Sciences, University of Nebraska-Lincoln, Lincoln, NE, United States, [8]Department of Molecular Biology & Genetics, Democritus University of Thrace, Alexandroupolis, Greece

Watercress (*Nasturtium officinale*) is a rich source of secondary metabolites with disease-preventing and/or health-promoting properties. Herein, we have utilized extraction procedures to isolate fractions of polyphenols, glucosinolates and isothiocyanates to determine their identification, and quantification. In doing so, we have utilized reproducible analytical methodologies based on liquid chromatography with tandem mass spectrometry by either positive or negative ion mode. Due to the instability and volatility of isothiocyanates, we followed an ammonia derivatization protocol which converts them into respective ionizable thiourea derivatives. The analytes' content distribution map was created on watercress flowers, leaves and stems. We have demonstrated that watercress contains significantly higher levels of gluconasturtiin, phenethyl isothiocyanate, quercetin-3-*O*-rutinoside and isorhamnetin, among others, with their content decreasing from flowers (82.11 ± 0.63, 273.89 ± 0.88, 1459.30 ± 12.95 and 289.40 ± 1.37 ng/g of dry extract respectively) to leaves (32.25 ± 0.74, 125.02 ± 0.52, 1197.86 ± 4.24 and 196.47 ± 3.65 ng/g of det extract respectively) to stems (9.20 ± 0.11, 64.7 ± 0.9, 41.02 ± 0.18, 65.67 ± 0.84 ng/g of dry extract respectivbely). Pearson's correlation analysis has shown that the content of isothiocyanates doesn't depend only on the bioconversion of individual glucosinolates but also on other glucosinolates of the same group. Overall, we have provided comprehensive analytical data of the major watercress metabolites thereby providing an opportunity to exploit different parts of watercress for potential therapeutic applications.

KEYWORDS

watercress, polyphenolic acids, flavonoids, glucosinolates, isothiocyanates, thiourea derivatives, UPLC-ESI-MS/MS

10.3389/fpls.2022.998755

1 Introduction

Watercress (*Nasturtium officinale*) is a widely and extensively studied perennial cruciferous plant. As a member of the *Brassicaceae* family, it is a rich source of various phytochemicals including polyphenols (phenolic acids, flavonoids and proanthocyanins), pigments (chlorophylls, lycopene and carotenoids), and isothiocyanates (ITCs). In addition to this, watercress is a source of soluble sugars, proteins and vitamins (Zeb, 2015; Ma et al., 2021; Kyriakou et al., 2022a; Kyriakou et al., 2022b).

Glucosinolates (GLs), the precursors of ITCs, belong to of nitrogenous-sulfur- enriched phytochemicals. Most of the time, GLs are overexpressed in cruciferous vegetables and are structurally characterized by the presence of a β-D-thioglucosidic bond which is linked to ***D***-glucose (Barba et al., 2016; Prieto et al., 2019). GLs can be segregated into various classes (over 200) according to the different types of groups which are linked to them with the major ones being sulfoxides, methyl sulfites, methyl sulphates, saturated aliphatics, allylics, aromatics and indolyls (Ishida et al., 2014). The bioconversion of GLs into active ITCs is regulated by the presence of an endogenous myrosinase which is activated upon tissue disruption as part of plant defense mechanism (Kyriakou et al., 2022a). However, variations on pH, ferrous ions, ascorbate availability and epithiospecifier protein(s) expression can drive rearrangement into nitriles, thiocyanates, epithionitriles and oxazolidine-2-thiones rather than ITCs (Wentzell and Kliebenstein, 2008; Hanschen and Schreiner, 2017; Wang et al., 2019).

The biochemical importance of watercress supplementation (or its extracts) has been reported numerous times in the literature. For instance, it has been previously suggested that watercress juice controls hyperglycemia by restricting the activity of α-glucosidase while enhances that of lipase and α-amylase (Spínola et al., 2017). In another study, it has been demonstrated that watercress consumption was associated with augmentation of superoxide dismutase and catalase activities with a concomitant reduction of hepatic glutathione (GSH), glutathione reductase (GR), glutathione peroxidase (GPx) and malondialdehyde (MDA) contents in hypercholesterolaemic rats (Azarmehr et al., 2019). In addition, others have demonstrated that the administration of raw watercress increases the accumulation of plasma antioxidants including lutein and β-carotene (Gill et al., 2007). Finally, other studies have shown that watercress extracts enriched in either phenethyl isothiocyanate; PEITC or polyphenolic compounds induce cytotoxicity in various human cancer cell lines including melanoma, prostate, leukemia, cervical, liver, colon, lung, myeloma and breast (Gao et al., 2011; Kim et al., 2011; Sun et al., 2019; Mitsiogianni et al., 2021a; Mitsiogianni et al., 2021b).

To conclude, this study focuses into the descriptive characterization and evaluation of the main phytochemicals (glucosinolates, polyphenolics and isothiocyanates) present in the aerial parts of watercress including watercress flowers, leaves and stems. The information obtained also highlights the possibility of exploitation of watercress by products, such as stems.

2 Materials and methods

2.1 Reagents

Solvents: Methanol LC-MS, grade ≥ purity 99.9% (34860), water HPLC grade (34877), acetonitrile HPLC grade, purity ≥ 99.9 (34851), were purchased from Honeywell (Medisell Nicosia, Cyprus). Formic acid LC-MS grade (85178) and trifluoroacetic acid LC-MS grade (85183) were purchased from Thermofisher Scientific (G. Georgiou, Nicosia, Cyprus). 2M ammonia solution in isopropanol (392693) was purchased from Sigma Aldrich (Vouros, Nicosia, Cyprus). The analytical standards: gluociberin potassium salt (2513S), glucoraphanin potassium salt (2509S), glucocamelinin potassium salt (2517S), glucoarabin potassium salt (2516S), homoglucocamelinin potassium salt (2518), glucoraphanin potassium salt (2514S), glucocheirolin potassium salt (2524S), glucolepidiin potassium salt (2505S), glucoerucin potassium salt (2504S), glucoberteroin potassium salt (2501S), sinigrin potassium salt (7295S), gluconapin potassium salt (2507S), glucobrassicanapin potassium salt (2502S), progoitrin potassium salt (2515S) epiprogoitrin potassium salt (2512S), glucotropaeolin potassium salt (2510S), gluconasturtiin potassium salt (2508S), sinalbin potassium salt (2511S), glucolimnanthin potassium salt (2520S), glucobrassicin potassium salt (2520S), neoglucobrassicin potassium salt (2519S), 4-methoxyglucobrassicin potassium salt (2526), glucomorignin potassium salt (2506S), gallic acid (4993S), chlorogenic acid (4991S), ferulic acid (4753S), ellagic acid (6075), vanillin (6110S), caffeic acid (6034S), syringic acid (6011), p-coumaric acid (4751S), rosmarinic acid (4957S), 4-hydroxybenzoic acid (6099), protocatechuic acid (6050), 2'-hydroxyflavanone (1180), 7-hydroxyflavanone (1212), 4'-methoxyflavanone (1185), 5-methyxyflavanone (1186), apigenin-7-*O*-glucoside (1004S), luteolin-7-*O*-glucoside (1126S), isorhamnetin (120S), quercetin-3-*O*-rhamnoside (1236S), hyperoside (1027S), myricetin-3-*O*-galactoside (1355S), kaempferol-3-*O*-rutinoside (1053), ipriflavone (1328), naringin (1129S), were purchased from Extrasynthese (Lyon, France). Iberin (ab141944) and Sulforaphane (ab141969) were purchased from Abcam (Cambridge, UK. Allyl isothiocyanate (36682), benzyl isothiocyanate (89983), phenethyl isothiocyanate (68488) and indole-3-carbinol (I7256) were purchased from Sigma Aldrich (Vouros, Nicosia, Cyprus).

2.2 Plant material cultivation, processing and storage

Fresh watercress samples were kindly provided by the Watercress Company, Dorchester, Dorset, UK. The aerial part of watercress plants including flowers, lateral buds, petioles and stems were kept at -20°C until further use. Then, they were immersed in liquid nitrogen prior to being dehydrated in a freeze-drier (Alpha 1-4 LSC Basics, Christ) at -55°C, 0.05 mbar for 96 hrs. The dried parts were sprayed with liquid nitrogen and milled to fine powder using a domestic blender. The freeze-dried watercress powdered samples were stored at -80°C in a sealed bag protected from air, humidity and light until further use.

2.3 Extraction of polyphenolic compounds

The extraction of polyphenolic compounds was performed according to Kyriakou et al. (2022b). Briefly, one (1.0) gr of each of the examined watercress samples were extracted with exhaustive maceration at 80°C, with aqueous methanol 80% (*v/v*) for 48 hrs (Kyriakou et al., 2022b). The resulting mixture was filtered through a Whatman filter paper (pore size: 4.0-12 μm). The process was repeated twice. The combined methanolic solutions were lyophilized on a freeze-drier (Alpha 1-4 LSC Basics, Christ). The reconstructed (in 100% methanol) extracts, were filter (0.22 μm) (mixed cellulose esters, MCE) syringe and directly injected into UPLC-MS/MS for analysis.

2.4 Extraction of GLs

The extraction of GLs was accomplished according to Yu et al. (2022) with some modification. Briefly, one (1.0) gr of each watercress sample (either flowers or leaves or stems) were mixed with 100 mL aqueous methanol 70% (*v/v*). The resulting suspension was heated at 80°C for 30 mins and then it was sonicated for further 30 mins at room temperature (25°C). The resulting extract was centrifuged at 3000 x *g* for 20 min and 1 mL of the supernatant was diluted with UPLC grade water (1:10 dilution). The diluted solution was passed through a 0.22 μm (mixed cellulose esters, MCE) syringe filter, and directly analyzed *via* UPLC MS/MS.

2.5 Hydrolysis of GLs and extraction of ITCs

The hydrolysis of GLs was performed based on modified previously published procedure (Kyriakou et al., 2022b). Briefly, five (5.0) g of each watercress sample (flowers, leaves or stems) was dissolved in 350 mL phosphate buffer saline (PBS) (pH 7.0) containing 0.5 mmol of ascorbic acid. The formed suspension was heated at 37°C and stirred vigorously for 2 hrs. Then, the hydrolyzed mixture was extracted by stirring at 37°C for further 2 hrs with either 400 mL hexane (for PEITC extraction) or dichloromethane [for iberin (IBN), sulforaphane (SFN), indole-3-carbinol and benzyl isothiocyanate (BITC)] or diethyl ether for allyl isothiocyanate (AITC). Upon completion of the extraction, the organic phase was isolated, dried over magnesium sulfate and concentrated under reduced pressure at 40°C. The formed oils were reconstituted in acetonitrile, filtered twice through a 0.22 μm (mixed cellulose esters, MCE) membrane and rapidly mixed with 500 μL of 2M ammonia is isopropanol (apart from indole-3-carbinol enriched extract). The formed solutions were allowed at 25°C for 24 hrs. Then, the solvents were evaporated to dryness under reduced pressure. The dried thiourea derivatives were taken up in methanol containing 0.1% TFA.

2.6 Preparation of standards and samples

Stock solutions of 4-hydroxybenzoic acid, protocatechuic acid, gallic acid, vanillin, syringic acid, p-coumaric acid, caffeic acid, ferulic acid, rosmarinic acid, chlorogenic acid, ellagic acid, 7-hydroxyflavanone, 4'-methoxyflavanone, apigenin-7-*O*-glucoside, isorhamnetin, quercetin-3-*O*-rhamnoside, quercetin-3-*O*-rutinoside (rutin), naringin, kaempferol-3-*O*-rutinoside, hyperoside, myricetin-3-galactoside, were prepared in methanol. Luteolin-7-*O*-glucoside in acetonitrile/water mixture (1:1) and 2'-hydroxyflavanone, 5-methoxyflavanone and ipriflavone in methanol/acetonitrile mixture (1:1). Glucolepidiin, sinigrin, gluconapin, glucobrassicanapin, progoitrin, epiprogoitrin, glucotropaeolin, glucoerucin, gluconasturtiin, glucoiberin, glucoraphenin, glucoberteroin, glucoraphanine, glucocheirolin, glucobrassicin, neoglucobrassicin, 4-methoxyglucobrassicin, glucocamelinin, homoglucocamelinin, glucoarabin, glucomoringin, sinalbin and glucolimnanthin in 70% (*v/v*) methanol. All stock solutions were 1000 ppm.

For the content analysis of ITCs and indole-3-carbinol, the analytical standards were added with 2M ammonia in isopropanol for 24 hrs and then were evaporated to dryness under a nitrogen stream and reconstituted in a methanol: 0.1% (*v/v*) TFA mixture (4:1) whereas indole-3-carbinol was directly dissolved in methanol: 0.1% (*v/v*) TFA mixture (4:1) at a concentration of 1000 ppm. Working standard solution was made by diluting the individual standard stock solutions with ice cold methanol. Watercress extracts were diluted with ice cold methanol at final concentration of 25 ppb. Each solution was kept in the dark, protected from light in order to minimize the autooxidation of polyphenols. In addition, stock, standard and sample solutions were stored at -20°C before use. All prepared

solutions were passed through 0.22 µm (mixed cellulose esters, MCE) membrane filtered prior UPLC-QqQ-ESI-MS/MS analysis.

2.7 Quantification of polyphenolic compounds, intact GLs and ITCs

2.7.1 Liquid chromatography (LC) conditions

For the detection and quantification of the listed polyphenols, a Waters Acquity UPLC system (Waters Corp., Milford, MA, USA) equipped with an autosampler chamber, two pumps and a degasser, was used. The chromatographic separation was performed on an ACQUITY UPLC BEH C18 (100 x 2.1 mm, particle size: 1.7 µm) column (Waters Corp., Milford, MA, USA), heated at 30°C and eluted as it was previously reported with some modification for polyphenolic compounds, intact GLs and ITCs respectively (Song et al., 2005; Tian et al., 2005; Zhu et al., 2020).

2.7.1.1 Polyphenolic compounds

Briefly, the mobile phase was consisted of a solution of acetonitrile (eluent A) and formic acid 0.1% (*v/v*) (eluent B). A flowrate of 0.3 mL/min was used and the linear gradient conditions applied consisted of 5-100% A (0-4 min), 100-90% A (4.0-4.1 min), 90%A (4.1-5 mins), 90-5% A (5-5.01 mins) and 5% A (5.1-6 mins).

2.7.1.2 Intact GLs

The mobile phase consisted of methanol (eluent A) and formic acid 0.1% (*v/v*) (eluent B). A flowrate of 0.2 mL/min was used and the linear gradient conditions were 10% A (0-3 mins), 25% A (3-5 mins), 60% A (5-6 mins) and 0% A (6-6.2 mins), 10% A (6.2-9 mins).

2.7.1.3 Derivatized ITCs and indole-3-carbinol

The mobile phase was consisted of a solution of 80% methanol (eluent A) and trifluoroacetic acid (TFA) 0.1% (*v/v*) (eluent B). A flowrate of 0.3 mL/min was used and the linear gradient conditions applied consisted of 50% A (0-5 mins), 60% A (5-10 mins), 70% A (10-20 mins) and 80% A (20-35 mins). The injection volume for all standards and analytes was 10 µL, and the autosampler temperature was set at 4°C during all analyses.

2.7.2 MS/MS conditions

For the MS/MS experiments, a Xevo Triple Quadrupole (TQD) Mass detector (Waters Corp., Milford, MA, USA) was operated in either positive or negative ionization mode (ESI±). Detection of the various analytes was performed using selected ion recording (SIR) mode utilizing the collision voltage (MS1), individually determined for each analyte. Quantitative analysis was accomplished using selected multiple reaction monitoring (MRM) mode. The MRM conditions were optimized for each standard, by MS manual tuning of each standard prior to sample analysis at a concentration of 1 ppm (Tables S1–S3). In order to acquire maximum signals, the optimized tuning parameters were as follow: capillary voltage: 2.5-3.0 kV; cone voltage: 36 V; source temperature: 150°C; disolvation temperature: 500°C; source disolvating gas flow: 1000 L/h and gas flow: 20 L/h. High-purity nitrogen gas was used as the drying and nebulizing gas, whereas ultrahigh-purity argon was used as a collision gas. The data acquisition and processing were performed on MassLynx software (version 4.1).

2.8 Statistical analysis

All statistical analyses and Pearson's correlation plotswere generated *via* R Statistical Software (2022), heatmaps were plotted using heatmap.2 from the gplots package in R. Statistical comparison of the analyte content among the aerial parts of watercress was performed by one-way ANOVA test, with post-hoc Tukey test for multiple comparisons utilizing appropriate software (GraphPad Prism version 8.0.1).

3 Results

The experimental set-up involved the segregation of flowers, leaves and stems followed by their dehydration and pulverization into a fine powder. Various extraction procedures were followed in order to obtain fractions of polyphenolic acids, GLs and ITCs (in the form of thiourea derivatives) (Figure 1).

3.1 Detection of GLs and polyphenolic compounds in aerial watercress parts

The first part of the analysis involved the detection of GLs with the prosthetic group been either sulfoxide or methyl sulphate or saturated aliphatic or methyl sulphite or allylic or aromatic. For this purpose, a selected ion recording (SIR) experiment was run involving watercress flower, leaves and stems in their GL-enriched extract in order to identify the masses of all tested GLs using the collision energy obtained from the manual tuning of each of the respective GL standards (Figure 2A; S1).

The scanning results revealed that among all 23 screened GLs, in the flowers, the levels of sinalbin and glucolimnanthin were below the detection limit of our method, in addition to those obtained from the leaves (e.g., glucoiberin, glucoberteroin, glucoraphanine, neoglucobrassicin and glucolimnanthin). Moreover, those obtained from the stems, levels of

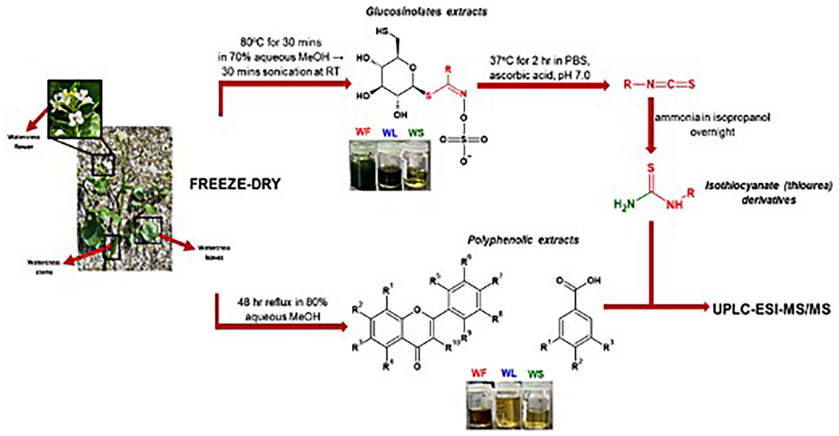


FIGURE 1
The methodological pipeline for the determination of polyphenolic and GL contents in the aerial parts of watercress: WF, watercress flowers; WL, watercress leaves; and WS, watercress stems.

glucoberteroin, glucoraphanine, glucocheirolin, sinalbin and glucolimnanthin were not detected as well. In all three extract samples, gluconasturtiin and glucobrassin were at the highest abundance. Overall, the flower samples contained the highest abundance of GLs compared to leaves and stems. Finally, the same approach was followed for the polyphenolic (e.g., phenolic acids, flavanones, flavones, flavonols and isoflavone) compounds (Figures 2B; S2) as well as derivatised ITCs (thioureas)/indole-3-carbinol (Figures 2C; S3).

3.2 Standardization of UPLC and MS conditions

Afterwards, we sought into the quantification of each of the GLs individually by employing MRM transitions and by utilising commercially available reference standards. Finally, the combination of the mobile phase, elution mode, flow rate and column used for the separation were chosen in order to acquire the optimal signal for all the analytes.

For the determination of the optimum mobile phase several combinations were applied including methanol/water and acetonitrile/water in different ratios, with none of them being effective in the improvement of the shape of the peaks. Acidification of water with 0.1% (*v/v*) formic acid, was chosen since it allowed to obtain peaks with better symmetry and shape. Additionally, the presence of formic acid facilitated the ionization of the compounds. In the case of derivatized ITCs, TFA was preferred to rather than formic acid, as it was lower the pH of the mobile phase and improved the width of the peaks. Improvement of the shape of the peaks was achieved by increasing the column temperature to 30°C rather than 20°C for GLs and polyphenolic compounds and to 35°C for ITCs as it was previously suggested by others (Pilipczuk et al., 2017; Ruslin et al., 2022). For the ionization of polyphenols, the electrospray ionization (ESI) with either negative or positive ($ESI^{\pm}$) mode was used while GLs and derivatized ITCs were ionized under ESI^{-} and ESI^{+} respectively.

3.3 Method validation

The analytical method was validated according to the guidelines of European Medicines Agency (European Medicine Agents, 1995). Namely, parameters including, linearity, limit of detection (LOD), limit of quantification (LOQ), precision, accuracy were determined (Tables S4–6). The generated calibration curves of the standards were plotted with linear regression equation of peak areas versus various concentrations ranging from 0.65 to 505.60 ppb for polyphenolic compounds (Figure S4), 1.95-250 ppm for GLs (Figure S5) and derivatized ITCs (Figure S6). All polyphenolic compounds and GLs demonstrated good linearity in the range of 0.81-513.20 ppb and1.50-258.20 ppb respectively whereas the correlation coefficients (R^2) were >0.99 for all the analyzed standards (Table 1).

Finally, we evaluate the reproducibility of the UPLC-QqQ-ESI-MS/MS method by means of determining the % of recovery (Tables S4–S6) based on the quantification protocol followed for GLs, polyphenolic compounds and derivatized ITCs respectively. For this purpose, each watercress extract was spiked with mixtures of standard solutions of various GLs or polyphenols or ITCs. Spike samples were prepared in triplicates and the results were of at least six repetitions. The % recovery was calculated according to equation (**1**), where A is the final amount detected, A_0 is the initial amount and A_a is the

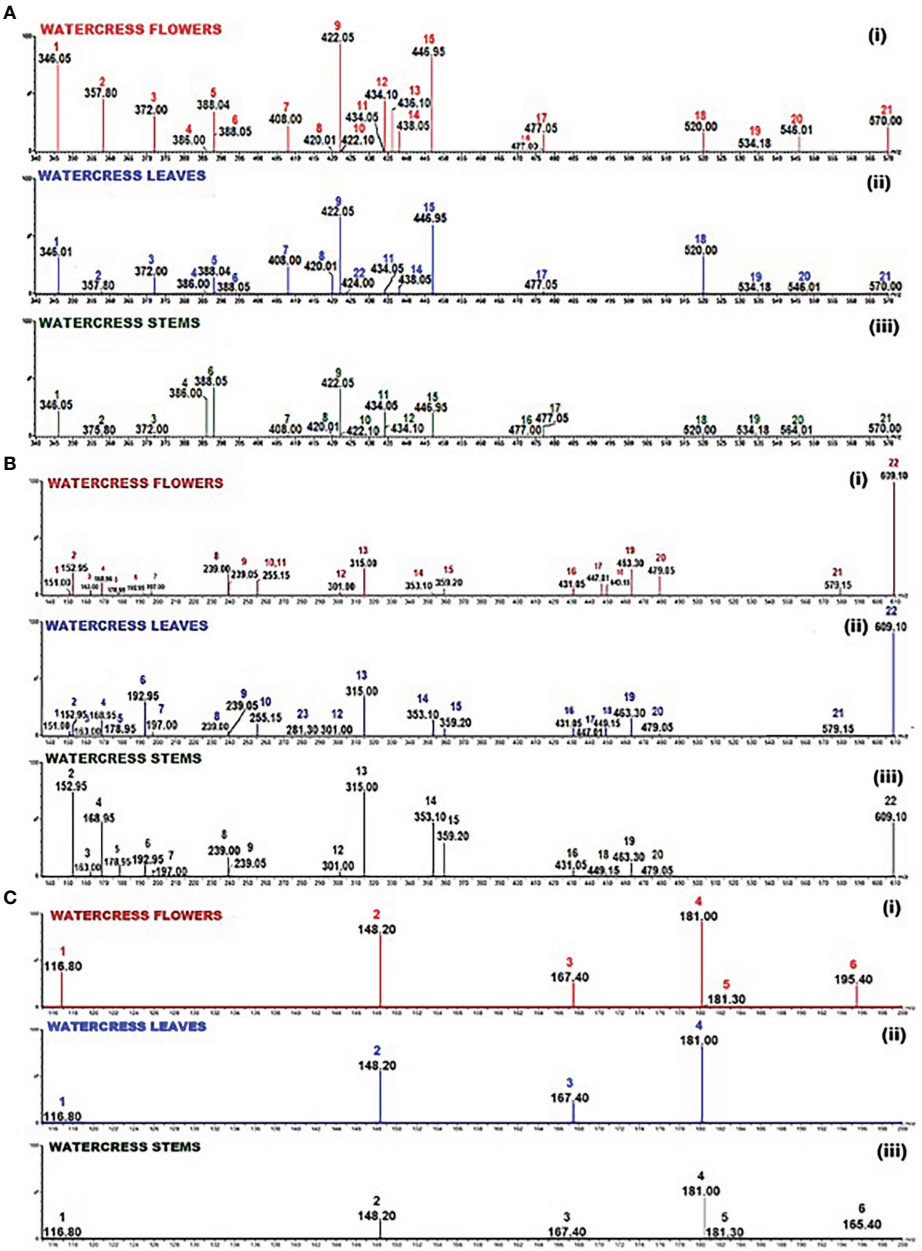


FIGURE 2
Selected Ion Recording (SIR) spectrum of: **(A)** the GL-enriched fractions of watercress (i) flowers, (ii) leaves and (iii) stems. Scanning (m/z; 340-570) was performed using collisions energies and m/z in the negative electrospray ionisation (ESI-) mode according to the collision energy as pointed in Table S1; 1 - glucolepidiin, 2 - sinigrin, 3 - gluconapin, 4 - glucobrassicanapin, 5 - progoitrin, 6 - epiprogoitrin, 7 - glucotropaeolin, 8 - glucoerucin, 9 - gluconasturtin, 10 - glucoiberin, 11 - glucoraphenin, 12 - glucoberteroin, 13 - glucoraphanine, 14 - glucocheirolin, 15 - glucobrassicin, 16 - neoglucobrassicin, 17 - 4-methoxyglucobrassicin, 18 - glucocamelinin, 19 - homoglucocamelinin, 20 - glucoarabin, 21 - glucomoringin, 22 – sinalbin; **(B)** the polyphenolic compounds in the hydromethanolic fraction of watercress flowers, leaves and stems. Scanning (m/z; 135-610) in both positive and negative electrospray ionisation (ESI±) mode according to the collision energy as pointed in Table S2; 1 - vanillin, 2 - protocatechuic acid, 3 - p-coumaric acid, 4 - gallic acid, 5 - caffeic acid, 6 - ferulic acid, 7 - syringic acid, 8 - 2′-hydroxyflavanone, 9 - 7-hydroxyflavanone, 10 - 4′-methoxyflavanone, 11 - 5′-methoxyflavanone, 12 - ellagic acid, 13 - isorhamnetin, 14 - chlorogenic acid, 15 - rosmarinic acid, 16 - kaempferol-3-O-rutinoside, 17 - quercetin-3-O-rhamnoside, 18 - luteolin-7-O-glucoside, 19 - hyperoside, 20 - myricetin-3-galactoside, 21 - naringin, 22 - quercetin-3-O-rutinoside, 23 – ipriflavone; **(C)** the derivatised into thiourea ITCs. Scanning (m/z 116.80-195.40) in positive electrospray ionisation mode according to the collision energy listed in Table S3; 1 - allyl thiourea, 2-indol-3-carbinol, 3 - benzyl thiourea, 4 - phenethyl thiourea, 5 - iberin thiourea, 6 - sulforaphane thiourea.

10.3389/fpls.2022.998755

TABLE 1 Quantitative data demonstrating the screened glucosinolates, polyphenolic compounds and isothiocyanates in the aerial parts (flowers, leaves and stems) of watercress.

Analyte	Watercress flowers	Watercress leaves	Watercress stems
		ng of analyte/g of dry extract	
		GLUCOSINOLATES	
Glucoiberin	0.37 ± 0.06^{b}	N.D.	0.14 ± 0.04^{a}
Glucoraphanine	17.98 ± 0.50^{b}	N.D.	0.04 ± 0.01^{a}
Glucocamelinin	5.97 ± 0.58^{b}	10.44 ± 0.32^{c}	0.040 ± 0.006^{a}
Homoglucocamelinin	0.41 ± 0.03^{b}	0.22 ± 0.03^{a}	N.D.
Glucoarabin	6.20 ± 0.14	N.D.	N.D.
Glucoraphenine	0.48 ± 0.08^{a}	3.57 ± 0.11^{b}	4.12 ± 0.06^{b}
Glucocheirolin	10.26 ± 0.96^{b}	0.23 ± 0.03^{a}	0.010 ± 0.005^{a}
Glucolepidiin	67.78 ± 0.84^{c}	8.30 ± 0.38^{b}	3.53 ± 0.16^{a}
Glucoberteroin	21.13 ± 0.92^{b}	N.D.	0.65 ± 0.03^{a}
Glucoerucin	0.19 ± 0.02^{b}	3.15 ± 0.10^{c}	0.020 ± 0.002^{a}
Sinigrin	30.63 ± 0.37^{c}	0.21 ± 0.02^{a}	0.030 ± 0.002^{a}
Gluconapin	14.95 ± 0.08^{c}	3.24 ± 0.16^{b}	0.0100 ± 0.0006^{a}
Glucobrassicanapin	0.090 ± 0.001^{b}	0.05 ± 0.01^{a}	5.25 ± 0.09^{c}
Epiprogoitrin	2.13 ± 0.08^{c}	0.34 ± 0.03^{a}	7.48 ± 0.05^{b}
Progoitrin	15.38 ± 0.47^{b}	3.96 ± 0.47^{a}	N.D.
Glucolimnanthin	N.D.	N.D.	N.D.
Sinalbin	N.D.	0.02 ± 0.01	N.D.
Gluconasturtiin	82.11 ± 0.63^{c}	36.25 ± 0.74^{b}	9.20 ± 0.11^{a}
Glucotropaeolin	9.59 ± 0.48^{c}	6.22 ± 0.04^{b}	0.0200 ± 0.0008^{a}
Glucobrassicin	74.93 ± 0.85^{c}	21.11 ± 0.10^{b}	1.61 ± 0.24^{a}
Glucomoringin	9.42 ± 0.12^{b}	0.020 ± 0.001^{a}	0.10 ± 0.14^{a}
4-methoxyglucobrassicin	7.76 ± 0.77^{c}	2.29 ± 0.13^{b}	0.95 ± 0.03^{a}
Neoglucobrassicin	0.48 ± 0.02^{a}	N.D.	0.45 ± 0.07^{a}
4-hydroxybenzoic acid	N.D.	N.D.	N.D.
Protocatechuic acid	134.72 ± 1.21^{c}	96.79 ± 0.79^{b}	73.59 ± 0.48^{a}
Gallic acid	60.74 ± 0.44^{c}	56.04 ± 1.93^{b}	32.25 ± 0.26^{a}
Vanillin	11.20 ± 0.92^{b}	3.60 ± 0.07^{a}	N.D.
Syringic acid	13.83 ± 0.09^{c}	3.94 ± 0.06^{b}	1.59 ± 0.07^{a}
p-coumaric acid	17.30 ± 0.40^{b}	22.52 ± 4.22^{c}	3.58 ± 0.06^{a}
Caffeic acid	30.89 ± 0.12^{c}	13.44 ± 0.35^{b}	7.63 ± 0.07^{a}
Ferulic acid	17.70 ± 0.38^{b}	306.98 ± 2.44^{c}	7.94 ± 0.07^{a}
Rosmarinic acid	33.65 ± 1.01^{c}	26.54 ± 0.41^{b}	24.77 ± 0.19^{a}
Chlorogenic acid	41.34 ± 1.00^{c}	59.34 ± 0.17^{b}	38.84 ± 0.43^{a}
Ellagic acid	6.51 ± 1.06^{c}	2.38 ± 0.23^{b}	0.65 ± 0.04^{a}
2'-hydroxyflavanone	49.47 ± 0.71^{c}	32.75 ± 0.78^{b}	12.71 ± 0.30^{a}
7-hydroxyflavanone	39.94 ± 0.67^{c}	19.33 ± 0.53^{b}	7.31 ± 1.15^{a}
4'-methoxyflavanone	12.96 ± 0.81^{b}	6.12 ± 0.10^{a}	N.D.
5-methoxyflavanone	6.15 ± 0.02	N.D.	N.D.
Apigenin-7-*O*-glucoside	15.64 ± 0.24^{b}	3.29 ± 0.03^{a}	N.D.
Luteolin-7-*O*-glucoside	35.18 ± 1.14^{c}	7.39 ± 0.09^{b}	1.00 ± 0.07^{a}
Isorhamnetin	289.40 ± 1.37^{c}	196.47 ± 3.65^{b}	65.67 ± 0.81^{a}
Quercetin-3-*O*-rhamnoside	35.08 ± 0.70^{b}	2.22 ± 0.24^{a}	N.D.
Quercetin-3-*O*-rutinoside	1459.30 ± 12.95^{c}	1197.86 ± 4.24^{b}	41.02 ± 0.18^{a}
Hyperoside	125.73 ± 1.75^{c}	87.96 ± 0.53^{b}	6.51 ± 0.26^{a}
Myricetin-3-galactoside	63.15 ± 1.94^{c}	12.46 ± 0.33^{b}	0.81 ± 0.04^{a}

(Continued)

TABLE 1 Continued

Analyte	Watercress flowers	Watercress leaves	Watercress stems
		ng of analyte/g of dry extract	
Kaempferol-3-*O*-rutinoside	257.54 ± 2.31^{c}	10.41 ± 0.52^{b}	2.25 ± 0.13^{a}
Ipriflavone	N.D.	N.D.	N.D.
Naringin	21.72 ± 0.39^{b}	11.16 ± 0.11^{a}	N.D.
Sulforaphane	9.43 ± 0.24^{b}	N.D.	0.16 ± 0.01^{a}
Iberin	0.49 ± 0.01^{b}	N.D.	0.060 ± 0.007^{a}
Benzyl isothiocyanate	13.32 ± 0.39^{c}	0.77 ± 0.05^{b}	0.020 ± 0.006^{a}
Phenethyl isothiocyanate	273.89 ± 0.88^{c}	125.02 ± 0.52^{b}	64.70 ± 0.90^{a}
Allyl isothiocyanate	16.04 ± 0.81^{c}	0.75 ± 0.13^{b}	0.040 ± 0.004^{a}
Indole-3-carbinol	191.44 ± 1.99^{c}	95.55 ± 0.94^{b}	16.64 ± 0.28^{a}

N.D. indicates the non – detected metabolites. Data are the means of six independent experiments ± SD. Means ± SD followed by distinct letters in the same row statistically differ according to Tukey's *post hoc* test, ($p< 0.05$).

added amount:

$$\% \text{ recovery} = \left[\frac{A - A_0}{A_a}\right] \text{x } 100\,\% \quad (1)$$

The average recovery of all of the polyphenolic compounds ranged between 89.2% and 102.6% while the respective range for all intact GLs was 84.0-101.2%. In the case of derivatized ITCs (thioureas) the recovery ranged between 80.2-96.7% thereby demonstrating the accuracy and reproducibility of our methodological approach.

3.4 Linearity, accuracy and precision of the methodology

Limit of detection (LOD) and quantification (LOQ) values were calculated based on the signal to noise ratio (*S/N*) which was set at 3 and 10 respectively. The range of LOD and LOQ determination for polyphenolic compounds were 0.65-109.90 ppb and 1.21-105.20 ppb respectively, whereas for intact GLs was 0.27-6.36 and 0.79-22.65 ppb. Furthermore, the calculated LOD and LOQ ranges for the various thioureas (which correspond to ITCs) were between 2.73-24.43 ppb and 9.12-81.45 ppb respectively. Based on the values presentenced in Table S5, from all the polyphenols, the LOD of *p*-coumaric acid was the lowest and that of gallic acid was the highest suggesting that the sensitivity of *p*-coumaric acid was better than for the other polyphenols ionized in ESI^-. On the contrary, the highest sensitivity was shown to be for luteolin-7-*O*-glucoside and 7-hydroxyflavanone compared to the others being ionized in ESI^+. Furthermore, the LOD for glucoiberin was the lowest whereas the highest LOD was denoted for epiprogoitrin thus indicating a higher sensitivity in detecting glucoiberin among all other GLs (Table S4). Moreover, the LOD of indole-3-carbinol was the highest whereas the respective one for AITC was the lowest suggesting better sensitivity among the thioureas ionized *via* ESI^+ (Table S6).

Finally, the degree of similarity of multiple samples within the same family (bearing the same functional groups) was evaluated by means of percent relative standard deviation (% RSD). In order determine the % RSD, six replicated samples, at the same concentration, were analyzed within one day for intra-day precision and within five continuous days for inter-day precision. More specifically, % RSD results of both intra- and inter-day were 1.89-4.6% (for intact GLs) (Table S4) 0.5-4.2% (for polyphenolic compounds) (Table S5), and 1.04-4.02% (Table S6) (for derivatized ITCs), respectively.

3.5 Polyphenolic compounds and intact GLs fragmentation

The collision gas fragmentation of the precursor molecules led to the formation of increasingly lower m/z (fragments) ions at *m/z* 93-197. In particular, the fragmentation pattern of the phenolic acids [e.g., 4-hydroxybenzoic acid (136.95>93), protocatechuic acid (*m/z* 152.9>108.95), gallic acid (*m/z* 68.95>124.95), *p*-coumaric acid (*m/z* 163>119) and caffeic acid (*m/z* 178.95>134.95)] is characterized by loss of the carboxyl group (*m/z* 44) $[M\text{-}H\text{-}C(=O)OH]^-$. On the other hand, in some phenolic acids [e.g., vanillin (*m/z* 151>136), syringic acid (*m/z* 197>182)], anionic species were produced by the loss of their methyl group; $[M\text{-}H\text{-}CH_3]^-$ whereas in the case of ferulic acid (*m/z* 192.95>178) a loss of both methyl and carboxyl acid moieties occurred $[M\text{-}H\text{-}59]^-$. The gas-induced fragmentation of chlorogenic acid resulted in loss of the caffeyl functionality $[M\text{-}H\text{-caffeyl}]^-$ yielding *m/z* 359.2>145. Whereas, the fragment produced from rosmarinic acid was generated upon dehydration of the cleaved ester products $[M\text{-}H\text{-}179]^-H_2O$ or $[M\text{-}H\text{-}197]^-2H_2O$, *m/z* 359.2>161. Since chlorogenic and rosmarinic acid share the same base peak, their identification was performed

based on the fragmentation pattern, which included fragments in the range of 65-173 *m/z*. The quantification of each acid was then performed on the most abundant fragment of each compound, different from the base peak one.

Similar observations were also noted in flavanones. Namely, 2'-hydroxyflavanone, 7-hydroxyflavanone, apigenin-7-*O*-glucoside, isorhamnetin, quercetin-3-*O*-rutinoside, quercetin-3-*O*-rhamnoside, hyperoside, myricetin-3-galactoside, kaempferol-3-*O*-rutinoside and nariginin produced $[M-H]^-$ at *m/z* 239-609.1 whereas 4'-methoxyflavanone, 5'-methoxyflavanone, luteolin-7-*O*-glucoside, and ipriflavone produced $[M+H]^+$ at *m/z* 255.15-449.15. On the other hand, MS^2 fragmentation generated fragment ions with signals at *m/z* 91.15-300 and *m/z* 151.3-287.1 for negative and positive ionization modes respectively. Namely, collision-induced fragmentation of 2'-hydroxyflavanone resulted in the loss of chromone $[M-H\text{-chromone}]^-$, whereas fragmentation of 7-hydroxyflavanone resulted in the loss of chromenone $[M-H\text{-chromenone}]^-$ thus generating signals *m/z* 240.47>93.1 and *m/z* 240>91.15, respectively. In the case of 4'-methoxyflavanone (*m/z* 255.15>240) cleavage of the methyl group (*m/z* 15) was noticed. Furthermore, during MS^2 fragmentation, apigenin-7-*O*-glucoside (*m/z* 431.15>268.35), luteolin-7-*O*-glucoside (*m/z* 449.15>187.1) and hyperoside (*m/z* 463.3>300) yielded aglucon fragments by the loss of glucoside $[M \pm H\text{-glucoside}]^{\pm}$. In respect to quercetin-3-*O*-rutinoside and quercetin-3-*O* rhamnoside, the collision-induced fragmentation generated signals $[M-2H\text{-rutinoside}]^{2-}$ at *m/z* 447.01>300 and $[M-H\text{-rhamnoside}]^-$ at *m/z* 609.1>300, respectively. A similar fragmentation pattern was also noticed in the case of kaempferol-3-*O*-rutinoside and naringin as the parental molecules lost rutinoside and glucorhamnoside respectively thus allowing the detection of the fragment ionsat *m/z* 431.05>255.3 and *m/z* 579.15>271.1. Finally, fragmentation of isorhamnetin led to elimination of 2,4,6-trihydroxybenzaldehyde $[M-H-164]^-$ thus generating a signal at *m/z* 315>151 whereas elimination of isopropyl moiety of ipriflavone $[M-2H-i\text{Pr}]^{2-}$ yielded a signal at *m/z* 281.3>240.

With respect to the MS^2 fragmentation of GLs, it appears that all of them follow the same fragmentation pattern. Collision gas fragmentation of the precursor GLs led to the production of a maximum abundant precursor ion at *m/z* 97. Therefore, the MS^2 ionization induced a cleavage at the sulfate functionality forming an ion $[SO_3H^-]$ with *m/z* 97. Glucomoringin is the only GL with a distinct MS^2 fragmentation pattern and which produced a strong signal at *m/z* 570.

In general, ITCs can be detected in the ESI^- mode, however, the analysis is of low sensitivity thus preventing the accurate quantification of the ITC content. Therefore, ITCs were derivatized into their respective thioureas upon treatment with alcoholic solution of ammonia. This modification increased the sensitivity and accuracy of the method and thus facilitating their quantification (Song et al., 2005). In addition, in this study, we utilized indole-3-carbinol as a representative hydrolysis product of the elevated glucobrassicin. This is because, the respective indolyl-3-methyl ITC is unstable in the aqueous environment of hydrolysis and it is rapidly rearranged into the respective indolyl alcohol (Błoch-Mechkour et al., 2010). The fragmentation of all derivatized ITCs followed the same pattern as the MS^2 fragmentation promoted the loss of ammonia (*m/z* 18) thereby resulting in signals for allyl thiourea *m/z* 99.8, iberin thiourea *m/z* 163.4, benzyl thiourea *m/z* 148.9 and sulforaphane thiourea *m/z* 177.2. In the case of phenethyl thiourea the collision energy fragmentation led to the loss of thiourea $[M\text{-thiourea}]^+$ (*m/z* 76) allowing the detection of the transition *m/z* 181.0>104.99. Finally, fragmentation of indole-3-carbinol resulted in the loss of the methanol (*m/z* 32) $[M\text{-MeOH}]^+$ thus generating a transition at *m/z* 148.0>117.2.

3.6 Quantification of intact GLs, polyphenolic compounds and ITCs

The content of intact GLs, polyphenolic compounds and ITCs was assessed by means of both qualification and quantification. Absolute quantities of individual analytes are reported in Table 1 and representation of the total content of the analytes in the aerial parts is illustrated in the heatmap as part of Figure 3. Overall, our findings suggest that among all screened GLs, watercress flowers lack of two aromatic GLs namely glucolimnanthin and sinalbin. In addition, glucoiberin, glucoraphanin, glucoarabin, glucoberteroin, glucolimnanthin and neoglubrassicin were not present in watercress leaves. Finally, homoglucocamelinin, glucoarabin, progoitrin, glucolimnanthin and sinalbin were not detected in watercress stems as well.

Overall, the highest content of GLs was detected in watercress flowers over watercress leaves and stems as proved by statistical analysis. The next aerial part with the highest contest was leaves followed by stems. Among all screened GLs, gluconasturtiin was the most dominant GL in all aerial components of watercress [82.11 ± 0.63 ng/g of dry extract – for watercress flowers; 36.25 ± 0.74 ng/g of dry extract – for watercress leaves; and 9.20 ± 0.11 ng/g of dry extract for watercress stems]. Interestingly, on watercress flowers, glucolepidiin (67.78 ± 0.84 ng/g of dry extract) and glucobrassicin (74.93 ± 0.85 ng/g of dry extract) were also found in high abundance. On the contrary, in watercress leaves, increased glucocamelinin (10.44 ± 0.32 ng/g of dry extract) was observed whereas on watercress stems, glucoraphenin (4.12 ± 0.06 ng/g of dry extract), glucobrassicanapin (5.25 ± 0.09 ng/g of dry extract) and epiprogoitrin (7.48 ± 0.05 ng/g of dry extract) were in the highest proportion. The difference in the content of glucoraphanine, glucocheirolin, glucoberteroin, sinigrin and glucomorignin in watercress leaves and stems, respectively, were not of statistical significance. Similar observations were also noted in the content of sinalbin and neoglucobrassicin as there wasn't any statistical alteration of their content between watercress flowers and stems. The highest levels of glucobrassinapin were

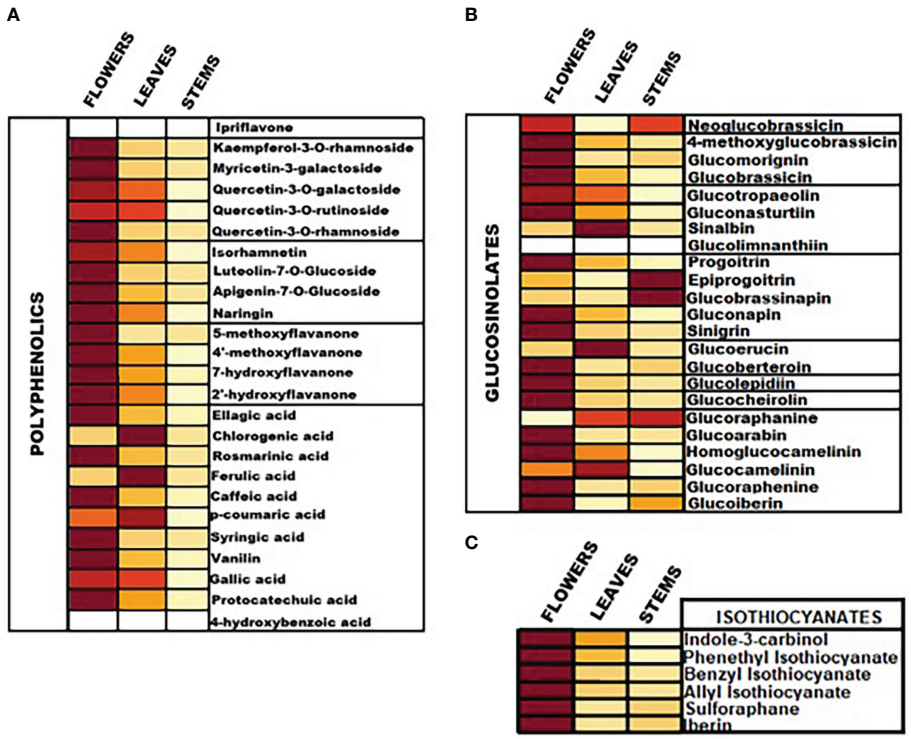


FIGURE 3
Heatmap representation of the relative content of **(A)** polyphenolic compounds, **(B)** intact GLs and **(C)** ITCs in the aerial parts of watercress including flowers, leaves and stems, upon LC-MS/MS quantification.

observed in watercress stems, whereas there wasn't any significant difference in the respective levels of watercress flowers and stems.

In addition, the content of polyphenolic compounds (phenolic acids and flavonoids) was determined in the aerial parts of watercress. The compounds with the highest abundance, in watercress flowers, were quercetin-3-*O*-rutinoside (1459.30 ± 12.95 ng/g of dry extract), isorhamnetin (289.40 ± 1.37 μng/g of dry extract) and kaempferol-3-*O*-rutinoside (257.54 ± 2.31 ng/g of dry extract) The difference in the detected levels of the above-mentioned metabolites was of statistical significance when compared with the respective levels of watercress leaves and stems. The same pattern was followed in watercress leaves, however, sharp elevated levels of ferulic acid (306.98 ± 2.44 ng/g of dry extract) ($p<0.0001$ between watercress flowers-leaves and stems-leaves) were also recorded. For watercress stems, the highest concentration of polyphenolic compounds was determined in protocatechuic acid (73.59 ± 0.48 ng/g of dry extract), chlorogenic acid (38.84 ± 0.43 ng/g of dry extract) and isorhamnetin (65.67 ± 0.81 ng/g of dry extract). However, those levels were scientifically lower when compared to watercress flowers and stems.

On another note, we determined that in all aerial parts of watercress tested, the PEITC content was of the highest abundance followed by indole-3-carbinol. Between the three aerial parts, the highest accumulation of PEITC and indole-3-carbinol was noted on flowers (273.89 ± 0.88 and 191.44 ± 1.99 ng/g of dry extract respectively) ($p<0.0001$) while the lowest one on stems (64.70 ± 0.90 and 16.64 ± 0.28 ng/g of dry extract respectively). Interestingly, our results shown that SFN and IBN while being present in watercress flower and leaves samples this was not the case in watercress stems, with the different in their content in watercress leaves and stems been non-statistical significance

3.7 Correlation of intact GLs and ITCs

We performed a Pearson's correlation analysis in order to observed any linear relationship between the GL and ITC contents in the aerial parts of watercress (Figure 4). Our results demonstrate a strong correlation between the precursors GLs and their direct hydrolysis by-products (e.g., sinigrin – AITC, glucoraphane – SFN, glucoiberin – IBN, glucotropaeolin – BITC, gluconasturtiin – PEITC and glucobrassicin – indole-3-carbinol). In addition, a strong correlation between each ITC and the general class of each GL was noted suggesting a potential contribution of more than one GL in the biosynthesis of specific ITCs. On the contrary, there is no correlation between GLs and ITCs of other classes. The same

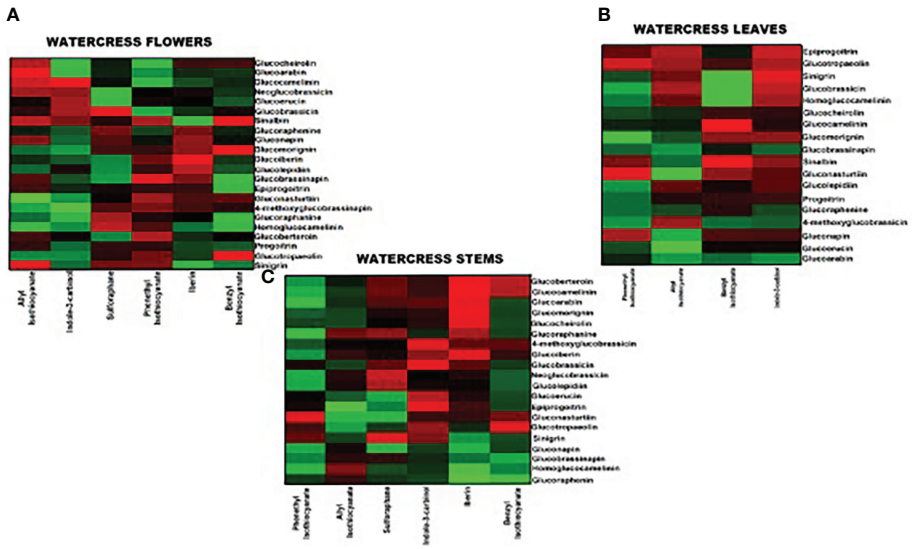


FIGURE 4
Pearson's correlation analysis between intact GL and ITC contents on watercress **(A)** flowers, **(B)** leaves and **(C)** stems. Red indicates a positive correlation (+1.0) whereas bright and dark green denote no correlation or negative correlation (–1.0), respectively.

trends are followed in all aerial parts of watercress. On the other hand, Spearman's correlation analysis (Figure S7) points that there is a correlation between GLs and ITCs of other classes. More specifically, it appears that allylic and indolyl ITCs are strongly correlated with aliphatic, sulfides and methyl sulfoxide-based GLs. This patterns also seems to be appeared in watercress flowers (Figure S7A) and watercress stems (Figure S7C). In addition to this, strong correlation also appears between aromatic and indolyl based ITCs and aliphatic, and allylic based GLs (Figures S7B, C).

4 Discussion

Quantitative determination of the various classes of GLs (e.g., sulfoxides, methyl sulphates, saturated aliphatics, methyl sulfites, allylics, aromatic and indolyls) and polyphenolic compounds (e.g., phenolic acids, flavanones, flavones, isoflavones and flavanols) in the aerial parts of watercress (e.g., flowers, leaves and stems) was performed by the detection of the corresponding masses of intact GLs and polyphenolic compounds utilizing SIR acquisition mode. On the other hand, their quantification was based on the employment of the MRM mode, which is characterized by sensitivity and selectivity. The co-elution of several GLs (e.g., glucoraphanine, glucoraphenine, glucoiberin, quercetin-3-*O*-rhamnoside and apigenine-7-*O*-glucoside) was observed during the UPLC-MS/MS analysis however, successful separation with MRM detection by-passed the potential effect of a simultaneous quantification. In addition, the MS/MS technique reduced the need for chromatographic resolution of individual analytes, minimizing the analytical time thus allowing a higher sample throughput.

GLs are highly polar metabolites, particularly due to the presence of thioglucosyl which behaves as a strong acid (Andini et al., 2020). Hence, the desulfation process was thought to be an essential step for their detection and quantification (Bennett et al., 2002; Förster et al., 2015). However, this methodology has many limitations including the self-degradation, self-dimerization and degradation. In addition, others have attempted to quantify GLs, by utilizing gas chromatography approaches, however, it was found that GLs are thermally unstable and so gas chromatography was withdrawn (Śmiechowska et al., 2010; Chaudhary et al., 2012; Wu et al., 2021). On the other hand, our approach is characterized by the considerably lower time of sample preparation, faster chromatographic separation and also the stability of GLs is not affected by the analysis.

In addition, many studies have been performed in order to quantify ITCs by means of utilizing the cyclo-condensation assay as well as liquid chromatography coupled to UV detector (LC/MS-UV) methodologies. However, these techniques are often accompanied with difficulties in the separation and identification of individual ITCs. On the other hand, the increased volatility of ITCs prevents them from gas chromatography analysis as they rapidly degraded (Hanschen et al., 2015; Pilipczuk et al., 2017). Consequently, derivatization methodologies have been employed based on the electrophilic center of the ITC moiety being used for further derivatization (e.g., *N*-acetyl cysteine bridging or glutathione bridging or mercaptoethanol) leading to the formation of mercapturic acid pathway intermediates which can be easily ionized with ESI and

thus facilitate their quantification (Agerbirk et al., 2015; Pilipczuk et al., 2017; Nacca et al., 2021). In our study, we have chosen derivation with ammonia in order to form stable thioureas. The ammonia derivatization of ITCs is a clean fast and almost quantitatively approach which can be followed for the quantification of individual ITCs (Ji and Morris, 2003; Agerbirk et al., 2015). Our results are in agreement with Jeon et al. (2017) showing that the highest content of all accumulated ITCs was that of gluconasturtiin (Jeon et al., 2017). Moreover, the same group indicated that watercress flowers had the highest content of ITCs among all other plant parts examined. However, our data also shown the presence of increased levels of quercetin-3-rutinoside, isorhamnetin and kaempferol-3-*O*-rutinoside in watercress flowers. Overall, our data are in agreement with the work of others although they are lacking data regarding the localization of these metabolites (Zeb, 2015; Ma et al., 2021; Panahi Kokhdan et al., 2021). To the best of our knowledge, this is the first report which provides a comprehensive analysis of the content of a library of GLs, ITCs, phenolic acids and flavonoids in the aerial parts of watercress.

Finally, according to our data, watercress flowers are the most concentrated source of phenolic acids, flavonoids, GLs and ITCs followed by those of leaves and stems. The elevated quantities of polyphenols (flavonoids) and volatiles (including ITCs) in flowers can be attributed to the fact that they contain the reproductive organs which makes them more valuable for plant fitness (Strauss et al., 2004; Schiestl, 2014). Also, their role defers from that of leaves and stems despite the fact that they share similar chemistry (Kessler et al., 2013). For instance, it has been suggested that the existence of ITCs does not function entirely as a defense mechanism but also to attract pollinator signals related to plant reproduction (Bouwmeester et al., 2019). On the contrary, the existence of volatiles and phytochemicals on leaves and plants are primarily responsible for the activation of the plant defense mechanism upon herbivore of parasites attack (Dixon and Shaw, 2011; Palliyaguru et al., 2018).

Overall, throughout this report, we suggest a strong correlation between ITCs with their respective GLs within the same group. Pearson's association analysis reveals that bioconversion of GLs into ITCs might be regulated either by the rearrangement of other ITCs or by the contribution of other GLs (which belong to the same group) into ITCs biosynthesis, depending on plant necessities. This report provides substantial evidence for the potential exploitation of various aerial parts of the watercress plant in order to enhance its capacity towards the development of pharmaceutical applications. Although mainly in watercress flowers and leaves (edible parts) there appear to be a plethora of nutrients and phytochemicals, stems (non-edible part) also provides sources of these metabolites, thereby makes it a promising source of waste recycling and valorization in the context of circular economy.

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

Conceptualization, SK, MP. Methodology, SK, KM, MP. Software, SK. Formal analysis, SK, KM, MP. Investigation, SK. Resources, TA, KS, PW, MP. Data curation, SK. Writing - original draft preparation, SK, MP. Writing - review and editing, SK, TA, KS, PW, DT, RF, AP, MP. Supervision, MP. Project administration, MP. Funding acquisition, MP. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a grant provided by the Cyprus Institute of Neurology and Genetics (Telethon Cyprus), Nicosia, Cyprus (MIP).

Conflict of interest

Author TA is employed by The Watercress Company, authors KS and PW are co-founders of Watercress Research Limited.

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.2022.998755/full#supplementary-material

References

Agerbirk, N., De Nicola, G. R., Olsen, C. E., Müller, C., and Iori, R. (2015). Derivatization of isothiocyanates and their reactive adducts for chromatographic analysis. *Phytochemistry* 118, 109–115. doi: 10.1016/j.phytochem.2015.06.004

Andini, S., Araya-Cloutier, C., Sanders, M., and Vincken, J.-P. (2020). Simultaneous analysis of glucosinolates and isothiocyanates by reversed-phase ultra-High-Performance liquid chromatography–electron spray ionization–tandem mass spectrometry. *J. Agric. Food Chem.* 68, 3121–3131. doi: 10.1021/acs.jafc.9b07920

R Statistical Software (2022). *R foundation for statistical computing r version 4.2.0 r: A language and environment for statistical computing* (Vienna, Austria). Available at: https://www.rR-project.org/.

Azarmehr, N., Afshar, P., Moradi, M., Sadeghi, H., Sadeghi, H., Alipoor, B., et al. (2019). Hepatoprotective and antioxidant activity of watercress extract on acetaminophen-induced hepatotoxicity in rats. *Heliyon* 5, e02072–e02072. doi: 10.1016/j.heliyon.2019.e02072

Barba, F. J., Nikmaram, N., Roohinejad, S., Khelfa, A., Zhu, Z., and Koubaa, M. (2016). Bioavailability of glucosinolates and their breakdown products: Impact of processing. *Front. Nutr.* 3. doi: 10.3389/fnut.2016.00024

Bennett, R. N., Mellon, F. A., Botting, N. P., Eagles, J., Rosa, E. A. S., and Williamson, G. (2002). Identification of the major glucosinolate (4-mercaptobutyl glucosinolate) in leaves of *Eruca sativa* l. (salad rocket). *Phytochemistry* 61, 25–30. doi: 10.1016/S0031-9422(02)00203-0

Błoch-Mechkour, A., Bally, T., Sikora, A., Michalski, R., Marcinek, A., and Gębicki, J. (2010). Radicals and radical ions derived from indole, indole-3-carbinol and diindolylmethane. *J. Phys. Chem. A* 114, 6787–6794. doi: 10.1021/jp912121y

Bouwmeester, H., Schuurink, R. C., Bleeker, P. M., and Schiestl, F. (2019). The role of volatiles in plant communication. *Plant J.* 100, 892–907. doi: 10.1111/tpj.14496

Chaudhary, A., Rampal, G., Sharma, U., Thind, T. S., Singh, B., Vig, A., et al. (2012). Anticancer, antioxidant activities and GC-MS analysis of glucosinolates in two cultivars of broccoli. *Med. Chem. Drug Discovery* 2, 30–37.

Dixon, M. J., and Shaw, P. J. (2011). Watercress and water quality: The effect of phenethyl isothiocyanate on the mating behaviour of *Gammarus pulex*. *Int. J. Zool.* 2011, 328749. doi: 10.1155/2011/328749

European Medicine Agents (1995) *Validation of analytical procedures: Text and methodology Q2(R1)*. Available at: https://www.ema.europa.eu/en/documents/scientific.

Förster, N., Ulrichs, C., Schreiner, M., Müller, C. T., and Mewis, I. (2015). Development of a reliable extraction and quantification method for glucosinolates in moringa oleifera. *Food Chem.* 166, 456–464. doi: 10.1016/j.foodchem.2014.06.043

Gao, N., Budhraja, A., Cheng, S., Liu, E. H., Chen, J., Yang, Z., et al. (2011). Phenethyl isothiocyanate exhibits antileukemic activity *in vitro* and *in vivo* by inactivation of akt and activation of JNK pathways. *Cell Death Dis.* 2, 1–9. doi: 10.1038/cddis.2011.22

Gill, C. I. R., Haldar, S., Boyd, L. A., Bennett, R., Whiteford, J., Butler, M., et al. (2007). Watercress supplementation in diet reduces lymphocyte DNA damage and alters blood antioxidant status in healthy adults. *Am. J. Clin. Nutr.* 85, 504–510. doi: 10.1093/ajcn/85.2.504

Hanschen, F. S., and Schreiner, M. (2017). Isothiocyanates, nitriles, and epithionitriles from glucosinolates are affected by genotype and developmental stage in *Brassica oleracea* varieties. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01095

Hanschen, F. S., Yim, B., Winkelmann, T., Smalla, K., and Schreiner, M. (2015). Degradation of biofumigant isothiocyanates and allyl glucosinolate in soil and their effects on the microbial community composition. *PloS One* 10, e0132931. doi: 10.1371/journal.pone.0132931

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

Jeon, J., Bong, S. J., Park, J. S., Park, Y.-K., Arasu, M. V., Al-Dhabi, N. A., et al. (2017). *De novo* transcriptome analysis and glucosinolate profiling in watercress (*Nasturtium officinale* r. br.). *BMC Genomics* 18, 401. doi: 10.1186/s12864-017-3792-5

Ji, Y., and Morris, M. E. (2003). Determination of phenethyl isothiocyanate in human plasma and urine by ammonia derivatization and liquid chromatography–tandem mass spectrometry. *Anal. Biochem.* 323, 39–47. doi: 10.1016/j.ab.2003.08.011

Kessler, D., Diezel, C., Clark, D. G., Colquhoun, T. A., and Baldwin, I. T. (2013). Petunia flowers solve the defence/apparency dilemma of pollinator attraction by deploying complex floral blends. *Ecol. Lett.* 16, 299–306. doi: 10.1111/ele.12038

Kim, S.-H., Sehrawat, A., Sakao, K., Hahm, E.-R., and Singh, S. V. (2011). Notch activation by phenethyl isothiocyanate attenuates its inhibitory effect on prostate cancer cell migration. *PloS One* 6, e26615. doi: 10.1371/journal.pone.0026615

Kyriakou, S., Trafalis, D. T., Deligiorgi, M. V., Franco, R., Pappa, A., Panayiotidis,, et al. (2022a). Assessment of methodological pipelines for the determination of isothiocyanates derived from natural sources. *Antioxidants* 11:1–32. doi: 10.3390/antiox11040642

Kyriakou, S., Tragkola, V., Alghol, H., Anestopoulos, I., Amery, T., Stewart, K., et al. (2022b). Evaluation of bioactive properties of lipophilic fractions of edible and non-edible parts of *Nasturtium officinale* (Watercress) in a model of human malignant melanoma cells. *Pharmaceuticals* 15, 1–21. doi: 10.3390/ph15020141

Ma, X., Ding, Q., Hou, X., and You, X. (2021). Analysis of flavonoid metabolites in watercress (*Nasturtium officinale* r. br.) and the non-heading Chinese cabbage (*Brassica rapa* ssp. chinensis cv. aijiaohuang) using UHPLC-ESI-MS/MS. *Molecules* 26:1–12. doi: 10.3390/molecules26195825

Mitsiogianni, M., Anestopoulos, I., Kyriakou, S., Trafalis, D. T., Franco, R., Pappa, A., et al. (2021a). Benzyl and phenethyl isothiocyanates as promising epigenetic drug compounds by modulating histone acetylation and methylation marks in malignant melanoma. *Invest. New Drugs* 39, 1460–1468. doi: 10.1007/s10637-021-01127-0

Mitsiogianni, M., Kyriakou, S., Anestopoulos, I., Trafalis, D. T., Deligiorgi, M. V., Franco, R., et al. (2021b). An evaluation of the anti-carcinogenic response of major isothiocyanates in non-metastatic and metastatic melanoma cells. *Antioxidants* 10, 1–14. doi: 10.3390/antiox10020284

Nacca, F., Cozzolino, C., Carillo, P., Woodrow, P., Fuggi, A., and Ciarmiello, L. F. (2021). An hplc-automated derivatization for glutathione and related thiols analysis in *Brassica rapa* l. *Agronomy* 11:1–13. doi: 10.3390/agronomy11061157

Palliyaguru, D. L., Yuan, J.-M., Kensler, T. W., and Fahey, J. W. (2018). Isothiocyanates: Translating the power of plants to people. *Mol. Nutr. Food Res.* 62, e1700965. doi: 10.1002/mnfr.201700965

Panahi Kokhdan, E., Khodabandehloo, H., Ghahremani, H., and Doustimotlagh, A. H. (2021). A narrative review on therapeutic potentials of watercress in human disorders. evidence-based complement. *Altern. Med.* 2021, 5516450. doi: 10.1155/2021/5516450

Pilipczuk, T., Kusznierewicz, B., Chmiel, T., Przychodzeń, W., and Bartoszek, A. (2017). Simultaneous determination of individual isothiocyanates in plant samples by HPLC-DAD-MS following SPE and derivatization with *N*-acetyl-*L*-cysteine. *Food Chem.* 214, 587–596. doi: 10.1016/j.foodchem.2016.07.125

Prieto, M. A., López, C. J., and Simal-Gandara, J. (2019). Glucosinolates: Molecular structure, breakdown, genetic, bioavailability, properties and healthy and adverse effects. *Adv. Food Nutr. Res.* 90, 305–350. doi: 10.1016/bs.afnr.2019.02.008

Ruslin,, Yamin,, Rahma, N. A., Irnawati,, and Rohman, A. (2022). UPLC MS/MS profile and antioxidant activities from nonpolar fraction of patiwala (*Lantana camara*) leaves extract. *Separations* 9, 1–12. doi: 10.3390/separations9030075

Schiestl, F. P. (2014). Correlation analyses between volatiles and glucosinolates show no evidence for chemical defense signaling in *Brassica rapa*. *Front. Ecol. Evol.* 2. doi: 10.3389/fevo.2014.00010

Śmiechowska, A., Bartoszek, A., and Namieśnik, J. (2010). Determination of glucosinolates and their decomposition products–indoles and isothiocyanates in cruciferous vegetables. *Crit. Rev. Anal. Chem.* 40, 202–216. doi: 10.1080/10408347.2010.490489

Song, L., Morrison, J. J., Botting, N. P., and Thornalley, P. J. (2005). Analysis of glucosinolates, isothiocyanates, and amine degradation products in vegetable extracts and blood plasma by LC-MS/MS. *Anal. Biochem.* 347, 234–243. doi: 10.1016/j.ab.2005.09.040

Spínola, V., Pinto, J., and Castilho, P. C. (2017). *In vitro* studies on the effect of watercress juice on digestive enzymes relevant to type 2 diabetes and obesity and antioxidant activity. *J. Food Biochem.* 41, e12335. doi: 10.1111/jfbc.12335

Strauss, S. Y., Irwin, R. E., and Lambrix, V. M. (2004). Optimal defense theory and flower petal colour predict variation in the secondary chemistry of wild radish. *J. Ecol.* 92, 132–141. doi: 10.1111/j.1365-2745.2004.00843.x

Sun, M., Shi, Y., Dang, U. J., and Di Pasqua, A. J. (2019). Phenethyl isothiocyanate and cisplatin Co-encapsulated in a liposomal nanoparticle for treatment of non-small cell lung cancer. *Molecules* 24, 1–13. doi: 10.3390/molecules24040801

Tian, Q., Rosselot, R. A., and Schwartz, S. J. (2005). Quantitative determination of intact glucosinolates in broccoli, broccoli sprouts, Brussels sprouts, and cauliflower by high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry. *Anal. Biochem.* 343, 93–99. doi: 10.1016/j.ab.2005.04.045

Wang, J., Yu, H., Zhao, Z., Sheng, X., Shen, Y., and Gu, H. (2019). Natural variation of glucosinolates and their breakdown products in broccoli (*Brassica oleracea* var. italica) seeds. *J. Agric. Food Chem.* 67, 12528–12537. doi: 10.1021/acs.jafc.9b06533

Wentzell, A. M., and Kliebenstein, D. J. (2008). Genotype, age, tissue, and environment regulate the structural outcome of glucosinolate activation. *Plant Physiol.* 147, 415–428. doi: 10.1104/pp.107.115279

Wu, W., Chen, J., Yu, D., Chen, S., Ye, X., and Zhang, Z. (2021). Analysis of processing effects on glucosinolate profiles in red cabbage by LC-MS/MS in multiple reaction monitoring mode. *Molecules* 26, 1–11. doi: 10.3390/molecules26175171

Yu, X., He, H., zhao, X., Liu, G., Hu, L., Cheng, B., and Wang, Y. (2022). Determination of 18 intact glucosinolates in brassicaceae vegetables by UHPLC-MS/MS: Comparing tissue disruption methods for sample preparation. *Molecules* 27(231), 1–17. doi: org/10.3390/molecules27010231

Zeb, A. (2015). Phenolic profile and antioxidant potential of wild watercress (*Nasturtium officinale* l.). *Springerplus* 4, 714–721. doi: 10.1186/s40064-015-1514-5

Zhu, Y., Yin, Q., and Yang, Y. (2020). Comprehensive investigation of *Moringa oleifera* from different regions by simultaneous determination of 11 polyphenols using UPLC-ESI-MS/MS. *Molecules* 25:1–15. doi: 10.3390/molecules25030676

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 05 December 2022
DOI 10.3389/fpls.2022.1021907

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Sumit Ghosh,
Council of Scientific and Industrial Research (CSIR), India
Hikaru Seki,
Osaka University, Japan

*CORRESPONDENCE
Søren Bak
bak@plen.ku.dk

†PRESENT ADDRESS
Aldo Almeida,
Carlsberg Research Laboratory,
Copenhagen, Denmark

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

RECEIVED 17 August 2022
ACCEPTED 26 October 2022
PUBLISHED 05 December 2022

CITATION
Almeida A, Dong L, Thorsen TH, Raadam MH, Khakimov B, Carreno-Quintero N, Kampranis SC and Bak S (2022) Metabolic engineering of cucurbitacins in *Cucurbita pepo* hairy roots. *Front. Plant Sci.* 13:1021907. doi: 10.3389/fpls.2022.1021907

Metabolic engineering of cucurbitacins in *Cucurbita pepo* hairy roots

Aldo Almeida[1†], Lemeng Dong[2], Theis H. Thorsen[1], Morten H. Raadam[1], Bekzod Khakimov[3], Natalia Carreno-Quintero[4], Sotirios C. Kampranis[1] and Søren Bak[1*]

[1]Department of Plant and Environmental Sciences, University of Copenhagen, Frederiksberg, Denmark, [2]Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, Netherlands, [3]Department of Food Science, University of Copenhagen, Frederiksberg, Denmark, [4]Innovation for Crops, Keygene, N.V., Wageningen, Netherlands

In this paper we show that metabolic engineering in *Cucurbita pepo* hairy roots can be used to both effectively increase and modify cucurbitacins. Cucurbitacins are highly-oxygenated triterpenoids originally described in the Cucurbitaceae family, but have since been found in 15 taxonomically distant plant families. Cucurbitacin B, D, E and I are the most widespread amongst the Cucurbitaceae and they have both important biological and pharmacological activities. In this study *C. pepo* hairy roots were used as a platform to boost production and alter the structures of the afore mentioned cucurbitacins by metabolic engineering to potentially provide new or more desirable bioactivities. We report that the ability to induce cucurbitacin biosynthesis by basic Helix-Loop-Helix transcription factors is partially conserved within the Cucurbitaceae and therefore can potentially be used as a biotechnological tool to increase cucurbitacins in several genera of this family. Additionally, overexpression of a novel acyltransferase from cucurbitacin producing *Iberis amara* generates a hitherto undescribed acetylation at the C3-hydroxyl group of the cucurbitadienol backbone. While overexpression of the cytochromes P450 *CsCYP88L2* and *McCYP88L7* from *Cucumis sativus* and *Momordica charantia* (respectively), results in accumulation of new spectral feature as revealed by High resolution liquid chromatography mass spectroscopy analysis; the m/z of the new peak supports it might be a cucurbitacin hydroxylated at the C19 position in *C. pepo* hairy roots. Finally, this paper is a case study of how hairy roots can be used to metabolically engineer and introduce novel modifications in metabolic pathways that have not been fully elucidated.

KEYWORDS

Rhizobium rhizogenes, triterpenoids, Acyl transferase, P450, Iberis amara, Cucurbita pepo, Ecballium elaterium, Cucumis sativus

10.3389/fpls.2022.1021907

1 Introduction

The aim of this paper was to explore the potential of hairy roots as a production platform for cucurbitacins using metabolic engineering mediated by overexpression of transcription factors and candidate genes. Cucurbitacins are plant specialized metabolites initially developed in the plant for herbivore deterrence (Ferguson and Metcalf, 1985; Dinan et al., 1997), but they have also been shown to have important pharmaceutical bioactivities such as anti-cancer, anti-inflammatory and anti-diabetic activities (Tan et al., 2008; Abdelwahab et al., 2011; Rios et al., 2012). Structurally they are a diverse type of triterpenoids; they are categorized into 17 groups, named from A to T, according to their great variety of functional groups and side-chains (Kaushik et al., 2015). It is this diversity in cucurbitacin structures that impact functionality, and it has been reported that even single modification on a triterpenoid backbone can have great impact on the structure-activity relationships (Sun et al., 2005; Almeida et al., 2020). Cucurbitacins were initially discovered in the Cucurbitaceae family but have now been shown to exist in 15 taxonomically distant families of plants, including the brassicaceae *Iberis amara* (Chen et al., 2005), indicating that the ability to biosynthesize cucurbitacins has evolved more than once in the plant kingdom (Dong et al., 2021). Cucurbitacins may also affect actin polymerization, accordingly they are also cytotoxic to normal cells, and excess amounts cause vomiting in mammals (Wattj and Breyer-brandwijk, 1962), which limits their pharmacological use. Accordingly, introduction of novel modifications in cucurbitacins could generate bioactive cucurbitacins with novel or improved pharmacological properties; however, this approach is limited due to only partial elucidation of genes encoding for enzymes in the cucurbitacin pathway.

Cucurbitadienol is the backbone scaffold of cucurbitacins and can be oxidized at several positions, often multiple times, leading to highly oxygenated triterpenoid structures. Cucurbitacin B (CucB), D (CucD), E (CucE) & I (CucI) are the most widespread amongst the Cucurbitaceae (Figure 1) (David and Vallance, 1955; Rehm et al., 1957). Due to the structural complexity of cucurbitacins, their biosynthetic pathway after the initial cyclization step has only been partially elucidated. Five steps in the Cuc E pathway have been identified in members of the Benincaseae tribe, one of the 15 cucurbitaceous tribes, and found to reside in a metabolic gene cluster (Shang et al., 2014) (Figure 2). Four of the characterized steps are performed by Cytochromes P450 (P450s): CYP88L2 hydroxylates the C19 position of cucurbitadienol, while CYP87D20 oxidizes the C20 and C11 positions by adding a hydroxyl and carbonyl group to these carbons, respectively; CYP81Q58 hydroxylates the C25 position of cucurbitadienol and they CYP81Q59, while inactive in cucumber, paralogs of this P450 in other members of the Benincaseae tribe hydroxylate the C2 position (Figure 2A). The final characterized gene in the cucurbitacin E pathway encodes an acyl transferase which acetylates the C25 hydroxyl group performed by the CYP81Q58 (Figure 2A). Despite these advances, at least four steps in the cucurbitacin E pathway remain uncharacterized: the

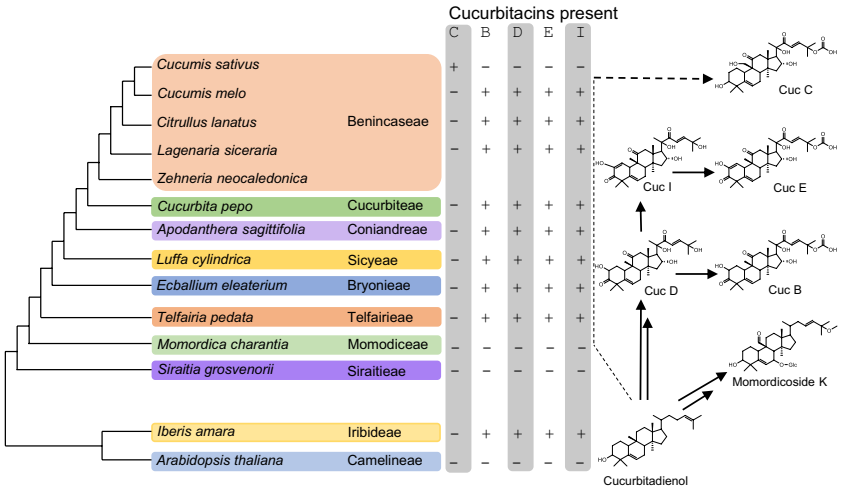


FIGURE 1
Cucurbitacins B, D, E and I are the most widespread amongst the Cucurbitaceae. On the left is a representation of the phylogenetic relationships within the Cucurbitaceae and Brassicaceae families. The presence of major types of cucurbitacins in each tribe is indicated by a + sign whilst the absence by a – sign. The right shows that cucurbitacins B, D, E and I first appear in members of the Telfairieae up to the Benincaseae tribe, but are absent in the Momodiceae tribe and *Cucumis sativus*. Within Brassicecea only *Iberis amara* is known to accumulate cucurbitacins.

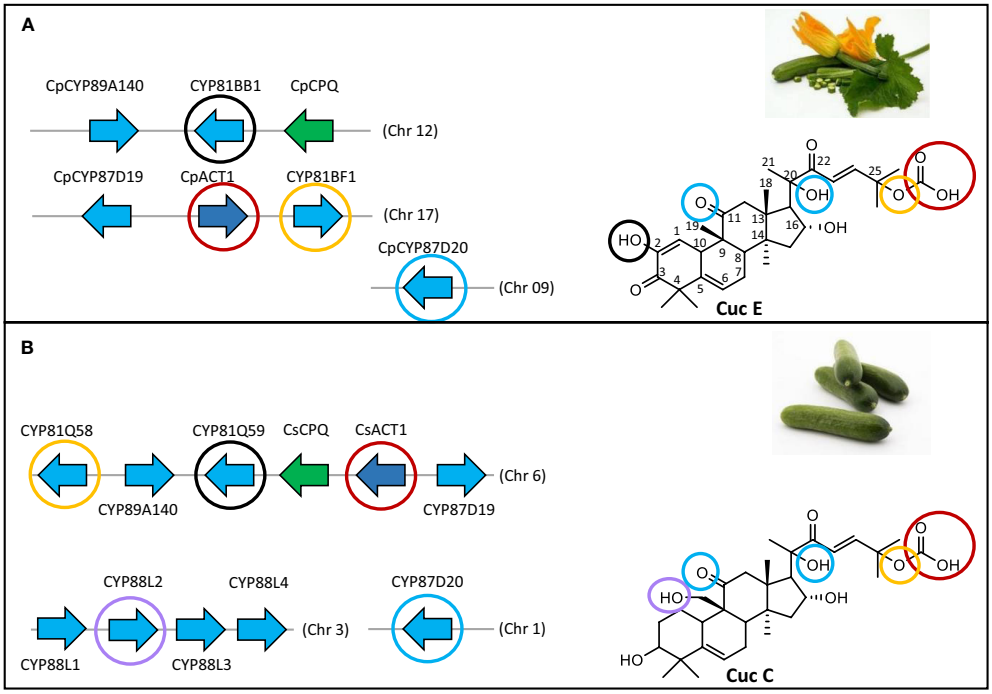


FIGURE 2
The main cucurbitacin biosynthetic cluster in *Cucurbita pepo* is split in two chromosomes. **(A)** Chromosomal organization of cucurbitacin biosynthetic genes in *C. pepo*. Modifications on the cucurbitadienol backbone performed by each gene are highlighted in distinct colored circles. **(B)** Chromosomal organization of cucurbitacin biosynthetic genes in *Cucumis sativus*. Modifications on the cucurbitadienol backbone performed by each gene are highlighted in distinct colored circles. CYP81Q58 catalyzes the C25 hydroxylation; CYP81Q59 the C2 hydroxylation; CpCPQ the first committed step in cucurbitacin biosynthesis making cucurbitadienol backbone; ACT1 transfer an acetyl group to the C25 hydroxyl, CYP87D20 is a C11 carboxylase and hydroxylates the C20 position, CYP88L2 hydroxylates the C19 position.

oxidation of the C3 hydroxyl to a carbonyl, introduction of the carbonyl at the C22 position, hydroxylation of the C16 position and formation of the C1-C2 double bond (Figure 2A). Furthermore, six orthologous basic helix-loop-helix (bHLH) transcription factors (TFs) that induced the expression of Cucurbitacins biosynthetic genes in the Benincaseae tribe in an organ-specific manner have been described (Shang et al., 2014; Zhou et al., 2016); these bHLH TFs are differentially expressed in different organs. Additionally, cucurbitacin profiles differ within species of the Cucurbitaceae; i.e, Cucurbitacins from *Momordica charantia* (Momordica) are not hydroxylated at the cucurbitadienol C16 position, but are oxidized at the C19 position. The C19 hydroxylation is not observed in species of the Cucurbiteae like *Cucurbita pepo* (squash) nor in most species of the Benincaseae, with the exception of *Cucumis sativus* (cucumber) (Figure 1). We recently showed that the cucurbitacin pathway evolved independently in *Iberis amara* of the Brassicaceae (Dong et al., 2021). This convergent evolution extends the existing limited molecular toolbox of genes from Curcurbitaceae species for metabolic engineering of cucurbitacins.

Many Cucurbitaceae species accumulate cucurbitacins in the roots, accordingly, this study explored if hairy roots are suitable targets for cucurbitacin biosynthesis induction and metabolic engineering. Hairy roots (HRs) are induced when *Rhizobium rhizogenes* transforms a set of open reading frames (ORFs) and root loci (rol) genes contained in the hairy root inducing plasmid into a wounded plant cell; the ORFs/rol genes participate in the proliferation of neoplastic roots and root hairs. Aside from the advantage of growing on phytohormone-free media, HRs are considered stable, with the capacity to grow indefinitely and thus can be contained and potentially scaled up. In addition, the use of binary vector systems can take advantage of the *R. rhizobium* DNA-integration mechanism to also deliver genes of interest into the genome of the host plant. Therefore, HRs can be especially useful for metabolic engineering of phytochemicals located in the roots for which the entire pathway has not been elucidated, such as for cucurbitacins.

We have previously shown that HRs from squash can be developed on agar plates and accumulate cucurbitacins. Furhtermore, we have reported that the approach of overexpressing squalene epoxidases e.g. CpSQE2 and OsSQE2

from squash and *Ononis spinosa* can increase the production of cucurbitacins and α-onocerin, respectively, by increasing abundance and availability of the precursor 2,3-oxidosqualene (Almeida et al., 2018; Dong et al., 2018). In this paper, squash HRs were established as a platform for metabolic engineering of cucurbitacins. This paper also reports that the ability to induce cucurbitacin biosynthesis by basic Helix-Loop-Helix transcription factors is partially conserved in the Cucurbitaceae and serves as a cucurbitacin-metabolic-engineering switch for this family, that overexpression of a novel acyltransferase from *I. amara* generates an undescribed acetylation at the C3-hydroxyl group of the cucurbitadienol backbone, and that overexpression of *CYP88L2* and *CYP88L7* from *Cucumis sativus* and *Momordica charantia*, respectively, results in accumulation of cucurbitacins hydroxylated at the C19 postion which is not normally observed in squash.

2 Materials and methods

2.1 Cloning of candidate genes

Total RNA was isolated from *Cucurbita pepo* cultivar golden glory and *Cucumis sativus* var. beit alpha ten-day-old roots using Spectrum Plant Total RNA Kit (Sigma-Aldrich). First-strand cDNA was synthesized from total RNA of roots using the GoScriptTM Reverse transcriptase kit (Promega) following the manufacture's instructions.

Squash orthologs of genes in the cucurbitacin cluster and pathway were searched for in the Cucurbit Genomics Database (http://cucurbitgenomics.org/) by BLASTP analysis using cucumber sequences as query. Annotations of the genomic region were corrected using FgenesH gene-finder (Solovyev et al., 2006).

Primers designed for cloning candidate genes are given in Table S1. Amplification was done by polymerase chain reaction (PCR) in a SensoQuest Labcycler using homemade Phusion X7 polymerase. The following program was applied: 95°C for 1 min; 30 cycles of 95°C for 20 s, 50°C for 30 s, 72°C for 1 min and 30 s; 72°C for 5 min (elongation time for CpCPQ was increased to 2 min). Gel purified amplicons were ligated to the respective cloning system: i) Gateway cloned into pDONR 207 (Invitrogen) and subsequently to the pJCV51 binary vector (VIB-UGent Center for Plant Systems Biology, Vector ID: 4_60) ii) the USER modified binary vector pEAQ : HT-DEST (ref) by USER™ Enzyme (New England Biolabs) iii) pJET 1.2/blunt vector from CloneJet PCR cloning kit (Thermo Fisher Scientific). Ligation reactions were transformed into *E. coli* competent cells using the heat shock method and successful constructs were retrieved 16 hours after. Sequences of constructs were confirmed by Sanger sequencing service (Macrogen, Europe). PJCV51 empty vector from Dong et al. (2018) was used in this study.

Two uracil-specific excision-reagent (USER) vectors with uracil, tryptophan, leucine, or histidine as auxotrophic markers were used for genomic integration. The expression plasmids pESC-URA-USER (pCfB132), pESC-LEU-USER (pCfB220), and pESC-HIS-USER (pCfB291) were gift from Prof. Irina Borodina (Technical University of Denmark, Kongens Lyngby, Denmark) and pESC-TRP-USER (pWUS) was constructed by replacement of the URA3 gene with TRP1 on pCfB132. In addition, the synthetic fragments bear flanking regions containing specific restriction sites and a generic sequence compatible for USER cloning. By applying the USER cloning technique, one or two genes of interest in a uni- or bidirectional promoter of choice were simultaneously integrated into the backbone vectors by specific annealing of the overhang created between the PCR parts and digested USER cassette (Geu-Flores et al., 2007). The CsCPR and Cp P450 genes were amplified and cloned by USER cloning into the yeast expression vectors using primers listed in Table S1 to introduce compatible USER overhangs with the corresponding promoter part (Table S2) and the digested AsiSI/Nb.BsmI USER cassette of the pESC-URA/LEU/HIS/TRP-USER backbone.

Yeast integrative vectors were constructed in *Escherichia coli* cells (E. cloni®; Lucigen) utilizing USER cloning technique. Isolated ORFs of squash genes in the cucurbitacin cluster and pathway and *S. cerevisiae* tHMG1, ERG1, ERG12 and ERG20 were amplified by PCR using primers given in Table S1. Three integrative vectors (pAS1_X-3, pAS2, pAS3_X-3) (Table S2) were used to insert Oxidosqualene cyclases (OSCs) and terpene pathway boosting genes on site 3 of chromosome X, a region known for high gene expression with low to no impact on growth rate (Mikkelsen et al., 2012). Empty integration vectors were kindly provided by Victor Forman, University of Copenhagen, Denmark. pAS1_X-3 contains an ampicillin resistance gene (AmpR), a Kluyveromyces lactis URA3 gene and a region homologous to the insertion site, X3-DOWN, and a homologous region to pAS2. pAS2 contains homologous regions to pAS1_X-3 and pAS3_X-3. pAS3_X-3 contains homologous regions to pAS2 and the integration site X3-UP. PCR fragments were USER cloned into AsiSI/Nb.BsmI digested empty vectors with PCR amplified GAL1/GAL10 promoters from *S. cerevisiae, S. mikae* and *S. arboricola*, to form pInt1-X (Table S2). All generated vectors were confirmed by sequencing.

Prior to yeast transformation, integrative vectors were linearized by digestion with *Not*I according to manufacturer's protocol.

2.2 Hairy root transformation

A list of the plant species used and their providers can be found in Table S3.

Seeds for *Cucurbita pepo* cultivar golden glory and *Iberis umbellata* were first sterilized in a 50-ml falcon tube by washing in 30 mL of deionized water with 300 µL of Mild Cream Soap

(ABENA, Denmark) and shaken for 20 min. Seeds were then rinsed in 70% ethanol for 1 min. Afterwards, seeds were submerged in 1.5% NaOCl and shaken for no more than 15 min. Subsequently, the seeds were rinsed in sterile water 6 times. The wet seeds were placed on a sterile petri dish in the sterile bench to surface dry. When the seeds were surface dried, seeds were placed on plates with ½ Murashige and Skoog media (MS) agar supplemented with 3% glucose sealed with parafilm (Merck KGaA, Darmstadt, Germany) and placed in a climate chamber. The climate chamber was from Fitotron type SGC120-H from 2016 (Weiss Technik UK Ltd., Epinal Way, Loughborough, United Kingdom) and the settings were: light at 79 uE, 25°C, 80% humidity, 16-hour photoperiod.

Rizhobium rhizogenes strain LBA9402 was transformed with pJCV51, pJCV51:*CpCUCbH1* through electroporation and recovered on S.O.C medium for 2 hours at 28°C before being plated on YEB agar supplemented with spectinomycin (50 μg/ml) and rifampicin (50 μg/ml). A single colony was selected for each transformation and grown overnight in liquid YEB supplemented with spectinomycin and rifampicin at 28°C under constant agitation (180 rpm). Subsequently, 50 μL of the *R. rhizogenes* suspension was placed on solid YEB medium supplemented with spectinomycin and rifampicin, and grown for two days at 28°C and colonies resuspended into PS buffer [10 mM PIPES (Sigma-Aldrich P8203)/KOH pH 6.8, 200 mM Sorbitol] supplemented with acetosyringone to a concentration of 100 μM. Finally the OD_{600} of the suspension was adjusted to 0.4.

For squash, cotyledons were transformed with *R. rhizogenes* ten days after germination. Cotyledons were inoculated by bruising with a sterile syringe needle dipped in the *R. rhizogenes* suspension. Four incisions were cut on the abaxial side of the cotyledon perpendicular to the central vein and where placed with the abaxial side down onto ½ MS (3% glucose) agar plates without antibiotics and incubated in the dark for two days. Afterwards, the inoculated cotyledons were transferred to ½ MS (3% glucose) agar plates supplemented with cefotaxim (500 μg/ml), carbenicillin (250 μg/ml) and kanamycin (50 μg/ml) and incubated under light for a week, cotyledons were subsequently transferred to new plates where cefotaxim and carbenicillin concentrations were halfed. Transformed hairy roots expressing mRFP would start to emerge after three weeks of tissue culture. These roots were excised from leaves and subcultured every two weeks with cefotaxim and carbenicillin and kanamycin. The concentrations of cefotaxim and carbenicillin were reduced 50% in each subculturing step until no antibiotic was used, nevertheless kanamycin concentration remained constant.

In the case of *I. umbellata* optimal age for transformation was optimized to be 5 weeks after germination for leaves and stems. For the leaves a similar procedure to that of squash was used, but stems were inoculated by cutting them transversally with a sterile razor blade dipped in the *R. rhizogenes* suspension. Subsequent cultivation were performed the same way as described above for squash cotyledons. All incubation periods were done in the climate chamber.

2.3 Cultivation in temporary immersion reactors (TIR) and methyl jasmonate (MeJA) treatment

Squash HRs were grown 2 weeks in MS agar plates before being transferred to temporary immersion bioreactors (CIRAD Ltd., France) which were previoiusly sterilized by autoclaving at 121°C for 15 min. The TIR were filled with 400 ml of ½ MS media, HRs were flooded for intervals of 30 min every 8 hours for a period of 14 days. The flow rate of inlet air was 60 l/h.

For MeJ treatment, a 100mM stock solution of MeJ of 95% purity (Sigma Aldrich) was prepared in 96% ethanol and filter sterilized. After incubating HRs for 14 days in TIR, 400 μl of the MeJA stock solution was added to make a ∼100 μM MeJA treatment. In the same manner, 30 μl of MeJA sterilized stock solution were pipetted in the ½ MS agar near 14-day-old HR (square plates typically had 30 ml of agar). Samples were collected after 24 hours of adding MeJA, frozen in liquid nitrogen and stored at -70°C until further processing. All experiments in TIRs were carried out in triplicates and statistical significance was tested using one-way ANOVA ($p < 0.05$).

2.4 Transient expression experiments

pEAQ-HT-DEST expression vector (Sainsbury et al., 2009) constructs harboring *eGFP* and *CpCUCbH1* (described above) were transformed into *Rhizobium (Agrobacterium) tumefaciens* strain AGL1. Colonies of *A. tumefaciens* were picked in the morning and precultured in 5 ml of Luria-Bertani (LB) media supplemented with kanamycin (50 μg/ml). Afterward, 12 ml of LB media containing kanamycin was inoculated with 50 μL of the preculture and incubated at 28°C overnight without allowing OD_{600} to exceed 1.5. The cultures were subsequently centrifuged at 3000g, the resulting cell pellet resuspended in infiltration buffer (10 mM MgCl2, 10 mM MES, pH 5.6, and 100 μM acetosyringone), and the OD_{600} was adjusted to 0.4. The suspension was incubated for 1 h on a shaker at room temperature. Age of infiltration for tissues of each species was optimized: for cotyledons of squash and cucumber were infiltrated 5 days after germination and examined 5 days after infiltration, whereas cotyledons from Luffa were infiltrated 7 days after germination and examined 7 days after infiltration. In the case of *E. elaterium* leaves of 2-month-old plants were infiltrated and examined 5 days after infiltration, while for *Nicotiana benthamiana* (tabacco) leaves were infiltrated when plants were 5-weeks old and harverted 5 days after infiltration.

For infiltration experiments plants were grown in soil (Pindstrup substrate no. 2) in a glasshouse with a 16-h day at 28°C and an 8-h night at 28°C. Infiltrated tissue for confocal microscopy was harvested the same day while for metabolite analysis tissues were frozen in liquid N_2 and stored at −70°C until further processing. All experiments were carried out in triplicates and statistical significance was tested using one-way ANOVA ($p < 0.05$).

2.5 Gene expression analysis by qRT-PCR

Gel electrophoresis confirmed that the primers designed for qRT-PCR (Table S4) amplify only one amplicon. Each tissue was analyzed in triplicate, synthesizing cDNA from 1 µg of total RNA using the GoScript™ Reverse transcriptase kit (Promega), and final volume was adjusted to 10 ng/µl. All quantitative PCRs for the three primer sets were performed in the CFX384 real-time system (Bio-Rad) under the following conditions: 4 µL of PowerUp SYBR Green Master Mix (Thermo Fisher Scientific), 1 µL of forward primer (2 µm), 1 µL of reverse primer (2 µm), 1 µL of cDNA (10 ng/µl), and 1 µl of MilliQ water. The PCR conditions were initial 95°C for 5 min, followed by 40 cycles of 95°C for 15 s, 50°C for 15 s, and 72°C for 30 s.

2.6 Visualization of transient-expression efficiency

For visualization of *in planta* transient expression, confocal Laser Scanning Microscopy (CLSM) was done essentially as previously described (Laursen et al., 2016). Briefly, leaf discs from infiltrated tissue described above were excised, mounted in water, and observed by CLSM. Cell imaging was performed using an SP5x laser scanning confocal microscope equipped with a DM6000 microscope (Leica, Germany). Images were recorded using a 63x water immersion objective lens. Excitation/emission wavelengths were 488/500–550 nm for eGFP. The images were sequentially acquired and processed using the LAS X software (Leica).

2.7 Heterologous expression in *Saccharomyces cerevisiae*

The yeast strain EGY48 (Thomas and Rothstein, 1989; Ellerström et al., 1992; Ignea et al., 2011) was transformed with yeast integrative vectors using the lithium acetate method (Gietz and Schiestl, 2007), generating strains EGY48-EV, EGY48-CpCPQ. Subsequently, EGY48-CpCPQ was co-transformed with plasmids pCfB220 containing the *Cucumis sativus* Cytochrome P450 Reductase (pCfB220:CsCPR) and pCfB291 containing one of the following P450s: CpCYP88L4, CYP88L2, CYP88L7, CpCYP87D19, CpCYP89A140; a full list of the strains generated is presented in Table S4 and the description of the plasmids used is in Table S2. Colonies following transformation were genotyped for correct integration using PCR. For triterpenoid production, pre-cultures of the strains were grown overnight until saturation in 10 ml synthetic defined media containing 2% glucose. Then, the cells were washed twice with water and heterologous genes induced by inoculation into 10 mL synthetic defined media with 4% galactose and 2% raffinose with or without 10 mM β-methyl cyclodextrins and grown for 96 h at 30°C and 150 rpm.

2.8 Phylogenetic analysis

Protein-coding sequences were deduced using the Expasy translate tool (http://web.expasy.org/translate/) and manually curated. Models for the evolutionary histories for the bHLH TFs, CYP88s and ACT protein sequences were inferred separately by using the maximum-likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood was selected (−4579.11 for the bHLH, -6,515.95 for the CYP88 and −8,676.26 for the ACT trees, respectively). Initial trees for the heuristic search were obtained by applying the neighbor-joining method to a matrix of pairwise distances estimated using a JTT model. The trees are drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 21, 11 and 14 amino acid sequences for the bHLH TFs, CYP88s and ACT trees, respectively. All positions containing gaps and missing data were eliminated. In total, there were 187, 434 and 366 positions in the final data sets for the bHLH, CYP88 and ACT trees, respectively. The statistical significance of each node was tested by the bootstrap method using 1,000 iterations. The evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016). Accession numbers for sequences used in phylogenetic trees are given in Table S5.

2.9 Cucurbitacin-feeding experiments in *Saccharomyces cerevisiae*

Strains EGY48-CpCPQ+CsCPR+CYP88L2 and EGY48-CpCPQ+CsCPR+CYP88L2 were grown overnight until saturation in 10 ml synthetic defined media containing 2% glucose. Then, the cells were washed twice with water and heterologous genes induced by inoculation into 10 mL synthetic defined media with 4% galactose and 2% raffinose with 10 mM β-methyl cyclodextrins. In addition, 100 µl of either

1mM Cucurbitacin B, D, E or I were added to the cultures and grown for 96 h at 30°C and 150 rpm. Additionally, production of either Cuc B or E by feeding Cuc D or I to a cultures of strain EGY48-CpACT1 grown in a similar fashion as described above was used as controls, to make sure that feeding of yeast with cucurbitacins was indeed possible. All reactions were carried out in triplicates.

2.10 Root phenotyping

WinRHIZO, a root image analysis software (Arsenault et al., 1995) was used for determination of primary root length, number of root tips and total root length phenotypes of HR lines. Pictures for root phenotyping were taken from HRs in agar plates 5 days after sub-culturing using a Nikon D5600 DSLR camara with an AF-P DX 18-55 VR lens (Nikon, Japan).

2.11 Analysis of cucurbitacins by GC-MS

For GC-MS analysis, 50 µL of extract was aliquoted into a glass insert and evaporated under vacuum. The glass inserts were sealed with air-tight magnetic lids into GC-MS vials and derivatized by the addition of 30 µL of trimethylsilyl cyanide as described before (Khakimov et al., 2013). All steps involving sample derivatization and injection were automated using a MultiPurpose Sampler (MPS; Gerstel). After reagent addition, the sample was transferred into the agitator of the MPS and incubated at 40°C for 40 min at 750 rpm. Immediately after derivatization, 1 µL of the derivatized sample was injected in splitless mode. The spilt/splitless injector port was operated at 320°C. The septum purge flow and purge flow to split vent at 2.1 min after injection were set to 3 and 15 mL min−1, respectively. The GC-MS system consisted of an Agilent 7890A GC device and an Agilent 5975C series MSD (Agilent Technologies). GC separation was performed on an Agilent HP-5MS column (30 m × 250 µm × 0.25 µm) by using hydrogen carrier gas at a constant flow rate of 1.2 ml/min. The GC oven temperature program was as follows: initial temperature, 40°C; equilibration time, 2 min; heat up to 270°C at the rate of 12°C/min; heat at the rate of 6°C/min until 310°C; and hold for 10 min. Mass spectra were recorded in the range of 50 to 700 mass-to-charge ratio with a scanning frequency of 2.2 scans/s, and the MS detector was switched off during the first 20 min of the run, since all targeted molecules eluted after this retention time. The transfer line, ion source, and quadrupole temperatures were set to 290°C, 230°C, and 150°C, respectively. The mass spectrometer was tuned according to the manufacturer's recommendation by using perfluorotributylamine. The MPS and GC-MS devices were controlled using vendor software Maestro (Gerstel).

2.12 Analysis of cucurbitacins by LC-QToF-MS

First, 100 mg of frozen powdered material were extracted with 600 µL of methanol (containing 50 µM of protopanaxatriol as internal standard) in a 1.5 mL vials with screw cap. Vials were ultrasonicated for 20 min at room temperature, the vials were then centrifuged for 10 min at 13,000 rpm and the supernatant was filtered through 0.22 µm centrifugal filters (UFC30GV, Merck Millipore). All treatments consisted of at least three replicates (specific replicate numbers are specified in figure legends). Statistical significance was tested usin one-way ANOVA ($p < 0.05$).

For extraction of triterpenoids from yeast cultures incubated with β-methyl cyclodextrin, 10 mL of culture was extracted using 10 mL ethyl acetate. The mixture was vortexed for 2 minutes, then centrifuged at 3,000g for 10 minutes and the organic phase collected, evaporated under vacuum and the residues were resuspended in 200µL of methanol. For extraction of triterpenoids from yeast cultures not containing β-methyl cyclodextrin, the cell pellet from 10 ml culture was resuspended in 500µL 10% KOH, 80% EtOH and saponified for 2 hours at 70°C. The samples were then extracted three times with 500µL hexane, which was pooled, evaporated under vacuum and the residues resuspended in 200µL of Methanol.

A Dionex UltiMate™ 3000 RS UPLC system from Thermo Scientific™ (Waltham, MA USA) was used and equipped with a KINETEX® XB-C18 column (2.1 mm × 100 mm, 1.7 µm, Phenomenex). Mobile phase A was prepared using MilliQ water and 0.05% formic acid, whereas mobile phase B consisted of acetonitrile and 0.05% formic acid. The gradient was t = 0 min, 5% B; t = 40 min, 100% B; t = 45 min, 100% B; t = 46 min, 5% B; t = 50 min, 5% B. The chromatographic run lasted 50 min with a flow rate of 0.3 ml/min. For detection, a Bruker compact™ QTOF mass spectrometer (Bremen, Germany) was used and operated in negative mode. The mass spectrometer's settings were: dry gas flow rate 8 L min−1 at 220°C, capillary 4500 V, collision energy 7 eV and collision RF 500Vpp, transfer energy 100 µs, pre-Pulse storage 5µs. The QTOF operated with the mass range set from 50 to 1200 m/z and was calibrated with sodium formate clusters at the beginning of every injection. The injection volume was 5 µL per sample. Acquisition of LC-MS data was performed under Bruker DataAnalysis 4.3. Quantification of cucurbitacins was performed in more than three lines for each HR construct/treatment. Calibration curves with cucurbitacins B,D,E&I (Extrasynthese, Lyon, France) were made for quantification of cucurbitacin production in hairy roots lines.

Suggested chemical structures for the putative cucurbitacins in Figures 7; S7. have a metabolite identification confidence level of 4, according to the Metabolomics Standards Initiatives

(Sumner et al., 2007; Schrimpe-Rutledge et al., 2016) and was achieved by molecular formula generated from accurate mass spectrometry data (± 5 ppm) and MS/MS fragmentation pattern.

3 Results

3.1 Growth in temporary immersion reactors increases biomass and cucurbitacin content of HR as compared to agar plates and liquid cultures

Our previous work established a HR transformation protocol for Squash (Dong et al., 2018) based on their growth on agar plates. Recently we showed that *I. amara* produce Cucs B,D,E&I in all organs (Dong et al., 2021), suggesting that Iberis species could be an alternative to squash HR for metabolic engineering of cucurbitacins. Hence, in this study a HR transformation protocol for *Iberis umbellata* was developed, to enable a comparison between *I. umbellate* and squash as HR platforms for metabolic engineering of cucurbitacins (Figures 3A-C). Due to the ∼7-fold higher transformation efficiency and shorter time taken to generate analyzable HRs, further work was carried out in squash HRs as opposed to *I. umbellata* HRs (table in Figure S1).

To establish a reproducible and scalable production platform, growth of squash HRs on agar plates, in liquid culture and in temporary immersion reactors (TIRs) was compared. In terms of cucurbitacin production, metabolite profiling by LC-MS showed that accumulation of Cuc E significantly increased (∼2-fold) in empty vector (EV) HR lines grown in TIRs as compared to those in plates but not for Cucs B&I (Figure 3D). Subsequent attempts to increase cucurbitacin yields of HR grown on plates and TIRs involved inducing their production with Methyl-Jasmonate (MeJA). Although the mean values appear higher for HRs grown on plates treated with MeJA (5.78 ± 1.96, 9.16 ± 2.91 and 51.98 ± 12.87 ng/mg DW for Cuc I, Cuc B and Cuc E, respectively) when compared to untreated control (5.63 ± 1.14, 6.82 ± 0.74 and 31.93 ± 2.68 ng/mg DW for Cuc I, Cuc B and Cuc E, respectively), there was no significant difference in cucurbitacin production (Figure 3D). Similarly, addition of MeJA did not significantly elevate cucurbitacin levels in TIRs. In terms of biomass production, HRs on plates grew slowly and were limited in space, whilst growth in liquid media inside Erlen Meyer flasks resulted in a drastic variation of biomass production within lines (Figure S2A-C). In summary, upscaling and reproducible production of biomass as well as higher cucurbitacin yields is best in TIRs (Figure S2D), probably due to facilitated aeriation, and is not affected by addition of MeJA.

3.2 The transcription factor CpCUCbH1 induces cucurbitacin biosynthesis in several Cucurbitaceae clades

Our previous work showed that overexpression of squash squalene epoxidases can be used to increase cucurbitacin levels in both HRs and the tabacco transient-expression systems by increasing the triterpenoid precursor 2,3-oxidosqualene (Dong et al., 2018). In this paper the approach of increasing cucurbitacins through overexpression of transcription factors was attempted.

Previously, a TF named CsBL, belonging to the clade II of the bHLH superfamily was reported to increase cucurbitacin C levels in cucumber leaves. Hence, an orthologous bHLH TF that would induce cucurbitacin accumulation in squash HRs and could be used as a tool in metabolic engineering of cucurbitacins was searched for in squash. BLASTp was used to searched the squash genome for putative orthologous sequences of the CsBL TF. This identified two sequences in the loci Cp4.1LG05g03810.1 and Cp4.1LG09g01220.1 which were 57% and 36% identical at aa level, respectively. These genes were termed *Cucurbitacin inducing bHLH transcription factor 1* (*CpCUCbH1*) and *CpCUCbH1-like*.

A phylogenetic analysis showed that CpCUCbH1 was the most likely putative ortholog of CsBL while CpCUCbH1-like formed a separate clade from these sequences (Figure 3E). Overexpression of *CpCUCbH*1 in squash HR grown on plates increased ∼5-fold the production of Cuc I&B, and ∼3-fold of Cuc E (Figure 3D). To confirm that CpCUCbH1 was the functional ortholog of CsBl, the expression levels of known genes in the cucurbitacin pathway were quantified using RT-qPCR. With the exception of CpCYP87D19, there was a significant increase in the expression of all tested genes in the *CpCUCbH1* overexpressing lines as compared to the empty vector (EV) lines (Figure 4). These results suggest that CpCUCbH1 is indeed an ortholog of CsBl, and that the evolution of specific TFs regulating the production of cucurbitacins occurred at least before the split of the Cucurbitae and Benincasea tribes in the Cucurbitacea family. Interestingly, although growing HR lines in TIRs increased expression of genes in the cucurbitacin cluster and pathway as compared to HRs in agar plates, growth in TIR did not increase expression of CpCUCbH1 (Figure 4A). This suggests there are other mechanisms besides the clade II TFs of the bHLH superfamily which may be involved in the regulation of cucurbitacin biosynthesis which respond to different environmental stimuli.

It is noteworthy that the increase of cucurbitacins in *CpCUCbH1*-overexpressing HR lines was followed by a distinctive HR phenotype of reduced lateral roots, resulting in lines without the typical fish-bone phenotype of HRs, in the end this affected the biomass produced (Figure S3). Hence, propagation efforts to produce enough biomass to explore the

10.3389/fpls.2022.1021907

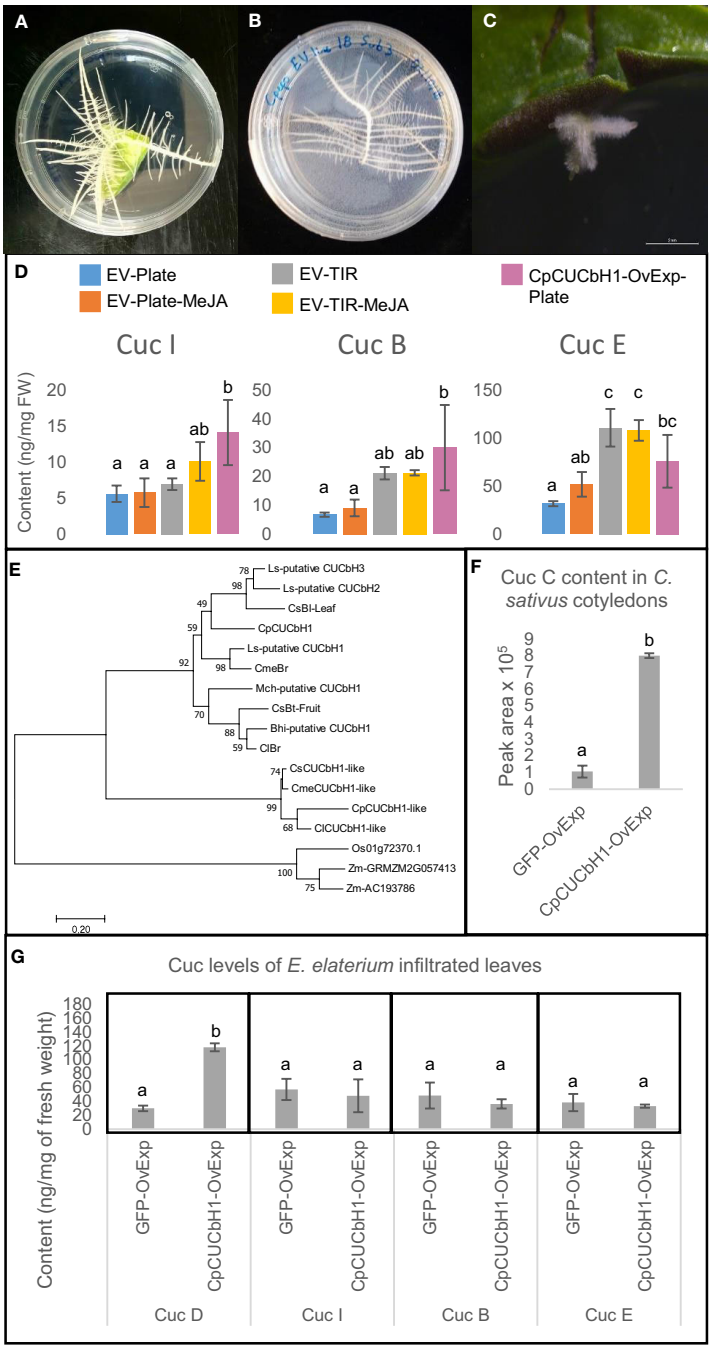


FIGURE 3
Cucurbitacin production in *C.ucurbita pepo* hairy roots can be increased by overexpression of *C.pCUCbH1* and cultivation methods. **(A)** Hairy roots emerging from *C. pepo* cotyledon 2 weeks after transformation. **(B)** *C. pepo* hairy roots propagation on agar plate. **(C)** hairy roots emerging from *Iberis umbellata* are smaller and grow slower than *C. pepo* hairy roots, the picture was taken 8 weeks after transformation. **(D)** Comparison of Cucurbitacin I, B and E content between *C. pepo* hairy root lines overexpressing different gene constructs and grown under different conditions i.e. agar plates (n = 5), temporary immersion reactors (n = 3). CpCUCbH1-OvExp lines were grown in plates (n = 3). **(E)** Phylogenetic tree of basic Helix Loop Helix transcription factors regulating cucurbitacin production in the Cucurbitaceae. Accession numbers for genes are in Table S5. Sequences of bHLH clade II transcription factors from Zea mays and Oriza sativa were used as an outgroup. **(F)** Relative abundance of Cucurbitacin C in *C.ucumis sativus* cotyledons infiltrated with different constructs (n = 3). **(G)** Concentration of Cucurbitacin D, I, B and E in *Ecballium elaterium* leaves infiltrated with constructs overexpressing GFP or CUCbH1 (n = 3). EV- empty vector control, TIR- grown in temporary immersion reactors, MeJA- induced with Methyl jasmonate, CpCUCbH1- samples overexpressing transcription factor CpCUCbH1, GFP- Green Fluorescent Protein control. Error bars are standard deviation. Statistical significance was tested using one-way ANOVA ($p < 0.05$).

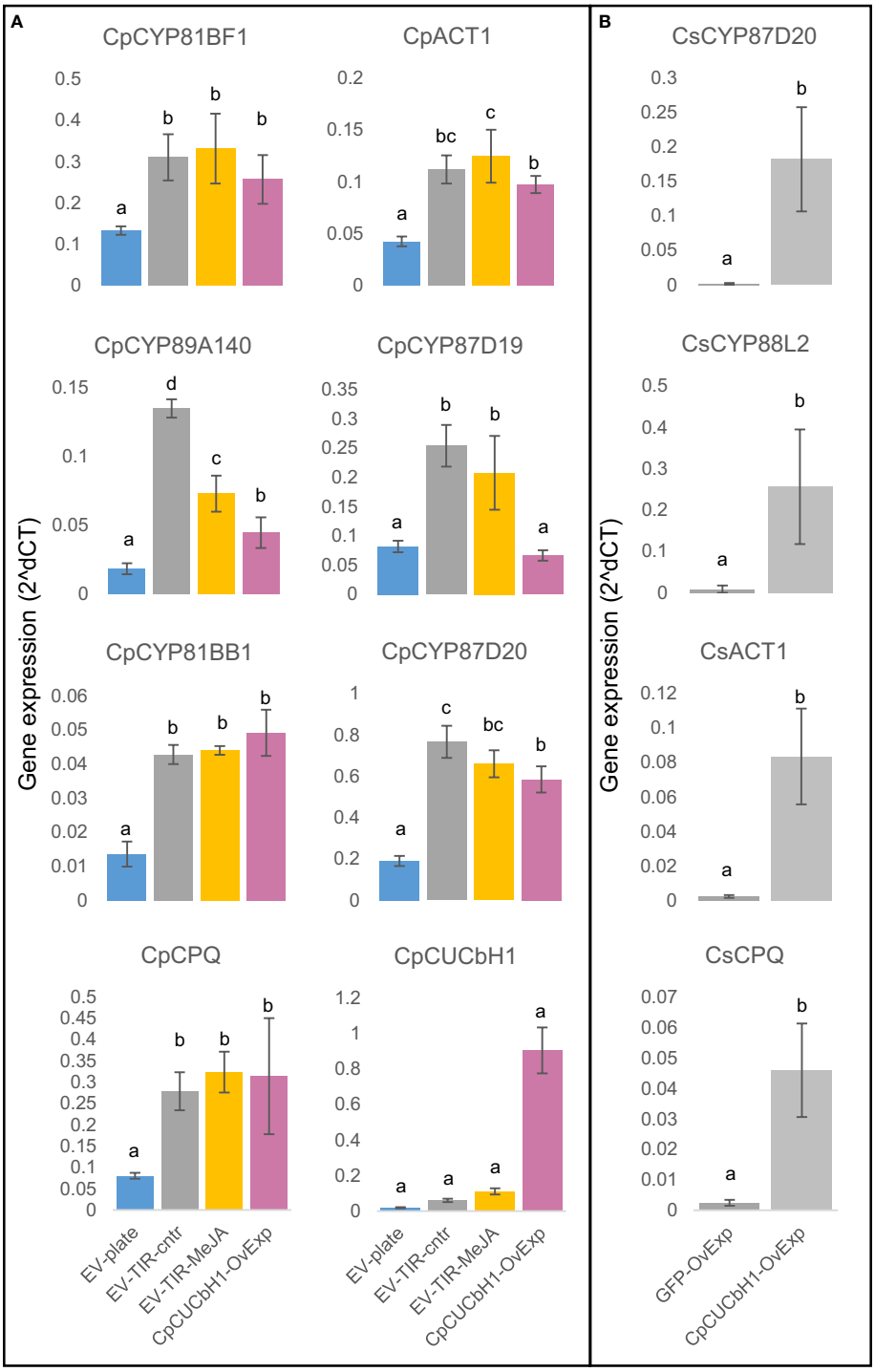


FIGURE 4
Overexpression of *C.pCUCbH1* increases expression of cucurbitacin biosynthetic genes in *C. pepo* hairy roots and *C. sativus* cotyledons. **(A)** Relative expression of selected cucurbitacin biosynthetic genes between *C. pepo* hairy root lines overexpressing *C.pCUCbH1* and grown under different conditions: *C.pCYP81BF1* (*CsCYP81Q58* ortholog *C.pCYP89A140* (*C.sCYP89A140* ortholog), *C.pCYP81BB1* (CsCYP81Q59 ortholog), *C.pCPQ* (*C.sCPQ* ortholog), *C.pACT1* (*C.sACT1* ortholog), *C.pCYP87D19* (*C.sCYP87D19* ortholog), *C.pCYP87D20* (*C.sCYP87D20* ortholog), *C.pCUCbH1* (*C.sBl* ortholog). **(B)** Relative expression for selected cucurbitacin biosynthetic genes between *C. sativus* cotyledons overexpressing different constructs: *C.sCPQ, CsACT1, CsCYP87D20, CsCYP88L2*.

cucurbitacin production of *CpCUCbH1*-overexpressing HR lines in TIRs were not attempted due to the poor growth reported by these lines.

An orthologous bHLH TF sequence of Momordica grouping with CsBl and CpCUCbH1 in the phylogetic tree (Figure 3E) made us suspect that the bHLH induction of cucurbitacins could be widespread in the Cucurbitaceae. To determine if CpCUCbH1 could be used as a biotechnological tool to induce cucurbitacins in different species of the Cucurbitaceae, cotyledons or leaves of several Cucurbitaceae species were transiently infiltrated, and cucurbitacin levels, and expression levels of selected cucurbitacin biosynthetic genes were measured. To test if the squash bHLH could induce cucurbitacin biosynthesis in other Cucurbitaceae species, *CpCUCbH1* was first transiently expressed in cucumber cotyledons. Since genes in the cucurbitacin biosynthesis pathway of cucumber have been reported, their expression could be monitored by qPCR compared to cucurbits in deeper Cucurbitaceae lineages that exhibit a lack of genetic resources. overexpression of CpCUCbH1 resulted in ~8 fold higher accumulation of Cuc C (Figure 3F) and higher expression of genes in the Cuc C pathway (Figure 4B). Having confirmed this, we next tested induction of cucurbitacins in deeper Cucurbitaceae lineages. *Luffa cylindrica* (Luffa) is a member of the Sicyeae tribe which split from the Cucurbitae (Figure 1) 39 ± 3 Million Years ago (MY) (Schaefer et al., 2009) and produces Cucs B,D,E&I in leaves (Rehm et al., 1957); cotyledons of Luffa were successfully infiltrated and constructs were overexpressed, as observed from fluorescence of the GFP control constructs (Figure S4A-D). However, cotyledons of Luffa did not accumulate Cuc B,D,E&I, and overexpression of *CpCUCbH1* did not induce production of cucurbitacins (Figure S4E). This incapacity of CpCUCbH1 to induce cucurbitacin biosynthesis was also observed for squash cotyledons (Figure S4F).

Finally, transient-expression experiments were performed on *Ecballium elaterium* to test if CpCUCbH1 would also increase cucurbitacin production in species of deeper phylogenetic clades in the Cucurbitaceae. *Ecballium elaterium* is a cucurbit which accumulates Cucs B,D,E&I (David and Vallance, 1955) in its leaves and belongs in the Bryonieae tribe, a tribe that split from the Cucurbitae 41 ± 3 Million Years ago (MY) (Schaefer et al., 2009). The transient-infiltration of *E. elaterium* leaves experiments for overexpression of CpCUCbH1 significantly increased levels of Cuc D ~4-fold but not of Cucs B,E&I (Figure 3G). Induction of cucurbitacins in the deeper lineage of the Momordiceae was also attempted, but unfortunately leaves of Momordica proved difficult to infiltrate.

These results suggest that induction of cucurbitacin biosynthesis by clade II bHLH TF may require additional factors and is partially conserved in the Cucurbitaceae. Nevertheless, here it is shown that *CpCUCbH1* overexpression can be used as a tool to induce cucurbitacins in selected clades of this family.

3.3 Modifying cucurbitacins through metabolic engineering

3.3.1 An acyl transferase from *Iberis amara* produces undescribed C3 acetylated cucurbitacins

We previously reported how the independent recruitment from the same gene families, I.e. OSCs and P450, lead to the convergent evolution of the cucurbitacin pathway in *I. amara* (Dong et al., 2021). Thus, we hypothesized that *I. amara* could harbor enzymes making unprecedented modifications to cucurbitacins from cucurbitacous plants. Therefore, to attempt to engineer potentially new cucurbitacins, our previously constructed transcriptome of *I. amara* was mined for candidate genes involved in modification of cucurbitacins, focusing first on acyl tranferases (ACTs). BLASTp was used to search for homologous ACTs in *I. amara* using as query the aa sequence of cucumber ACT1 (CsACT1, Csa6G088700) in the cucurbitacin gene cluster (Shang et al., 2014), as this enzyme acetylates the hydroxyl group at position C25 of Cuc B&I. This returned two sequences (Genbank: MZ695218 and MZ695219, named IaCUCA1 and IaACT2, respectively); IaACT2 showed 36% identity at amino acid level to CsACT1, while IaCUCA1 showed 56%. The function of IaCUCA1 was tested using *Rhizobium* (*Agrobacterium) tumefaciens*-mediated transient expression in tabacco. When IaCUCA1 was coexpressed with *I. amara* cucurbitadienol synthase (IaCPQ), the cucurbitadienol peak lowered in intensity but no new major peak could be detected in GC-MS (Figure 5A). This discrepancy could be interpreted in different ways: i) the result suggests that cucurbitadienol was not the *in planta* substrate for IaCUCA1 or ii) IaCUCA1 did act on cucurbitadienol, but the resulting triterpenoids were converted into other compounds that could not be detected with the GC-MS method used. However, when IaCUCA1 was coexpressed with IaCPQ and CYP708A16 (the *I. amara* P450 that makes 16-β-hydroxy cucurbitadienol) the intensity of the 16-β-hydroxy cucurbitadienol was lowered and a new peak could be detected (Figure 5A). This peak was isolated from infiltrated tobacco leaves using preperative HPLC, and the structure was identified as 3-acetyl-16-β-hydroxy-cucurbitadienol by NMR experiments (Supplementary File 1) (Figure 5). Acetylation at the C3 position of cucurbitacins has not been reported in species of the Cucurbitaceae nor in *I. amara*, making this an unexpected finding. A phylogenetic analysis of selected ACTs showed the two ACTs from *I. amara* are evolutionarily distinct from Cucurbitacea ACTs. IaCUCA1 and IaACT2 appear to have diverged from a common ancestor in the Brassicaceae which also gave rise to the clades containing AtTHAA1 and AtTHAA2 (Figure 5B), which both act on the triterpene thalianol of Arabidopsis (Huang et al., 2019). Thus, IaCUCA1 appears to be an acyl transferase which has a predisposition for triterpenoid scaffolds and could be used to create novel cucurbitacin structures through metabolic engineering.

Because acetylation at the C3 postion of cucbitacins from cucurbitaceous species had not been reported, metabolic

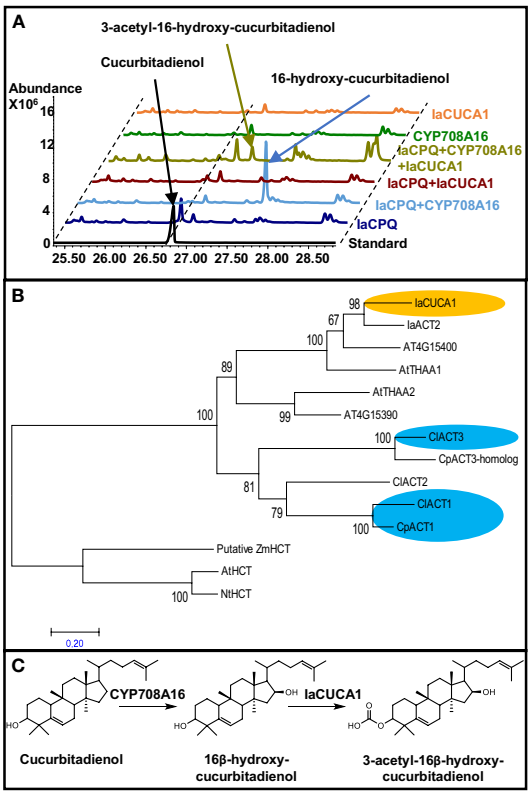


FIGURE 5
An acyltransferase from *Iberis amara* (IaCUCA1) acetylates the C3-hydroxyl position of 16-hydroxy-cucurbitadienol and has a different evolutionary origin than Cucurbitaceae acyl transferases acting on cucurbitacins. **(A)** Total ion chromatograms from GC-MS runs of *Nicotiana benthamiana* leaves infiltrated with different constructs. IaCPQ, *Iberis amara* cucurbitadienol synthase; CYP708A16, *Iberis amara* P450 making 16β-hydroxyl cucurbitadienol; IaCUCA1, *Iberis amara* acyltranferase making 3-O-acetyl-16β-hydroxyl-cucurbitadienol. **(B)** Maximum Likelihood tree constructed on deduced amino acid sequences of 14 acyl transferases aligned by ClustalW spanning 366 positions using the MEGA7 program. The statistical significance of each node was tested by the bootstrap method using 1,000 iterations. Representative names are as follows: IaCUCA1, *Iberis amara* acetyltransferase for the C3-hydroxy group of 16-hydroxy-cucurbitadienol; IaACT2, *Iberis amara* uncharacterized acyl transferase; At4G15400, *Arabidopsis thaliana* BAHD acyl transferase involved in brassinosteroid metabolism; AtTHAA1, *Arabidopsis thaliana* acetyltransferase for C15-hydroxy of 7,15-dihydroxy-16-keto-thalianol; AtTHAA2, *Arabidopsis thaliana* acetyltransferase for C3-hydroxy of different thalianol-derived compounds; At4G15390, *Arabidopsis thaliana* HXXXD-type acyl-transferase family protein; ClACT3, *Citrullus lanatus* acetyltransferase acting on C16-hydroxy of Cucs B,D,E&I; CpACT3, *Cucurbita pepo* homolog of ClACT3; ClACT2, *Citrullus lanatus* uncharacterized acyltransferase; ClACT1, *Citrullus lanatus* acetyltransferase acting on C25 hydroxyl of Cucs D&I; CpACT1, *Cucurbita pepo* functional homolog of ClACT1; the following sequences were used as an outgroup: AtHCT, Arabidopsis thaliana hydroxycinnamoyl transferase; NtHCT, Nicotiana tabacum hydroxycinnamoyl transferase; ZmHCT, Zea mays putative hydroxycinnamoyl transferase. Accession number of sequences are given in Table S5. **(C)** Structure and proposed biosynthesis of 3-acetyl-16-hydroxy cucurbitadienol.

engineering of cucurbitacins by overexpressing *IaCUCA1* in squash HR was attempted. Unexpectedly, only a single HR line was obtained from 120 cotyledons, which is a 0.8% transformation efficiency compared to the ~50% efficiency normally observed for squash (Figure S1); in addition, this root stopped growing before being large enough for further analysis by LCMS. This suggests that IaACT is making toxic compounds that are detrimental to squash HR growth.

3.3.2 Modification of Cucurbitacins by overexpression of genes in the CYP88L subfamily

Enzymes performing the C19 hydroxylation on cucurbitadienol have only been characterized in cucumber and Momordica catalyzed by the P450s CYP88L2 (Shang et al., 2014) and CYP88L7 (Takase et al., 2019), respectively. Despite squash being phylogenetically closer to cucumber than Momordica, there are no reports of cucurbitacins with a C19 hydroxylation in squash. To investigate if the absence of C19 hydroxylated cucurbitacins in squash was due to a lack of a CYP88L2 functional ortholog, we collected sequences and built a phylogenetic tree with CYP88 aa sequences from Cucurbitaceae species. Through this work we identified misannotations and mislabels in genomes of the cucurbit genomic database and previous papers (Figure S5). The reannotation of the 24kb long Cp4.1LG15g03520 gene model of squash containing homologous sequence to *CsCYP88L2*, resolved three CsCYP88L2 paralogs termed CpCYP88L4 (Genbank: MZ695222), and the probable pseudogenes CYP88L9P and CYP88L10P (Figure S5). The phylogenetic analysis revealed that CYP88s from the Cucurbiteae tribe form a separate clade arising from a CsCYP88L4 ancestor in the Cucurbitaceae. In addition, CsCYP88L2 apparently evolved more recently from a CsCYP88L4 ancestor that duplicated before the split of cucumber and *Citrullus lanatus*, and then underwent several duplications in cucumber; these duplicates form a different clade than that of McCYP88L7 (Figure 6). Theis suggests that C19-hydroxylation of cucurbitadienol evolved independently in cucumber and Momordica by recruitment from the CYP88L subfamily.

The CpCYP88L4 ORF sequence was amplified from squash root cDNA and tested for C19-hydroxylation activity by overexpression in a yeast strain producing cucurbitadienol, however, no activity was detected as compared to yeast overexpressing cucumber *CYP88L2* and momordica *CYP88L7* (data not shown).

Next, metabolic engineering of new cucurbitacins in squash HRs was attempted by overexpressing *CsCYP88L2* and *McCYP88L7*. Overexpression of either *CsCYP88L2* or *McCYP88L7* in squash HRs resulted in an obvious reduction of Cucs B,E&I compared to EV in some of the lines, however, cucurbitacins reduction was not statistically significant (Figure S6). No ions with m/z values corresponding to Cucs B,D,E&I with an additional hydroxyl group (Supplementary File 2) could be detected in the chromatograms of transgenic HRs overexpressing

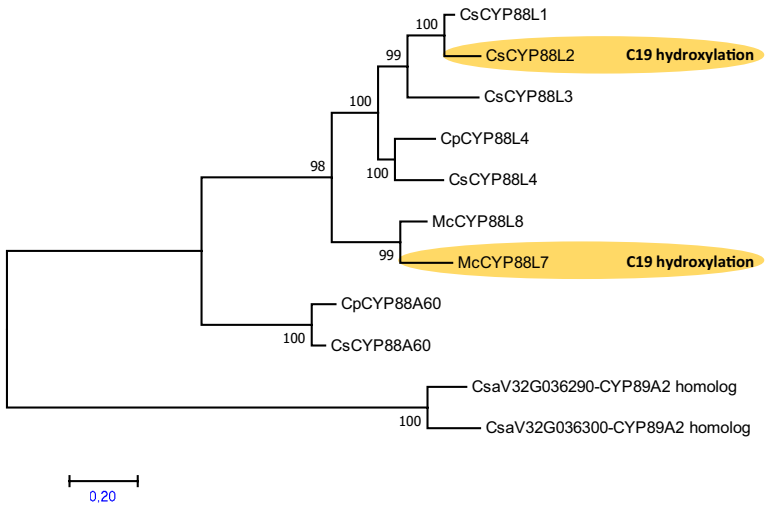


FIGURE 6
Phylogenetic analysis suggests that the P450 s the C19 hydroxylation in cucurbitadienol from *Cucumis sativus* and *Momordica charantia* evolved independently within the Cucurbitaceae. Maximum Likelihood tree constructed on deduced amino acid sequences aligned by ClustalW spanning 260 positions using MEGA7 program. The statistical significance of each node was tested by the bootstrap method using 1,000 iterations. Representative names are as follows: CsCYP88L3, *Cucumis sativus* uncharacterized P450; CsCYP88L1, *Cucumis sativus* uncharacterized P450; CsCYP88L2, P450 making cucurbitadienol C19 hydroxylation; CpCYP88L4, *Cucurbita pepo* CYP88L4 homolog; CsCYP88L4, *Cucumis sativus* uncharacterized P450; McCYP88L8, *Momordica charantia* P450 making cucurbitadienol C7 hydroxylation; McCYP88L7, *Momordica charantia* P450 making cucurbitadienol C19 and C7 hydroxylation; CpCYP88A60, *Cucurbita pepo* putative ent-kaurenoic acid hydroxylase; CsCYP88A60, *Cucumis sativus* putative ent-kaurenoic acid hydroxylase; CsCYP89A2-homologs, *Cucumis sativus* used as outgroup. Accession number of sequences are given in Table S5. Two Cucumis sativus sequences homologous to Arabidopsis thaliana CYP89A2 were used as an outgroup.

CsCYP88L2 or McCYP88L7. However, a thorough analysis of the LC-MS/MS chromatograms for putative cucurbitacin derived compounds revealed a peak eluding at 12.65 min in both the *CsCYP88L2* and *McCYP88L7* overexpressing lines (Figure 7A). The 12.65 min. peak is composed of several ions; the ion of m/z value of 469.2512 is found in both the EV and transgenic HR lines but the ion of m/z value 513.2452 $[M\text{-}H]^-$ is unique to transgenic lines (Figure S7A). The mass of the parental ion and fragmentation pattern for this candidate peak could correspond to a triterpenoid with a molecular formulas of: $C_{30}H_{42}O_7$ (containing four hydroxyl and three carbonyl groups as well as three double bonds), or $C_{32}H_{50}O_5$ with two different compositions of functional groups (two hydroxyl, one carbonyl and 1 acetyl as well as two double bonds or three hydroxyl and one acetyl as well as three double bonds) (Supplementary File 2). Possible cucurbitacin intermediate structures predicted for these molecular formulas are illustrated in Figures S7B-D. Due to the low accumulation levels of the putative new cucurbitacin, it was not possible to purify enough compound for structure elucidation using NMR.

Finally, to confirm the observation that Cucs B,D,E&I could not be hydroxylated by CYP88L2 and CYP88L7, the Cucs B,D,E&I were fed to yeast cells overexpressing these genes. As a positive control, when yeast cells overexpressing CpACT1 were fed Cuc I, acetylation of Cuc I to Cuc E was effectively observed (Figure S8). However, when Cucs B,D,E&I were individually fed to yeast cells overexpressing CsCYP88L2 or McCYP88L7, no additional peaks were observed (Figure S8). These results indicated an inability of CsCYP88L2 and McCYP88L7 to hydroxylate Cucs B,D,E&I, suggesting these enzymes work earlier in the cucurbitacin pathway, and may also explain the low level of the 12.65 min. peaked observed in *CsCYP88L2* and *McCYP88L7* overexpressing HR lines.

In summary, the phylogenetic analysis of the CYP88L subfamily in the Cucurbitacea suggests that the C19 hydroxylation trait by CsCYP88L2 and McCYP88L7 likely evolved independently in cucumber and Momordica. Both P450s proved capable of creating a new peak in the LC-MS analysis, suggesting cucurbitacins with unreported combinations of functional groups is possible in squash HRs through this strategy. Albeit, determination of this peak and strategies to optimize its production would need to be pursued in later research taking into account the fact that CsCYP88L2 and McCYP88L7 act early in cucurbitacin biosynthesis, as suggested by cucurbitacin-feeding experiments in yeast.

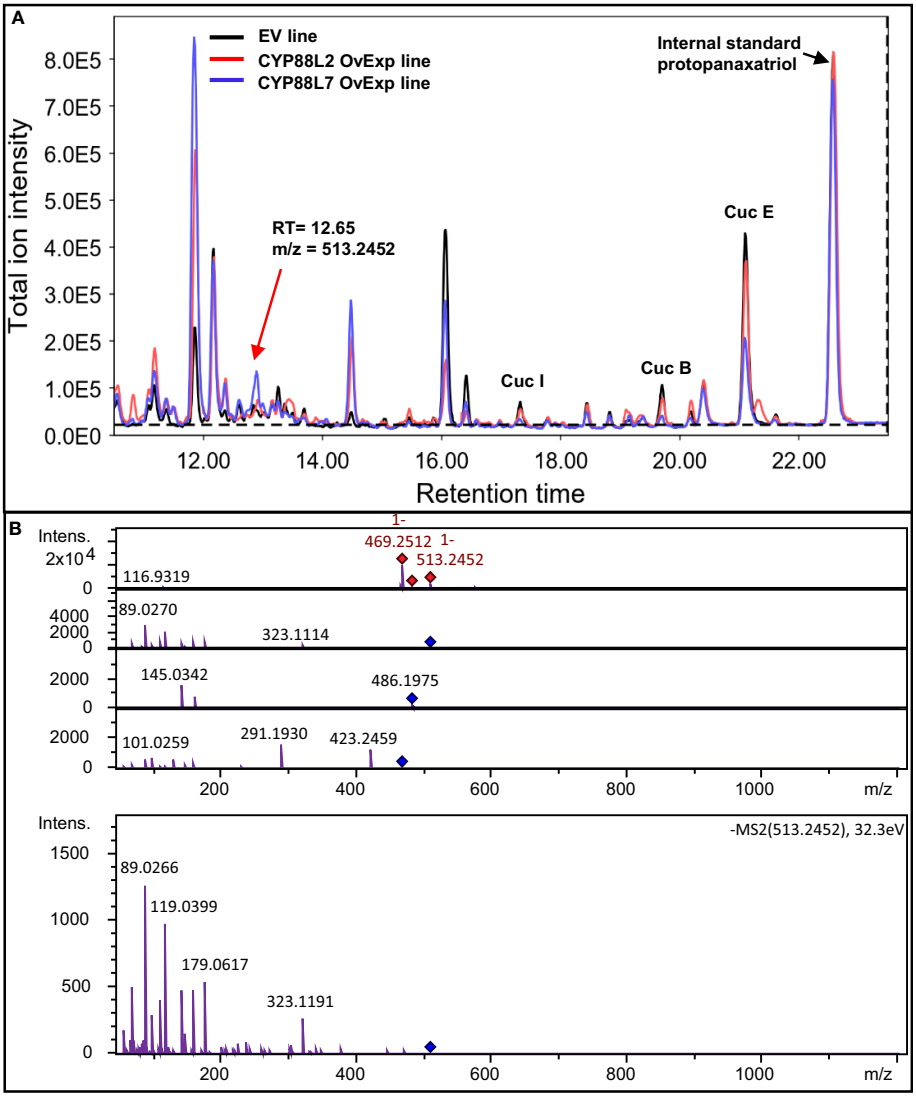


FIGURE 7
CYP88L2 and CYP88L7 appear to hydroxylate cucurbitacin intermediates in metabolically engineered *Cucurbita pepo* hairy roots. **(A)** Total ion chromatograms of HPLC-QToF analysis from *Cucurbita pepo* hairy roots empty vector, *CYP88L2* overexpression and *CYP88L7* overexpression lines. The new peak at 12,65 min is indicated with an arrow. **(B)** MS2 for the ion 513.2779 $[M+COOH]^-$ mass spectra at minute 12.65 *Cucurbita pepo* overexpressing *CYP88L2* and *CYP88L7*. Red diamonds represent ions detected in MS1 destined for fragmentation, while blue diamonds are these parental ion fragmented individually in the MS2 detector.

4 Discussion

4.1 bHLH clade II TFs can be used to induce cucurbitacin production across the Cucurbitaceae

The results in this paper show that cucurbitacin biosynthesis in Cucurbitaceae species can be increased by overexpression of *CpCUCbH1* in squash HRs and by environmental conditions such as growth in TIRs. Previously the class II bHLH TFs CsBl and CsBt in cucumber and ClBr in *Citrullus lanatus* were reported to regulate cucurbitacin biosynthesis in leaf, fruit and roots, respectively (Zhou et al., 2016). In this study it was observed that the bHLH class II in the Cucurbitaceae splits into two clades in which one contains functional homologs of CsBl. In particular, CpCUCbH1 can induce biosynthesis of Cucs B,D,E&I in squash HRs, Cuc C in cucumber cotyledons but only Cuc D in *E. elaterium* leaves. Furthermore, transient overexpression of CpCUCbH1 alone did not induce cucurbitacin biosynthesis in Luffa and squash cotyledons. Consistent with the latter observation, Brzozowski et al. (Brzozowski et al., 2020) reported that cucurbitacin

biosynthetic genes are not expressed in squash cotyledons and thus accumulation of cucurbitacins in cotyledons depends on transport of these metabolites from the roots; in addition, cucurbitacin accumulation trait in squash appears to be cultivar dependent.

It is known that bHLH TFs may make heterocomplexes with MYB TFs to be functional, and that the function of these heterocomplexes can be regulated by protein interactions such as by WD40 repeat proteins (Li, 2014; Pireyre and Burow, 2015). The absence of a MYB TF or presence of a WD40 protein could be a reason why cotyledons of squash and Luffa transiently overexpressing CpCUCbH1 did not produce cucurbitacins. When Shang et al. (Shang et al., 2014) restored Cuc C production in cotyledons of the non-bitter cucumber line XY-3 by transiently expressing CsBl, they were complementing a mutation in the CsBl TF, and since cucumber cotyledons typically produce Cuc C it is probable that the cofactors to induce Cuc biosynthesis were also present. In conclusion, further work is needed to understand the intricacies of cucurbitacin biosynthesis regulation throughout the Cucurbitaceae species and plant tissues.

When HRs were grown in different conditions, it was found that growth in TIR boosted Cuc production and induced the known biosynthetic genes (Figures 3, 4). The work described here was consistent with other papers where culture of HRs in TIR had been shown to increase yields of plant specialized metabolites (Thakore et al., 2017; Kochan et al., 2018); but perhaps, more interesting in this study was that this increase in expression of cucurbitacin biosynthetic genes and cucurbitacin production was not related to CpCUCbH1 expression (Figures 3, 4). In addition, class IV bHLH TFs inducing terpenoid production in *Medicago truncatula* and *Cataranthus roseus* are known to be induced by jasmonate signaling pathway, but this was apparently not the case for the expression of CpCUCbH1 since expression did not increase after MeJA treatment.

4.2 Potentially new cucurbitacin structures can be metabolically engineered in squash HRs and tabacco leaves

The cucurbitacin pathway continues to evolve in the Cucurbitaceae family, and thus not all Cucurbitaceae species have the same cucurbitacin profiles (Figure 1). Cucs B,D,E&I are the most widespread cucurbitacins, their presence can be observed starting in the Telfairieae tribe after the split from the Momodiceae (Rehm et al., 1957; Schaefer et al., 2009); nevertheless, cucumber in the Benincaseae tribe has lost the ability to produce Cucs B,D,E&I and instead biosynthesizes Cuc C. Cucurbitacins in the Momordiceae species in general lack the C16 hydroxylation (Okabe et al., 1982). In the case of cucumber, the absence of Cucs B,D,E&I is due to the lack of function of the CYP81Q59 catalyzing the C2 hydroxylation (Figure 2) (Shang et al., 2014). We identified an ACT, IaCUCA1, from *I. amara* that acetylates at the C3 postion. A phylogentic analysis showed that it evolved independently of the Cucurbitaceae ACTs that acetylate the C16 and C25 hydroxyl groups (Thakore et al., 2017), but evolved out of a clade in the Brassicaceae which has a predisposition for acetylating triterpenoids. To our knowledge, the acetylation of the hydroxyl group in C3 of cucurbitadienol has not been reported before. The fact that no 3-acetyl-16-hydroxy-cucurbitadienol was found in crude extracts of *I. amara*, most likely reflects the promiscuity of IaCUCA1; indeed, ACTs have been shown to be promiscuous acting in several substrates like the AtTHAA2 (Huang et al., 2019) from which IaCUCA1 diverged from. Nevertheless, the ability to create novel cucurbitacin modifications, reflects that *I. amara* represents an additional potential reservoir for cucurbitacin-modifying genes as the pathway has evolved independently in the two lineages.

Cucumber and Momordica diverged 46 ± 3 MY (Schaefer et al., 2009), but they both report the ability to hydroxylate at the C19 position. In contrast, there are no reports of cucurbitacins with C19 hydroxylation in squash despite of it diverging from cucumber only 30 ± 4 MY. The data presented in this paper shows that CsCYP88L2 and McCYP88L7 enzymes are responsible for the C19-hydroxylation in cucumber and Momordica, respectively, and suggests that CsCYP88L2 and McCYP88L7 must act early on the cucurbitadienol backbone as they are unable to hydroxylate highly oxygenated Cucs B,D, E&I (Figures 6; S8) possibly due to steric hindrance in the active site. Taking into consideration the suggested early action of these P450s and the lack of ions with m/z values for a deacetylated cucurbitacin in the MS2 spectra (Figure 7B), then a likely structure for the new observed peak could be one of those with molecular formula of $C_{30}H_{42}O_7$ in Figure S7B; especially the middle structure as this molecule would not be acetylated by the promiscuous CpACT1.

5 Conclusions

In this paper, we show that squash HR can be used to produce cucrbitacins and that through metabolic engineering cucurbitacin structures can be modified and even novel cucrbitacin structures can be generated. We show that overexpression of specific transcription factors is an effictive strategy to boost cucurbitacin levels in squash HR grown on agar plates. This complements our earlier findings that overexpression of squalene epoxidase in HR

and *Nicotiana benthamiana* increase cucurbitacin/triterpenoid levels by increasing the availability of the precursor 2,3-oxidosqualene (Almeida et al., 2018; Dong et al., 2018). In the present work we learned that the transcription factor CpCUCbH1 can induce cucurbitacins in several Cucurbitaceae species, indicating it can be used as a metabolic engineering tool in this family; however, the cucurbitacin-induction mechanism of CpCUCbH1 appears only partially conserved in *E. elaterium* in the Bryonieae tribe. In addition, convergent evolution of cucurbitacins in *I. amara* makes this plant a potential reservoir for genes that could generate novel cucucrbitacin structures with potentially new or improved bioactivities through metabolic engineering. Finally, using cucurbitacins as an example, this paper provides initial evidence that a HR platform can be used to modify and increase the production of specialized metabolites for which the biosynthetic pathway has not been fully elucidted.

Data availability statement

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

Author contributions

AA performed most of the experiments and analysis and drafted the article. TT was involved in the *Cucumis sativus* CYP88L2 experiments. MR was involved in the yeast experiments. LD and BK were involved in the *Iberis amara* experiments. NC-Q was involved in experimental design and construction of vectors. SK and SB supervised the work and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

SB, AA, and MR were supported by grants from the Independent Research Fund Denmark grant No. 7017-00275B and Novo Nordisk Foundation grant No. NNF17OC0027646.

Acknowledgments

The authors would like to thank Bruno Trevenzoli Favero and Henrik Vit Lütkin for providing the temporary immersion reactors used in this study. In the same manner, the authors thank Yadira Peña Garcia for providing the winrhizo data. The authors would also like to thank David Nelson for assisting in the annotation and naming of P450s in this manuscript.

Conflict of interest

The authors declare that this study received funding from the Novo Nordic Foundation. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1021907/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Cucurbita pepo has a higher transformation efficiency for hairy root line production than *Iberis umbellata*. Stereomicroscope images of *Cucurbita pepo* hairy roots overexpressing *mRFP*, on the left are the bright-field images **(A, C)** and on the right is the fluorescence through a 590 LongPass emission filter **(B, D)**. Stereomicroscope images of *Iberis umbellata* hairy roots overexpressing mRFP on the left are the bright-field images **(E, G, I)** and on the right is the fluorescence through a 590 LongPass emission filter **(F, H, J)**. **(K)** Example of hairy root lines generated from one transformation round. **(L)** Different tissues from *Iberis umbellata* transformed for hairy root production. The table below list the transformation frequencies of *C. pepo* versus *I. amara*.

SUPPLEMENTARY FIGURE 2
Temporary immersion reactors reduce within line variation of *Cucurbita pepo* hairy roots. Representative hairy root showing variation on growth in 200 ml Erlen Meyer flasks **(A-C)**. **(D)** Fresh weight yields of different hairy root lines transformed with different vectors, either EV or overexpression of CpACT, showing more consistent growth in TIR. Pictures of hairy roots grown in temporary immersion reactors **(E, F)**. Data points represent means with standard deviation (n = 6), ($p < 0.05$)

SUPPLEMENTARY FIGURE 3
Higher accumulation of cucurbitacins by overexpression of *CpCUCbH1* results in hairy roots with less branching and reduced total length. Representative hairy root lines overexpressing *CpCUCbH1* with three technical replications within each plate **(A-C)**. Representative empty vector hairy root lines with technical replicates within each plate **(D-F)**. Three root phenotypes: Primary root length **(G)**, Number of root tips **(H)** and Total root length **(I)** assessed by WinRhizo for empty vector and CpCUCbH1 overexpressing hairy roots

SUPPLEMENTARY FIGURE 4
Overexpression of *CpCUCbH1* alone does not induce *de novo* cucurbitacin biosynthesis in tissues that do not already accumulate cucurbitacins. Representative confocal image of *Cucurbita pepo* **(A)**, *Cucumis sativus* **(B)**, *Luffa cylindrical* **(C)** cotyledons and *Ecballium elaterium* leaves **(D)** overexpressing GFP to assess successful transformation. Confocal laser scanning microscopy images were

collected 5 days after agroinfiltration of cotyledons/leaves. **(E)** Total ion chromatograms of UHPLC-QToF analysis from infiltrated cotyledons of *Luffa cylindrica* showing absence of cucurbitacins. **(F)** Total ion chromatograms of UHPLC-QToF analysis from infiltrated cotyledons of *Cucurbita pepo* showing absence of cucurbitacins.

SUPPLEMENTARY FIGURE 5
Tandem duplications of the CYP88L subfamily in Cucurbitaceae species have resulted in expansion of CYP88L paralogs. **(A)** Alignment of CYP88 amino acid sequences showing 27 missing amino acids in the sequences misannotated on available databases (Cucurbit Genomics Data Base (Csa6G088690) and NCBI (NP_001295866)). The *CsCYP88L2* sequence obtained from cucumber leaf cDNA in this study (Genbank accession: MZ695220) encoded 27 aa more at the C-terminal region than sequences deposited in CuGenDB and NCBI. Multiple-pairwise aligments showed that the previously reported *CYP88L2* sequence had only 463 aa while all other CYP88s were 489-490 aa long, indicating it had been misannotated. **(B)** Genome browser image from Cucurbit Genomics Database displaying the chromosomal region of *Cucurbita pepo* where CYP88L paralog sequences are localized; the image shows that the region is misannotated as a single gene. The line at the left of the predicted transcript represents the start site (5'-end) while the arrow at the right represents the end (3'-end). **(C)** Graphical representation of chromosal region in *Cucurbita pepo* containing CYP88L sequences after reannotation with FgenesH software; this representation shows the region in Figure S6B actually contains 12 open reading frames out of which 3 are CYP88L paralogs. **(D)** Graphical representation of chromosomal region in *Cucurbita moschata* containing CYP88L sequences after reannotation with FgenesH software. **(E)** Graphical representation of chromosomaal region in *Cucurbita maxima* containing CYP88L sequences after reannotation with FgenesH software. **(F)** Graphical representation of chromosomal region in *Citrullus lanatus* containing CYP88L sequences after reannotation with FgenesH software. **(G)** Graphical representation of chromosomal region in *Cucumis sativus* containing CYP88L sequences after reannotation with FgenesH software.

SUPPLEMENTARY FIGURE 6
Overexpression of *CsCYP88L2* and *McCYP88L7* does not significantly lower levels of cucurbitacins in *Cucurbita pepo* hairy roots. Comparison of Cucurbitacin I **(A)**, B **(B)** and E **(C)** between *C. pepo* hairy root lines overexpressing different gene constructs shows there is no significant drop in the cucurbitacin levels compared to the EV lines. The same constructs were tested 2 and 3 months after transformation. Data points represent means with standard deviation (n = 6), (p < 0.05).

SUPPLEMENTARY FIGURE 7
The m/z ion 513.2452 is unique to transgenic hairy roots of squash overexpressing either *CYP88L2* or *CYP88L7*. **(A)** Extracted ion chromatogram (m/z= 513-513.5) for UHPLC-QToF analysis (negative mode) for EV and *CYP88L2* or *CYP88L7* overexpressing hairy roots. **(B)** Possible cucurbitacin structures for an ion of m/z= 513-513.5 with molecular formula $C_{30}H_{42}O_7$ containing four hydroxyl and three carbonyl groups as well as three double bonds. **(C)** Possible cucurbitacin structures for an ion of m/z= 513-513.5 with molecular formula $C_{32}H_{50}O_5$ containing two hydroxyl, one carbonyl and 1 acetyl groups as well as two double bonds. **(D)** Possible cucurbitacin structures for an ion of m/z= 513-513.5 with molecular formula $C_{32}H_{50}O_5$ containing three hydroxyl and one acetyl groups as well as three double bonds.

SUPPLEMENTARY FIGURE 8
McCYP88L7 acts early on in the cucurbitacin pathway as it does not act on highly oxygenated cucurbitacins like Cucs E&I. **(A)** Extracted ion chromatograms (m/z= 532-533; 574-575) for UHPLC-QToF runs (positive mode) of yeast cultures overexpressing Mc*CYP88L7* supplemented with Cucs E&I. **(B)** Extracted ion chromatograms (m/z= 559-560; 601-602) for UHPLC-QToF runs (negative mode) of yeast cultures overexpressing *CpACT* fed Cuc I to generate Cuc E to prove that supplementing with cucurbitacins to yeast is possible.

SUPPLEMENTARY TABLE 1
List of primers used in this study.

SUPPLEMENTARY TABLE 2
List of plasmids used for yeast experiments.

SUPPLEMENTARY TABLE 3
List of seed providers for Cucurbitaceae plant species used in this study.

SUPPLEMENTARY TABLE 4
List of yeast strains used and constructed in this study.

SUPPLEMENTARY TABLE 5
Details of sequences used for sequence alignments in Figure 2, Figure 4 and Figure 5.

SUPPLEMENTARY FILE 1
Structure elucidation of 3-O-acetyl-16β-hydroxycucurbitadienol using NMR.

SUPPLEMENTARY FILE 2
Calculated masses of potential intermediates in the biosynthesis of cucurbitacins.

References

Abdelwahab, S. I., Hassan, L. E. A., Sirat, H. M., Yagi, S. M. A., Koko, W. S., Mohan, S., et al. (2011). Anti-inflammatory activities of cucurbitacin e isolated from citrullus lanatus var. *citroides: Role of reactive nitrogen species and cyclooxygenase enzyme inhibition*. *Fitoterapia* 82 (8), 1190–1197.

Almeida, A., Dong, L., Khakimov, B., Bassard, J-E., Moses, T., Lota, F., et al. (2018). A single oxidosqualene cyclase produces the <em<Seco</em<-triterpenoid α-onocerin. *Plant Physiol.* 176 (2), 1469. doi: 10.1104/pp.17.01369

Almeida, A., Dong, L., Appendino, G., and Bak, S. (2020). Plant triterpenoids with bond-missing skeletons: biogenesis, distribution and bioactivity. *Natural Product. Rep.* 37 (9), 1207–1228. doi: 10.1039/C9NP00030E

Arsenault, J. L., Poulcur, S., Messier, C., and Guay, R. (1995). WinRHlZO™, a root-measuring system with a unique overlap correction method. *HortSci. HortSci.* 30 (4), 906D–9906. doi: 10.21273/HORTSCI.30.4.906D

Brzozowski, L. J., Gore, M. A., Agrawal, A. A., and Mazourek, M. (2020). Divergence of defensive cucurbitacins in independent cucurbita pepo domestication events leads to differences in specialist herbivore preference. *Plant. Cell Environ.* 43 (11), 2812–2825. doi: 10.1111/pce.13844

Chen, J. C., Chiu, M. H., Nie, R. L., Cordell, G. A., and Qiu, S. X. (2005). Cucurbitacins and cucurbitane glycosides: structures and biological activities. *Natural Product. Rep.* 22 (3), 386–399. doi: 10.1039/b418841c

David, A., and Vallance, D. K. (1955). Bitter principles of cucurbitaceæ. *J. Pharm. Pharmacol.* 7 (1), 295–296. doi: 10.1111/j.2042-7158.1955.tb12040.x

Dinan, L., Whiting, P., Girault, J-P., Lafont, R., Dhadialla, T., Cress, E. D., et al. (1997). Cucurbitacins are insect steroid hormone antagonists acting at the ecdysteroid receptor. *Biochem. J.* 327 (3), 643–650. doi: 10.1042/bj3270643

Dong, L., Pollierb, J., Bassard, J.-E., Ntallas, G., Almeida, A., Lazaridi, E., et al. (2018). Co-Expression of squalene epoxidases with triterpene cyclases boosts production of triterpenoids in plants and yeast. *Metab. Eng.* 49, 1–12. doi: 10.1016/j.ymben.2018.07.002

Dong, L., Almeida, A., Pollier, J., Khakimov, B., Bassard, J-E., Miettinen, K., et al. (2021). An independent evolutionary origin for insect deterrent cucurbitacins in iberis amara. *Mol. Biol. Evol.* 38 (11), 4659–4673. doi: 10.1093/molbev/msab213

Ellerström, M., Josefsson, L.-G., Rask, L., and Ronne, H.. (1992). Cloning of a cDNA for rape chloroplast 3-isopropylmalate dehydrogenase by genetic

complementation in yeast. *Plant Mol. Biol.* 18 (3), 557–566. doi: 10.1007/BF00040671

Ferguson, J. E., and Metcalf, R. L. (1985). Cucurbitacins. *J. Chem. Ecol.* 11 (3), 311–318. doi: 10.1007/BF01411417

Geu-Flores, F., Nour-Eldin, H. H., Nielsen, M. T., and Halkier, B. A. (2007). USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.* 35 (7), e55–e55. doi: 10.1093/nar/gkm106

Gietz, R. D., and Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2 (1), 31–34. doi: 10.1038/nprot.2007.13

Huang, A. C., Jiang, T., Liu, Y.-X., Bai, Y.-C., Reed, J., Qu, B., et al. (2019). A specialized metabolic network selectively modulates Arabidopsis root microbiota. *Science* 364 (6440), eaau6389. doi: 10.1126/science.aau638

Ignea, C., Cvetkovic, I., Loupassaki, S., Kefalas, P., Johnson, C. B., Kampranis, S. C., et al. (2011). Improving yeast strains using recyclable integration cassettes, for the production of plant terpenoids. *Microbial. Cell Factories.* 10 (1), 4. doi: 10.1186/1475-2859-10-4

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8 (3), 275–282. doi: 10.1093/bioinformatics/8.3.275

Kaushik, U., Aeri, V., and Mir, S. R. (2015). Cucurbitacins - an insight into medicinal leads from nature. *Pharmacognosy. Rev.* 9 (17), 12–18. doi: 10.4103/0973-7847.156314

Khakimov, B., Motawia, M. S., Bak, S., and Engelsen, S. B. (2013). The use of trimethylsilyl cyanide derivatization for robust and broad-spectrum high-throughput gas chromatography–mass spectrometry based metabolomics. *Anal. Bioanal. Chem.* 405 (28), 9193–9205. doi: 10.1007/s00216-013-7341-z

Kochan, E., Balcerczak, E., Lipert, A., Szymańska, G., and Szymczyka, P. (2018). Methyl jasmonate as a control factor of the synthase squalene gene promoter and ginsenoside production in American ginseng hairy root cultured in shake flasks and a nutrient sprinkle bioreactor. *Ind. Crops Products.* 115, 182–193. doi: 10.1016/j.indcrop.2018.02.036

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. doi: 10.1093/molbev/msw054

Li, S. (2014). Transcriptional control of flavonoid biosynthesis. *Plant Signaling Behav.* 9 (1), e27522. doi: 10.4161/psb.27522

Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E., Hansen, B. G., Mortensen, U. H., et al. (2012). Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* 14 (2), 104–111. doi: 10.1016/j.ymben.2012.01.006

Okabe, H., Miyahara, Y., and Yamauchi, T. (1982). Structures of momordicosides F1, F2, G, I, K and l, novel cucurbitacins in the fruits of momordica charantia l. *Tetrahedron. Lett.* 23 (1), 77–80. doi: 10.1016/S0040-4039(00)97537-3

Pireyre, M., and Burow, M. (2015). Regulation of MYB and bHLH transcription factors: A glance at the protein level. *Mol. Plant* 8 (3), 378–388. doi: 10.1016/j.molp.2014.11.022

Rehm, S., Enslin, P. R., Meeuse, A. D. J., and Wessels, J. H.. (1957). Bitter principles of the cucurbitaceae. VII.–the distribution of bitter principles in this plant family. *J. Sci. Food Agric.* 8 (12), 679–686. doi: 10.1002/jsfa.2740081203

Rios, L., Andujar, I., Escandell, M., Giner, M., and Recio, C. (2012). Cucurbitacins as inducers of cell death and a rich source of potential anticancer compounds. *Curr. Pharm. Design.* 18 (12), 1663–1676. doi: 10.2174/138161212799958549

Sainsbury, F., Thuenemann, E. C., and Lomonossoff, G. P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* 7 (7), 682–693. doi: 10.1111/j.1467-7652.2009.00434.x

Schaefer, H., Heibl, C., and Renner, S. S. (2009). Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proc. R. Soc. B.: Biol. Sci.* 276 (1658), 843–851. doi: 10.1098/rspb.2008.1447

Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., and McLean, J. A.. (2016). Untargeted metabolomics strategies–challenges and emerging directions. *J. Am. Soc. Mass. Spectromet.* 27 (12), 1897–1905. doi: 10.1007/s13361-016-1469-y

Shang, Y., Ma, Y., Zhou, Y., Zhang, H., Duan, L., Chen, H., et al. (2014). Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science* 346 (6213), 1084. doi: 10.1126/science.1259215

Solovyev, V., Kosarev, P., Seledsov, I., and Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 7 (1), S10. doi: 10.1186/gb-2006-7-s1-s10

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics* 3 (3), 211–221. doi: 10.1007/s11306-007-0082-2

Sun, J., Blaskovich, M. A., Jove, R., Livingston, S. K., Coppola, D., Sebti, S. M., et al. (2005). Cucurbitacin q: a selective STAT3 activation inhibitor with potent antitumor activity. *Oncogene* 24 (20), 3236–3245. doi: 10.1038/sj.onc.1208470

Takase, S., Kera, K., Nagashima, Y., Mannen, K., Hosouchi, T., Shinpo, S., et al. (2019). Allylic hydroxylation of triterpenoids by a plant cytochrome P450 triggers key chemical transformations that produce a variety of bitter compounds. *J. Biol. Chem.* 294 (49), 18662–18673. doi: 10.1074/jbc.RA119.009944

Tan, M.-J., Ye, M., Turner, N., Hohnen-Behrens, C., Ke, C.-Q., Tang, C.-P., et al. (2008). Antidiabetic activities of triterpenoids isolated from bitter melon associated with activation of the AMPK pathway. *Chem. Biol.* 15 (3), 263–273. doi: 10.1016/j.chembiol.2008.01.013

Thakore, D., Srivastava, A. K., and Sinha, A. K. (2017). Mass production of ajmalicine by bioreactor cultivation of hairy roots of catharanthus roseus. *Biochem. Eng. J.* 119, 84–91. doi: 10.1016/j.bej.2016.12.010

Thomas, B. J., and Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. *Cell* 56 (4), 619–630. doi: 10.1016/0092-8674(89)90584-9

Wattj, M., and Breyer-brandwijk, M. G. (1962). *The medicinal and poisonous plants of southern and Eastern Africa being an account of their medicinal and other uses, chemical composition, pharmacological effects and toxicology in man and animal* Vol. 1457 (Teviot Place, Edinburgh: E. & S. Livingstone Ltd. xii), 16–17.

Zhou, Y., Ma, Y., Zeng, J., Duan, L., Xue, X., Wang, H., et al. (2016). Convergence and divergence of bitterness biosynthesis and regulation in cucurbitaceae. *Nat. Plants* 2 (12), 16183. doi: 10.1038/nplants.2016.183

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/fpls.2022.1076871

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Miroslava Konstantinova Zhiponova,
Sofia University, Bulgaria
Ajay Kumar,
National Botanical Research Institute (CSIR), India

*CORRESPONDENCE
Shachi Singh
singhshachi@gmail.com

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

RECEIVED 22 October 2022
ACCEPTED 20 December 2022
PUBLISHED 09 January 2023

CITATION
Singh S, Kumar M, Dwivedi S, Yadav A and Sharma S (2023) Distribution profile of iridoid glycosides and phenolic compounds in two *Barleria* species and their correlation with antioxidant and antibacterial activity. *Front. Plant Sci.* 13:1076871. doi: 10.3389/fpls.2022.1076871

Distribution profile of iridoid glycosides and phenolic compounds in two *Barleria* species and their correlation with antioxidant and antibacterial activity

Shachi Singh[1]*, Mukesh Kumar[2], Seema Dwivedi[1], Anjali Yadav[1] and Sarika Sharma[1]

[1]Department of Botany, MMV, Banaras Hindu University, Varanasi, India, [2]Department of Statistics, MMV, Banaras Hindu University, Varanasi, India

Introduction: *Barleria prionitis* is known for its medicinal properties from ancient times. Bioactive iridoid glycosides and phenolic compounds have been isolated from leaves of this plant. However, other parts of a medicinal plants are also important, especially roots. Therefore, it is important to screen all organs for complete chemical characterization.

Method: All parts of *B. prionitis*, including leaf, root, stem and inflorescence in search of bioactive compounds, with a rapid and effective metabolomic method. X500R QTOF system with information dependent acquisition (IDA) method was used to collect high resolution accurate mass data (HRMS) on both the parent (MS signal) and their fragment ions (MS/MS signal). ESI spectra was obtained in positive ion mode from all parts of the plant. A comparative analysis of antioxidant and antibacterial activity was done and their correlation study with the identified compounds was demonstrated. Principal component analysis was performed.

Result: Iridoid glycosides and phenolic compounds were identified from all parts of the showing variability in presence and abundance. Many of the compounds are reported first time in *B. prionitis*. Antioxidant and antibacterial activity was revealed in all organs, root being the most effective one. Some of the iridoid glycoside and phenolic compounds found to be positively correlated with the tested biological activity. Principal component analysis of the chemical profiles showed variability in distribution of the compounds.

Conclusion: All parts of *B. prionitis* are rich source of bioactive iridoid glycosides and phenolic compounds.

KEYWORDS

iridoid glycoside, metabolomic, antibacterial, antioxidant, phenolic compounds

1 Introduction

Barleria, a member of Acanthaceae family, is a spiny shrub and known for its medicinal properties from ancient times (Chen et al., 1998; Amoo et al., 2009). It is found in India and also distributed in different parts of Asia and Africa (Chen et al., 1998). Aerial parts of this plant has long been used in the treatment of diseases such as, toothache, whooping cough, respiratory and gastrointestinal disorders, fever, swelling, artheritis, skin disorders, jaundice etc (Amoo et al., 2009; Gangaram et al., 2022). Extracts collected from aerial parts of the plant have shown to possess antioxidant, antibacterial, antiviral, anti-inflammatory, hepatoprotective, antidiabetic, antifertility, immune modulating activity and acetylcholinesterase inhibitory activities (Chen et al., 1998; Gupta et al., 2000; Singh et al., 2003; Singh et al., 2005; Verma et al., 2005; Kosmulalage et al., 2007; Amoo et al., 2009; Dheer and Bhatnagar, 2010; Chavan et al., 2014; Sharma et al., 2014; Gangaram et al., 2022).

Iridoid glycosides and phenolics are important class of chemical compounds isolated from leaves of this plant. Iridoids are oxygenated monoterpenes that commonly occur as glycoside attached with glucose. They have been reported in several plant families and have shown a broad spectrum of biological activities (Widyowatia et al., 2010; Zhang et al., 2018a). Isolated compounds from *Barleria* species (Taneja and Tiwari, 1975; Kanchanapooma et al., 2002; Leea et al., 2016; Gangaram et al., 2022) have demonstrated to possess glutathione S-transferase inhibitory activity, acetylcholinesterase inhibitory activity, free radical scavenging, antimicrobial, anti-inflammatory, immunomodulatory and gastroprotective activities (Ata et al., 2009; Ghule and Yeole, 2012; Jaiswal et al., 2014; Zhang et al., 2018b; Sun et al., 2022). Phenolic compounds are widely distributed throughout the plant kingdom. They are involved in a variety of biological activities such as antimicrobial, anti-inflammatory, antioxidant, antidiabetic, hepatoprotective, and anticancer properties (Tanase et al., 2019; Pannakal et al., 2022). Phenolics isolated from aerial parts of *Barleria* species have shown to possess biological activities (Vertuani et al., 2011; Gangaram et al., 2022).

Although active ingredients had been identified in leaves of *Barleria species*, limited work is done to explore other parts of the plant, such as root, bark or flowers, which could also be a rich source of bioactive compounds. Therefore, the objective of this study was to identify the active compounds present in all parts of the plant and study their distribution pattern so that the full pharmacological potential of a medicinal plant could be exploited.

Hyphenated liquid chromatography and tandem mass spectrometry (LC–MS/MS) is considered to be the most suitable technique for chemical characterization of natural extracts. Although the necessity of the technique in natural product science is unquestionable, however, it is time consuming and cannot meet the demand arising due to increase in number of samples. Therefore, it is required to find a more efficient technique which can perform high-throughput analysis by overcoming the time constrains of the conventional method. Application of direct mass measurement through techniques, such as electrospray ionisation (ESI-MS) have proven useful in characterizing crude extracts, such as medicinal plants (Mauri and Pietta, 2000; Kaur and Kaur, 2016; Liu et al., 2021) essential oils (Møller et al., 2007) etc. The speed is fast, however in absence of MS/MS signals, the information may not be sufficient to identify the compounds with accuracy. If a mass spectrometry instrument can select ion currents and generate their MS1 and MS2 spectra with a single injection, the necessity of separation based techniques can be overcome. In the present experiment we have used direct mass measurement coupled with information dependent acquisition (IDA) method for generation of MS and MS/MS signals. The method was found useful in analyzing multiple batches of sample within a short interval of time.

2 Methods

2.1 Plant material and preparation of extract

B. prionitis and *B. cristata* were collected from the Ayurvedic Garden of Banaras Hindu University, Varanasi, India, in the month of December, 2021. The plant material was washed thoroughly under running tap water and dried in air. Different parts of the plant, such as root, stem, leaves and inflorescence were separated and kept in oven with temperature not exceeding above 50°C. Dried plant material was grinded into fine powder. About 50 g of plant powder was extracted by absolute ethanol with sonication at 45°C for 15 min with frequency 40 kHz, power 100W, the procedure was repeated thrice. The decoction was filtered using the Whatman paper, the plant material was again extracted thrice with 70% ethanol; both the filtrates were pooled together and dried at 45°C. The obtained residue was extracted with hexane, to remove non-polar compounds, the remaining residue was used for further analysis.

2.2 2, 2-Diphenyl-1-picrythydrazyl (DPPH) radical scavenging assay

Total free radical scavenging capacity of the extract was estimated using the stable DPPH radical, as per the method described by Singh, 2016. Test sample was prepared by dissolving 10 mg extract in 1 ml methanol and 0.004% DPPH solution was prepared by dissolving 4 mg DPPH in 100 ml methanol. A 0.5 ml of methanolic extract was diluted by adding 2 ml methanol, followed by addition of 1 ml of DPPH solution. The mixture was shaken and incubated in dark for 30 minutes at room temperature. Absorbance was measured at 517nm

spectrophotometrically. Percentage of antioxidant activity was calculated using the following formula:

$$\text{Radical scavenging activity } \% = \text{Absorbance of Control} - \text{Absorbance of Sample}/\text{Absorbance of control} * 100$$

Where control was 3 ml methanol with 1 ml DPPH and sample was methanolic plant extract.

Antioxidant activity of the extracts was compared with standard antioxidant (Ascorbic acid). IC_{50} value was determined from the plotted graph of scavenging activity against the different concentrations of extracts, which is defined as the total antioxidant necessary to decrease the initial DPPH radical concentration by 50%.

2.3 Ferric- reducing antioxidant power (FRAP) assay

Frap assay was performed according to the method of Guo et al. (2003). Frap reagent was prepared freshly by addition of 2.5 ml of 10mmol/l TPTZ (2,4,6-Tripyridy- s-triazine, sigma) in 40 mmol/l HCl and 2.5 ml of 20mmol/l FeCl3 in 25 ml of 0.3mol/l acetate buffer, pH 3.6 and kept at 37^0C. Now a 0.5 ml of methanolic extract was taken in a test tube and diluted by 0.5 ml methanol, followed by addition of 2 ml frap reagent, the mixture was kept for 15 minutes in 37°C. Absorbance of this reaction mixture was measured at 593nm using UV/VIS spectrophotometer. Antioxidant activity was calculated by comparing with standard antioxidant (ascorbic acid).

2.4 Antibacterial activity

Antibacterial activity was performed by well diffusion method, according to the method decribed in literature (Singh et al., 2020), on the following test bacteria, obtained from MTCC (IMTECH, Chandigarh); *Escherichia coli, Staphylococcus aureus sub* sp. *aureus, Bacillus subtilis, Pseudomonas aeruginosa* and *Klebsiella pneumonia*. Plant extract of 10mg/ml concentration was prepared in distilled water. Nutrient agar plates were prepared and 0.5 cm wells were made on the solid media. Plates were inoculated by bacterial culture and wells were filled with 20 µl of the extract. The plates were incubated for 24h at 37° C. Antibacterial activity of the compound was determined by measuring the diameter of zone of inhibition of microbial growth. Antibiotic, Streptomycin at concentration of 1 mg/ml was used as positive control.

2.5 Mass spectrometry

Plant extract (10 mg/ml) of root, leaf, stem and inflorescence of *B.prionitis* and *B.cristata*, prepared in methanol, were directly analyzed by mass spectrometer. The SCIEX X500R QTOF system with the Turbo V^{TM} source was operated in positive electrospray ionization (ESI) mode. The TOF MS scan was conducted over a range of 100-1000 m/z. Following MS parameters were selected; ion source gas one 60 psi, ion source gas two 60 psi, curtain gas 40 psi, source temperature 400-500°C, ion spray voltage 5500 V, accumulation time 0.25 s, declustering potential 60-100 V and collision energy 7 V. An automated information dependent acquisition (IDA) approach was chosen for data collection. The TOF-MS data were acquired with high resolution mass measurement (HRMS) and isotopic resolution. Ion intensity signal less than 1000 cps (count per second) was not considered for evaluation. The resulting ion intensity matrices were normalized to be expressed as a percentage relative to the most intense ion in the spectrum (taken as 100%). Using the X500R's high-resolution mass spectrum TOF-MSIDA, with a single sample injection, TOFMS and TOF-MS/MS mass spectra were obtained. Using primary high-resolution mass numbers and the MS/MS database for comparison (e.g CAS registry/SciFinder, ChemSpider, Dictionary of Natural Products), published database searches and structural verification were performed for identification of the compounds.

2.6 Statistical analysis

All the assays were repeated five times and results were shown as mean ± standard deviation. Linear regression analysis was used to calculate the IC_{50} values. Pearson's correlation coefficient was calculated and statistical significance was determined with one way ANOVA test. A statistical significance of $p < 0.05$ was considered to be significant. The principal component analysis (PCA) was used to show the variation of multivariate data set in terms of components. This study includes ion intensity signals of components present in root, stem, leaf, and flower. The visual presentation PCA has been done through R software. In R software biplot function and library(ggfortify) has been used to plot PCA. Heatmaps were created using the statistical software Python. Ion intensity data peak were normalized and an average value of five replicate was used for construction of heatmap.

3 Result

3.1 Compound identification

ESI-MS in positive ion mode was applied for identification of active compounds from the aerial and underground parts of *Barleria* species. A total of 102 ions were detected from the mass spectra of leaf, root, stem and inflorescence of the plants. We were able to identify 58 ions, belonging to phenolic and iridoid

group. A clear difference was observed in the chemical profile of both the plants as well as within different parts of the plant, showing variation in distribution and abundance of the compounds (Table 1; Figure 1). Abundant sodium and potassium adducts prevailed in the positive spectra. Adducts were determined by their mass difference from their protonated ions. Compounds were identified based on their MS1 and MS2 signals as well as isotopic distribution.

3.2 Identification of iridoid glycosides

Major iridoid glycosides detected in both the plants were shanzhiside methyl ester, barlerin and acetylbarlerin. Along with them Shanzhiside and 7-methoxydiderroside were observed in *B. prionitis* and aucubin in *B. cristata* (Table 1). Their identity was confirmed by comparing their exact mass (m/z) and

TABLE 1 Compounds identified in different parts of *B. prionitis* and *B. cristata* and their relative abundance in percentage.

	Mass peak (m/z)	Compounds identified		MS/MS Diagnostic ions	Relative Abundance (%)							
					B. prionitis				*B. cristata*			
		Name	Ions selected		BYL	BYR	BYS	BYF	BWL	BWR	BWS	BWF
1.	167.0682	Melilotic acid	$[M+H]^+$	149	2.2	–	2.2	–	–	–	1.2	0.3
2.	195.0674	Ferulic acid	$[M+H]^+$	177	–	–	2.5	3.2	0.9	0.5	–	–
3.	191.0300	Vanillic acid	$[M+Na]^+$	173	5.1	5.0	6.0	15.3	–	–	–	–
4.	203.0843	Coumaric acid	$[M+K]^+$	186	–	–	–	–	10.2	11.6	15.3	18.7
5.	208.9808	Gallic acid	$[M+K]^+$	191	2.4	2.9	3.2	5.8	–	–	–	–
6.	219.0240	Caffeic acid	$[M+K]^+$	163	0.5	–	0.4	0.3	3.5	4.4	12.5	3.4
7.	221.0856	Syringic acid	$[M+Na]^+$	203	3.5	1.3	–	3.5	3.3	–	–	5.1
8.	223.0966	Methylgallic acid	$[M+K]^+$	–	–	0.5	0.3	–	0.2	–	0.5	0.3
9.	235.1651	Dihydroferulic acid	$[M+K]^+$	217	7.1	2.8	3.5	8.1	6.9	2.8	3.2	15.5
10.	293.0643	Apigenin	$[M+Na]^+$	275, 199	5.2	–	–	3.2	–	–	–	–
11.	315.1582	Hydroxy benzoic acid	[2M+K]+		0.3	0.2	–	0.3	1.5	–	–	–
12.	301.1396	Gossypetin	$[M\text{-}OH+H]^+$	–	0.02	1.0	0.5	1.2	–	–	1.5	–
13.	325.0939	Quercetin	$[M+Na]^+$	–	–	0.4	0.2	–	0.2	1.3	–	–
14.	339.0626	Methylquercetin	$[M+Na]^+$	–	–	0.5	0.4	–	0.5	1.9	3.1	–
15.	371.1324	Gossypetin 3methylether	$[M+K]^+$	340, 353	2.1	1.1	2.5	5.4	1.1	2.5	2.2	0.7
16.	381.0821	Sucrose	$[M+K]^+$	181	3.2	18.5	35.7	4.3	24.5	100	98.8	6.6
17.	385.0883	Aucubin	$[M+K]^+$		–	–	–	–	27.4	69.9	52.7	21.3
18.	431.1574	Shanzhiside	$[M+K]^+$		1.3	–	–	2.7	–	–	–	–
19.	445.1135	Shanzhiside methylester	$[M+K]^+$	283, 265, 247, 233, 229, 215, 187,159,	32.3	30.5	53.5	29.5	7.5	5.4	4.9	4.2
20.	471.1503	Barlerin	$[M+Na]^+$	411, 249, 231, 217, 203, 191, 177, 159, 131	83.3	100	5.5	90.3	1.2	2.7	2.1	7.5
21.	517.1443	7-Methoxydiderroside	$[M+K]^+$	457, 355	11.5	1.1	1.5	40	–	–	–	–
22.	529.1350	Acetylbarlerin	$[M+K]^+$	487, 427, 265	67.9	3.5	4.3	60.3	1.2	1.5	0.04	4.5
23.	543.1370	Caffeic acid derivative		163	0.2	5.4	4.1	0.3	3.4	20.5	18.7	1.2
24.	555.1131	Kaempferol	$[M+K]^+$	537	–	0.8	–	–	–	–	–	–
25.	591.1534	6-O-trans-p-Coumaroyl shanzhiside methyl ester	$[M+Na]^+$	–	–	1.5	–	–	–	–	–	–

(Continued)

TABLE 1 Continued

	Mass peak (m/z)	Compounds identified		MS/MS Diagnostic ions	Relative Abundance (%)							
					B. prionitis				*B. cristata*			
		Name	Ions selected		BYL	BYR	BYS	BYF	BWL	BWR	BWS	BWF
26.	593.1634	Saletpangponoside	$[M+K]^+$	–	–	–	–	7.2	0.5	–	–	3.4
27.	611.3609	Quercetin rutinoside		285, 325	–	–	1.5	–	–	–	–	–
28.	630.2427	Caffeic acid glycoside		612, 449	8.4	0.1	0.1	62.5	–	–	–	2.1
29.	663.4591	Acetoside	$[M+K]^+$	163, 517,495, 365, 383, 177	0.03	2.2	1.0	0.02	3.1	15.6	7.6	3.5
30.	705.1915	Acetylacetoside	$[M+K]^+$	663	–	10.5	10.8	–	0.9	30.6	18.5	2.2
31.	775.2332	Caffeic acid derivative		163	–	–	–	–	–	5.2	1.8	–
32.	793.2606	Poliumoside	$[M+Na]^+$	163, 630	–	4.5	1.2	0.04	–	–	–	–
33.	893.2803	Acetylpoliumoside	$[M+Na]^+$	793	4.5	4.0	8.3	9.2	–	–	–	–
34.	955.2951	Barlerinoside	$[M+Na]^+$	–	–	–	–	–	–	–	–	–

B.prionitis: BYL-leaf extract, BYR-root extract, BYS=-stem extract, BYF- inflorescence extract.
B.cristata: BWL-leaf extract, BWR-root extract, BWS=-stem extract, BWF- inflorescence extract.

fragmentation pattern (wherever applicable) from the databases and study of their isotopic distribution. Shanzhiside methyl ester, barlerin and acetylbarlerin were the most common compound present in the extracts, however their abundance was high in *B. prionitis* (Figure 1).

Shanzhiside methyl ester was observed as sodium adduct $[M+Na]^+$ with m/z 429, potassium adduct with m/z 445 and potassium dimer $[2M+K]^+$ with m/z 851. Isotopes of this compound observed at m/z 446 and 447 showed a normal distribution pattern of 10:3:1 (Figure 2). Product ion spectra for both sodium and potassium adduct were generated. Since both ions were present in abundant amount, hence were selected by the system for MS/MS analysis (Figure 3). Fragmentation pattern of potassium adduct of Shanzhiside methyl ester at m/z 445 is explained in Figure 4, where we can clearly observe the removal of sugar moiety (162 Da) from the iridoid part generating ion at m/z 283. The most abundant ion at m/z 247 is formed by loss of two water molecules from the iridoid part. Loss of one more water molecule gave rise to signal at m/z 229, loss of $C_2H_4O_2$ gave rise to ion at m/z 187 and loss of CO generates ion at m/z 159. A hydroxyl group is linked at C-6 position in shanzhiside methyl ester, so it easily losses a methanol molecule to form a lactone with the methylacetate ($COOCH_3$) group at the C-4 position, generating product ions at m/z 233 and 215.

Barlerin was observed as sodium adduct $[M+Na]^+$ at m/z 471, potassium adduct at m/z 487, sodium dimer $[2M+K]^+$ at m/z 919 and potassium dimer $[2M+K]^+$ at m/z 935, with a normal isotopic distribution pattern. MS/MS analysis of barlerin at m/z 471 $[M+Na]^+$ gave fragment ion at m/z 411 formed by cleavage of methylacetate ($COOCH_3$) group from the parent ion and at m/z 249 formed by removal of glucose group. Loss of water molecule gives a peak at m/z 231, whereas loss of methanol (32 Da) give rise to a peak at m/z 217. In barlerin also, a hydroxyl group is linked at C-6 position, so it losses a methanol molecule to form a lactone with the $COOCH_3$ group at the C-4 position. Loss of CO group (28 Da) from 231 ion give rise to peak at m/z 203. Replacement of sodium ion with hydrogen from 231 followed by loss of water gives peak at 191. Loss of water from 217 and replacement of sodium ion with hydrogen, yields a peak at m/z 177 Da. Further loss of water followed by removal of CO, yields a peak at m/z 159 and 131 respectively (Figure 5).

Acetylbarlerin was observed as sodium adduct $[M+Na]^+$ at m/z 513 and potassium adduct at m/z 529. Fragmentation pattern of potassium adduct of Acetylbarlerin at m/z 529 $[M+K]^+$, shows formation of barlerin at m/z 487 $[M+K]^+$ created by removal of acetyl group. Ion at m/z 427 is formed by loss of $COOCH_3$ group and ion at m/z 265 by loss of glucose molecule (Figure 5).

3.3 Identification of phenolic compounds

Major groups of phenolic compounds identified were phenolic acids, flavonoids and phenylethanoid glycosides. Phenolic acids observed were melilotic acid, caffeic acid, ferulic acid, vanillic acid, coumaric acid, gallic acid, syringic acid, methylgallic acid and dihydroferulic acid (Table 1). Their identity was confirmed by HRMS analysis and MS/MS data

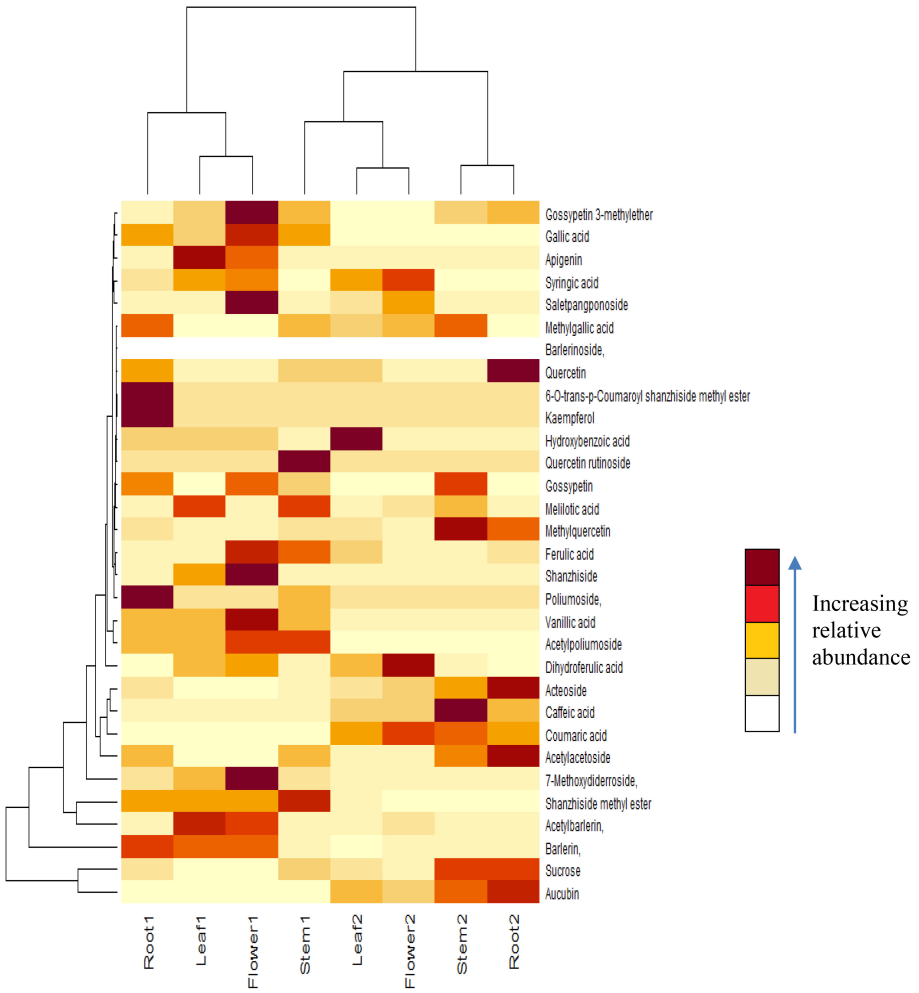


FIGURE 1
Heat map of compounds identified in different parts of *B prionitis* and *B cristata*. Root 1, Leaf 1, Flower 1 and Stem 1 are parts of *B prionitis*, whereas Leaf 2, Stem 2, Flower 2 and Root 2 are parts of *B cristata*.

comparison. MS2 signals indicating loss of water molecule (18Da) was the most important diagnostic ion generated from the parent ion peak; for example caffeic acid with m/z 181 $[M+H]^+$ gave MS2 signal at 163 $[M\text{-}H_2O]^+$. Ferulic acid with m/z 195 $[M+H]^+$ gave characteristic fragment ion peak at m/z 177, indicating loss of water molecule.

Flavonoids detected in *Barleria* were low in abundance (Table 1). Apigenin was detected by generation of fragment ion at m/z 275 formed by loss of water molecule from parent ion at m/z 293 $[M+Na]^+$ and ion at m/z 199 formed by elimination of C_6H_5OH (phenol) group from the parent ion. MS2 signal for Gossypetin 3-methylether was also observed due to its higher abundance (heat map). Fragment with m/z 340 was generated by loss of water molecule and 353 by loss of CH_3OH group from the parent ion. A flavonoid glycoside, quercetin rutinoside was identified by its diagnostic fragment ion 285 and 325 formed by breakage of sugar group from the parent ion.

Acetoside/Verbascoside, a caffeoyl phenylethanoid glycoside in which the phenylpropanoid caffeic acid and the phenylethanoid hydroxytyrosol form an ester and an ether bond respectively to the rhamnose part of a disaccharide, was detected in all parts of the plant, however it's abundance was high in *B. cristata* (Heat map). It was observed as sodium adduct $[M+Na]^+$ at m/z 647 and potassium adduct $[M+K]^+$ at m/z 663. Fragment ions of acetoside at m/z 663 $[M+K]^+$, generated characteristic ion of caffeoyl group (m/z 163). Ion at m/z 517 was produced by loss of caffeoyl group and addition of water molecule to the parent ion $[M\text{-caffeoyl}+H_2O+K]^+$. Ion at m/z 495 was generated by loss of rahmnose group from the parent ion with replacement of potassium ion with water molecule as adduct $[M\text{-Rha}+H_2O\text{-}K]^+$. Cleavage of hydroxyltyrosol group was observed, as indicated by appearance of ion at m/z 177 $[hydroxyltyrosol+K]^+$. Dissacharide sugar group was cleaved from the parent molecule to generate fragment at m/z 365

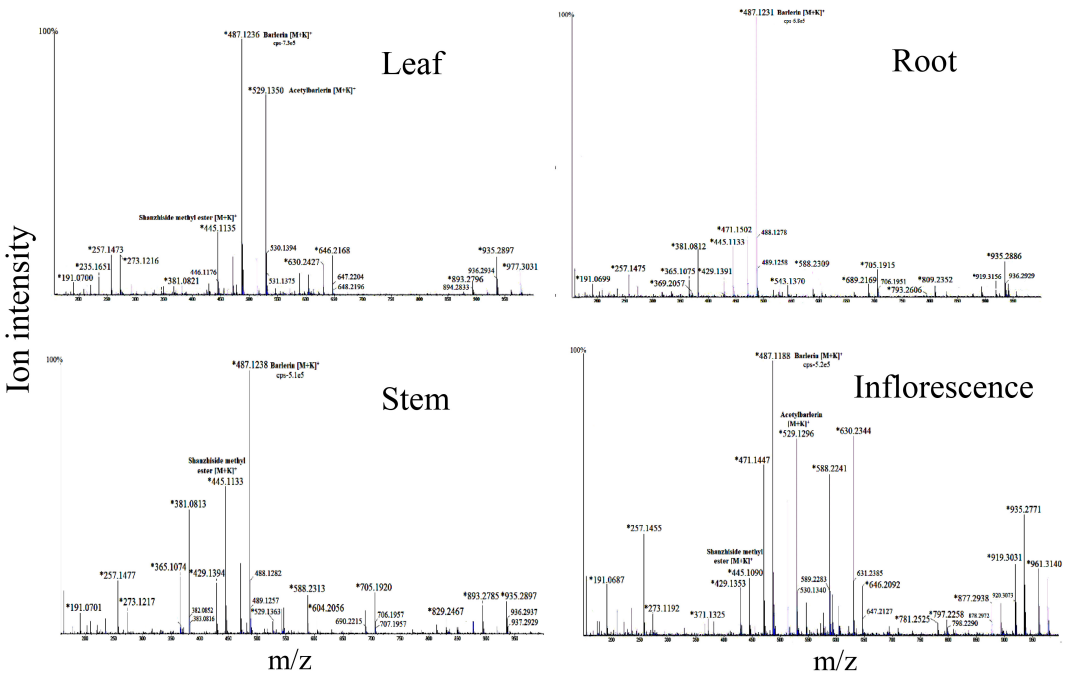


FIGURE 2
Mass spectra of root, leaf, stem and inflorescence of *Barleria prionitis* showing MS1 ions in positive ion mode. Ions marked with asterisk (*) symbol are the parent ion and without the symbol are their isotopes.

$[sugar+K]^+$ and addition of water generated fragment at m/z 383 (Figure 5).

Acetylacetoside was observed in significant amount in root and stem of the plant forming sodium and potassium ions at m/z 689 and 705 respectively. It was identified by formation of acetoside ion after removal of acetyl group. Poliumoside, a caffeoylated phenylpropanoid glycoside was also detected in significant amount in root and stem of *B.prionitis*, generating ions at m/z 793 $[M+Na]^+$ and 809 $[M+K]^+$. It was identified by its characteristic fragment ion m/z 630 formed by removal of caffeoyl group (163) from the parent ion (m/z 793). Acetylpoliumoside was detected by presence of its ions at m/z 893 $[M+Na]^+$ and 914 $[M+K]^+$ formation of poliumoside ion by loss of acetylgroup (Table 1).

Two ion peaks, at m/z 527 and 543 were identified as caffeic acid derivative because they generated fragment ion peak at m/z 163, indicating loss of caffeoyl group. Both signals were assumed to be generated by a single compound forming adduct with sodium and potassium ion. Similarly, ion with m/z 461, 775 and 795 observed in *B.cristata* also generated caffeoyl group fragment at m/z 163. MS/MS spectra of ions at m/z 257 and 273 were quite similar showing loss of hydroxyl group (18Da), indicating them to be a phenolic compound forming sodium and potassium adduct. Similarly, ions with m/z 604 and 630 observed in *B.prionitis* were also observed as signal of a similar compound indicating loss of water molecule (18Da) and glycosyl group; showing their probability to be a phenolic glycoside (Figures 2, 4).

3.4 Antioxidant and antibacterial activity

In our experiments, a strong antioxidant activity was observed in all parts of the plant by DPPH and FRAP assay, demonstrating free radical scavenging activity of the compounds (Table 2). A comparative analysis indicates highest activity in root and leaf of *B.cristata*. In *B.prionitis* highest activity was shown by inflorescence, followed by root. Antibacterial activity was shown in all parts of the plant, with root being the most effective one in both the species (Table 2). All parts were effective against *S. aureus-aureus*, with highest values shown by root extract. *B.cristata*, leaf, root and inflorescence were effective against most of the tested bacteria, whereas *B.prionitis* root extract was only effective against *P.aeruginosa, B.subtilis* and *E.coli*.

3.5 Correlation analysis

Correlation study was conducted to identity the contribution of various compounds in antioxidant and antibacterial activity. The correlation coefficients between the ion intensity peaks of the compounds and the bioactivity results, including FRAP assay, IC50 value for DPPH assay and antibacterial activity against *S. aureus-aureus* was calculated against 95% confidence level (Figure 6). As we can see from the table that most of the phenolic acid and flavonoids showed a positive correlation with

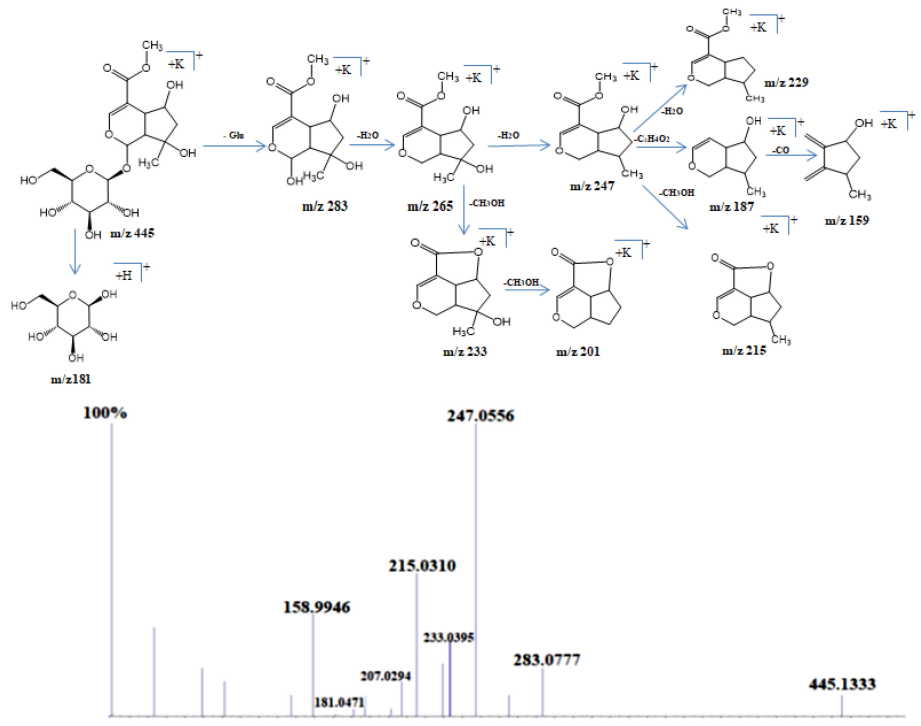


FIGURE 3
Schematic representation of fragmentation pattern of Shanzhiside methylester generated by MS/MS analysis.

the antioxidant activity. Phenylethanoid glycoside acetoside, poliumoside and acetylpoliumoside showed a positive correlation with antioxidant activity. Some of the phenolic compounds were positively correlated with antibacterial activity, among which caffeic acid and coumaric acid were the most important one. Among flavonoids gossypetin was positively correlated, phenylethanoid glycoside including acetoside, acetylacetoside and poliumoside were also positively correlated and among iridoid glycoside barlerin was slightly positively correlated with the antibacterial activity.

3.6 Principal component analysis

Thirteen ion peaks were ubiquitously distributed in all parts of both the plants, therefore they were selected as marker signals for PCA analysis, among which shanzhiside methylester, barlerin, acetylbarlerin and acetoside were the identified compounds. The total percentage of variation of data set was explained using four principal components (PC1, PC2, PC3 and PC4). The results showed that PC1 explains 48.26 percent, PC2 explains 29.29 percent, PC3 13.21 percent and PC4 explains 4.42 percent variation of dataset (Table 3). PC1 explains the maximum variation as compared to the other PCA in whole dataset. Scatter plot represents the variation of first two principal components PC1 and PC2 within the replicate 1 to 5 (Figure 7), demonstrating non-linear pattern of data which does not show normality. The given biplot in Figure 8 represents the first principal component which combines flower, leaf and some part of variation in stem and root with second principal component. The Screen plot represents a decreasing pattern of cumulative contribution of variation for eight principal components (Figure 9).

4 Discussion

A rapid metabolomic approach was used in the present research work for chemical profiling of active compounds in *B.prionitis* and *B.cristata*. X500R QTOF system uses information dependent acquisition (IDA) method to collect high resolution accurate mass data (HRMS) on both the parent and their fragment ions. IDA performs a non- targeted screen of MS1 fingerprint and separates MS1 signals into small ion currents, further selecting the most abundant ion current from the TOF MS total ion chromatogram to perform MS/MS fragmentation (Whitman and Lynch, 2019). The whole process was accomplished in a single run, thereby overcoming separation requirement of the extract. Compounds were identified on the basis of their HRMS analysis, fragmentation pattern and isotopic distribution, as reported in literature (Heffels et al., 2017;

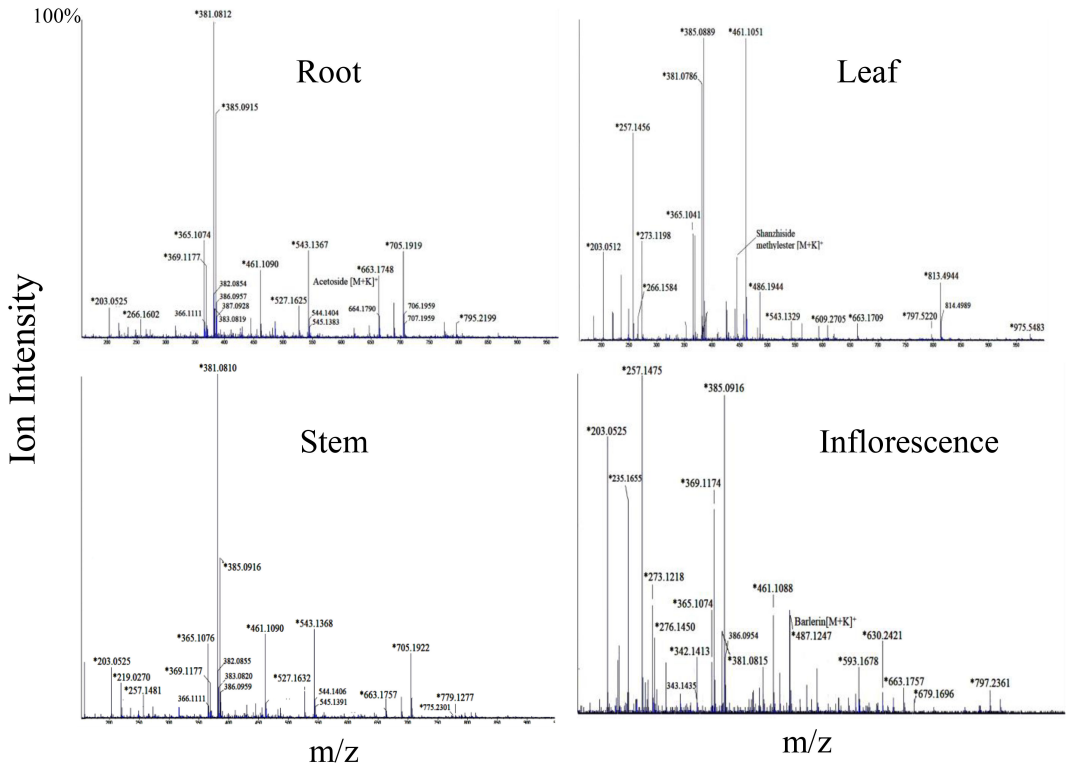


FIGURE 4
Mass spectra of root, leaf, stem and inflorescence of *Barleria cristata* showing MS1 ions in positive ion mode. Ions marked with asterisk (*) symbol are the parent ion and without the symbol are their isotopes.

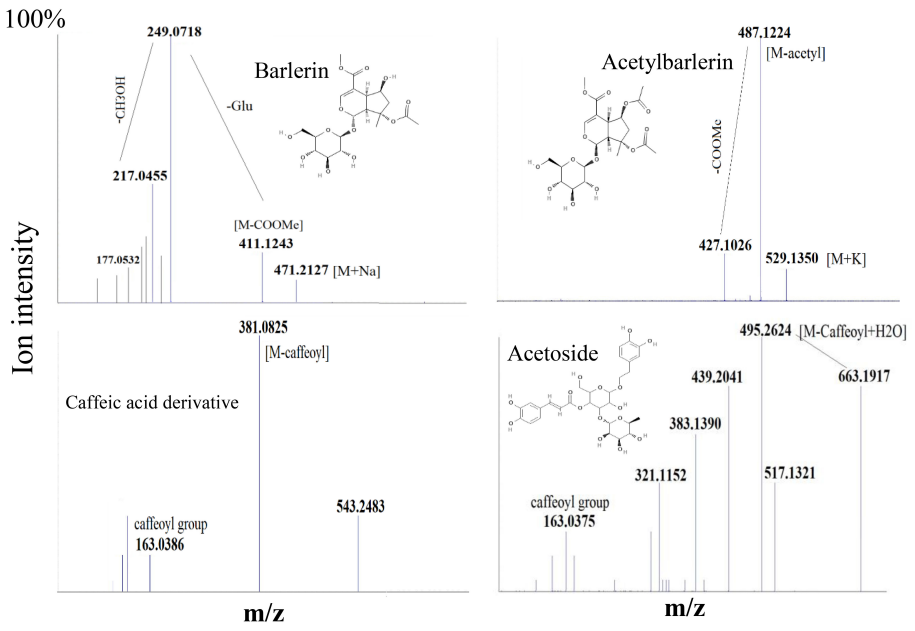


FIGURE 5
MS/MS spectra of some selected ions, showing fragmentation pattern of Barlerin, Acetylbarlerin, Caffeic acid derivative and Acetoside.

TABLE 2 Antioxidant and antibacterial activity of extracts obtained from different parts of *B.prionitis* and *B.cristata*.

	Antioxidant activity		Antibacterial activity (extracts 10 mg/ml)- zone of inhibition (cm)				
	FRAP assay (mgAA/g)	DPPH assay IC50 µg/mL	Staphylococcus aureus-aureus	Klebsiella pneumoniae	Psuedomonas aeruginosa	Bacillus subtilis	Escherichia coli
BYL	112.71 ± 5.67	62.2 ± 4.67	1.2 ± 0.3	–	–	–	–
BYR	141.92 ± 10.34	48.3 ± 3.45	2.0 ± 0.5	–	1.2 ± 0.1	1.3 ± 0.2	1.2 ± 0.2
BYS	104.85 ± 8.23	77.6 ± 6.35	1.1 ± 0.2	–	–	–	–
BYF	156.76 ± 12.24	38.5 ± 3.56	1.4 ± 0.3	–	–	–	–
BWL	174.69 ± 13.45	29.3 ± 4.53	1.7 ± 0.2	1.4 ± 0.2	1.5 ± 0.3	1.4 ± 0.3	1.2 ± 0.1
BWR	208.74 ± 14.36	18.4 ± 3,21	2.2 ± 0.4	1.2 ± 0.1	1.7 ± 0.4	1.5 ± 0.3	1.4 ± 0.2
BWS	143.04 ± 12.66	47.2 ± 6.57	1.6 ± 0.1				
BWF	128.17 ± 12.48	65.5 ± 6.55	1.7 ± 0.1	1.5 ± 0.4	1.4 ± 0.2		1.3 ± 0.1
Ascorbic acid		9.2 ± 2.38					
Streptomycin (1mg/ml)			3.1 ± 0.6	2.7 ± 0.5	2.3 ± 0.5	3.2 ± 0.6	2.6 ± 0.3

Values are expressed as mean ± SE (n = 5).

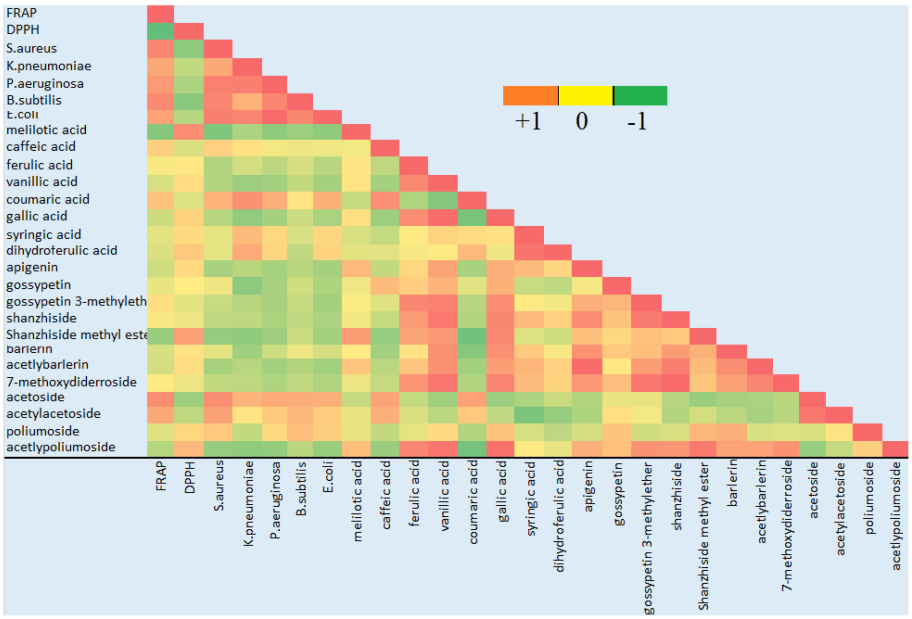


FIGURE 6
Correlation analysis of the identified compounds with antioxidant and antibacterial activity. Green color indicates negative correlation, red color indicates positive correlation and shades of yellow show no correlation.

Kite, 2020). Most of the natural compounds are made up of carbon, hydrogen, oxygen and nitrogen. Their protonated molecules have similar isotope patterns and the intensities of their isotopic ions are always in descending order: M > M+1 > M+2, following the ratio 10:3:1 (Zhang et al., 2012). Many of the detected compound such as shanzhiside methylester and barlerin showed a similar isotopic abundance, which helped us in further identification.

TABLE 3 Variance explained by the PC1-PC8 with standard deviation and cumulative proportion.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Standard Deviation	1.9648	1.5308	1.0282	0.5944	0.4652	0.3441	0.2125	0.0752
Proportion of variance	0.4826	0.2929	0.1321	0.0442	0.0271	0.0148	0.0057	0.0007
Cumulative proportion	0.4826	0.7755	0.9076	0.9518	0.9786	0.9937	0.9992	1.0000

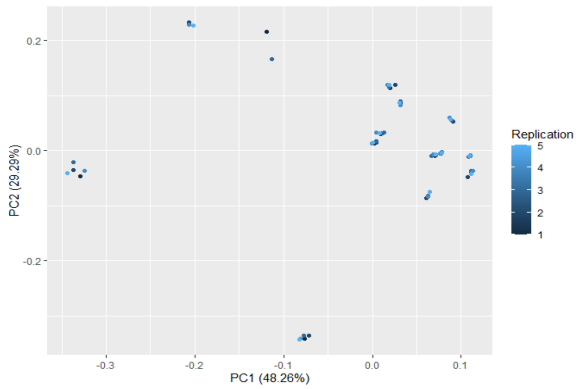


FIGURE 7
Scatter plot of first two principal components PC1 and PC2 with the replication of 1 to 5 presented. This plot shows variability in the distribution pattern of metabolites.

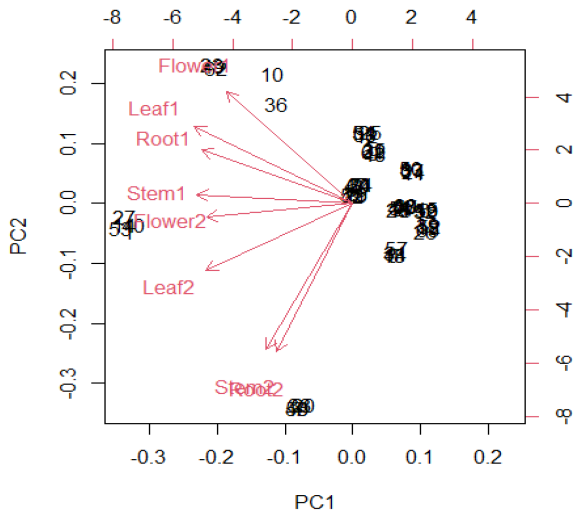


FIGURE 8
In R software biplot function has been used to plot PCA. The given biplot represents the first principal component which combine (flower, leaf) and little part of variation in (stem, root) in second principal component.

Iridoid glycosides and phenolic compounds have been reported from aerial parts of *Barleria* species (Gangaram et al., 2022), however as shown in our study, they are distributed in all parts of the plant, including leaf, root, stem and flower (Figure 2). The study also helped us to analyse the distribution pattern and abundance of these compounds in both the plants. Principal component analysis (PCA), conducted according to the methods described in literature (Abdi and Williams, 2010; Bro and Smilde, 2014; Lever et al., 2017), showed variation in the distributed pattern of identified metabolites. We can observe that *B. prionitis* had high relative abundance of iridoid glycosides, particularly shanzhiside methy ester, barlerin and acetylbarlerin, whereas *B. cristata* showed higher relative abundance of acetoside. These metabolites are active compounds with broad spectrum of biological activities. Extracts enriched with iridoid glycosides have shown glutathione S-transferase inhibitory activity, acetylcholinesterase inhibitory activity, free radical scavenging, antimicrobial, anti-inflammatory, immunomodulatory and gastroprotective activities (Kosmulalage et al., 2007; Ata et al., 2009; Ghule and Yeole, 2012; Jaiswal et al., 2014). Shanzhiside methyl ester and barlerin enriched fraction have shown to modulate specific and non-specific immune response in *in vivo* studies (Ghule and Yeole, 2012). Shanzhiside methyl ester have shown to reduce neuropathic pain by stimulating spinal microglial β-endorphin expression (Zhang et al., 2018), it also possesses anti-inflammatory properties and have shown to provide protection against depression by inhibiting inflammation (Sun et al., 2022). Iridoid glycoside 6-O-transp-coumaroyl-8-*O*-acetylshanzhiside methyl ester and its *cis* isomer have shown activity against respiratory syncytial virus (Chen et al., 1998). Similarly, phenylethanoid glycosides, a group of phenolic compound made up of phenylethyl alcohol, caffeic acid and glycosyl groups (Leea et al., 2016) shown in our results have reported biological activity. Barlerinoside, a phenylethanoid glycoside, isolated from aerial parts of *Barleria* sps. have shown to possess glutathaione S transferase and acetylcholinesterase inhibitory activity (Ata et al., 2009). Verbascoside/acetoside, a phenylpropanoid glycoside identified in *Barleria* species is known for its antioxidant, analgesic, anticancer, anti-inflammatory and photoprotective activity (Vertuani et al., 2011). Identification of these active metabolites in different parts of both the species will help in exploiting full pharmacological potential of the plant.

Some studies have demonstrated antioxidant and antibacterial activity of *Barleria* sps. (Amoo et al., 2009; Sunil

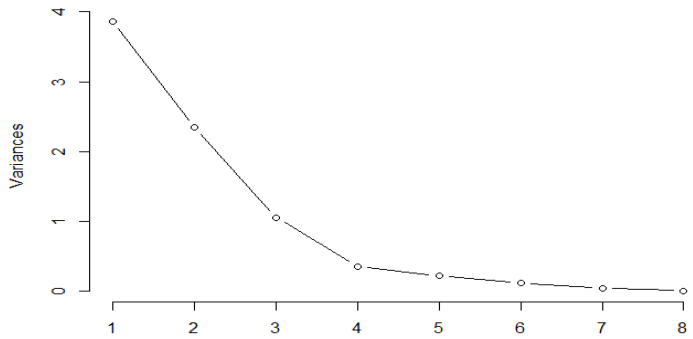


FIGURE 9
Screen plot represents decreasing pattern of cumulative contribution of variation for eight principal components.

et al., 2010; Chavan et al., 2014), however our study shows a comparative analysis, proving *B.cristata* to be more effective against tested bacterial species. As we can see that most of the phenolic acid and flavonoids showed a positive correlation with the antioxidant activity. Phenolic compounds are widely distributed plant substances and have been considered as significant contributors to antioxidant activity (Bittencourt et al., 2015; Kozyraa et al., 2019; Pannakal et al., 2022; Singh, 2023). Some of the phenolic compounds were positively correlated with antibacterial activity, among which caffeic acid and coumaric acid were strongly correlated. Earlier reports also confirms their role as antimicrobial (Bittencourt et al., 2015; Merlani et al., 2019). Acetoside has an antimicrobial activity, notably against *Staphylococcus aureus* (Bazzaz et al., 2018). None of the detected iridoid glycoside showed to be a major contributor in antioxidant and antibacterial activity, except barlerin having positive correlation with the tested bacterial strain. Owing to the abundance of these compounds in the tested extract it could be concluded that either alone or synergistically they might be responsible for the biological activity.

5 Conclusions

Phytochemical investigation of *B. prionitis* and *B. cristata* led to the identification of iridoid glycosides and phenolic compounds in different parts of the plant. Most abundant iridoid glycosides were barlerin, acetylbarlerin and shanzhiside methylester, whereas acetoside and acetylacetoside were the important phenolic compound present in the plants. Many of the chemical compounds are reported for the first time in *Barleria* species, such as ferulic acid, caffeic acid, kaempferol, acetlyacetoside. Most of the compounds detected in root, stem and inflorescence have not been reported earlier. Chemical compounds were found to be distributed in all parts of the plant with variation in their presence and abundance. They were shown to be correlated with antioxidant and antibacterial activity.

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

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Acknowledgments

The authors acknowledge Banaras Hindu University (BHU), Varanasi, India, for providing laboratory facilities and fund support. We also acknowledge Department of Chemistry, BHU, India for ESI-MS 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

References

Abdi, H., and Williams, L. J. (2010). Principal component analysis. *Wiley Interdiscip. Rev. Computat. Stat.* 2 (4), 433–459. doi: 10.1002/wics.101

Amoo, S. O., Finnie, J. F., and Staden, J. V. (2009). *In vitro* pharmacological evaluation of three *Barleria* species. *J. Ethnopharmacol.* 121 (2), 274–277. doi: 10.1016/j.jep.2008.10.035

Ata, A., Kalhar, K. S., and Samarasekera, R. (2009). Chemical constituents of barleria prionitis and their enzyme inhibitory and free radical scavenging activities. *Phytochem. Lett.* 2, 37–40. doi: 10.1016/j.phytol.2008.11.005

Bazzaz, B. S. F., Khameneh, B., Ostad, M. R. Z., and Hosseinzadeh, H. (2018). Evaluation of antibacterial activity of verbascoside, lemon verbena extract and caffeine in combination with gentamicin against drug-resistant *Staphylococcus aureus* and *Escherichia coli* clinical isolates. *Avicenna J. Phytomed.* 8 (3), 246–253.

Bittencourt, M. L. F., Ribeiro, P. R., Franco, R. L. P., Hilhorst, H. W. M., de Castro, R. D., and Fernandez, L. G. (2015). Metabolite profiling, antioxidant and antibacterial activities of Brazilian propolis: Use of correlation and multivariate analyses to identify potential bioactive compounds. *Food Res. Int.* 76, 449–457. doi: 10.1016/j.foodres.2015.07.008

Bro, R., and Smilde, A. K. (2014). Principal component analysis. *Anal. Methods* 6 (9), 2812–2831. doi: 10.1039/C3AY41907J

Chavan, C., Mulik, S., Chavan, M., Adnaik, R., and Patil, P. (2014). Screening of antioxidant activity and phenolic content of whole plant of *Barleria prionitis* Linn. *Int. J. Res. Ayurveda Pharm.* 2 (4), 1313–1319.

Chen, J. L., Blanc, P., Stoddart, C. A., Bogan, M., Rozhon, E., Parkinson, N., et al. (1998). New iridoids from the medicinal plant *Barleria prionitis* with potent activity against respiratory syncytial virus. *J. Nat. Prod.* 61 (10), 1295–1297. doi: 10.1021/np980086y

Dheer, R., and Bhatnagar, P. (2010). A study of the antidiabetic activity of *Barleria prionitis* Linn. *Indian J. Pharmacol.* 42 (2), 70–73. doi: 10.4103/0253-7613.64493

Gangaram, S., Naidoo, Y., Dewir, Y. H., and El-Hendawy, S. (2022). Phytochemicals and biological activities of barleria (Acanthaceae). *Plants* 11 (1), 82–118. doi: 10.3390/plants11010082

Ghule, B. V., and Yeole, P. G. (2012). *In vitro* and *in vivo* immunomodulatory activities of iridoids fraction from *Barleria prionitis* Linn. *J. Ethnopharmacol.* 141 (1), 424–443. doi: 10.1016/j.jep.2012.03.005

Guo, C., Yang, J., Wei, J., Li, Y., Xu, J., Jiang, Y., et al (2003). Antioxidant activities of peel, pulp and seed fractions of common fruits as determined by FRAP assay. *Nutrition research,* 23, 1719–1726. doi: 10.1016/j.nutres.2003.08.005

Gupta, R. S., Kumar, P., Dixit, V. P., and Dobhal, M. P. (2000). Antifertility studies of the root extract of the *Barleria prionitis* Linn in male albino rats with special reference to testicular cell population dynamics. *J. Ethnopharmacol.* 70 (2), 111–117. doi: 10.1016/S0378-8741(99)00150-6

Heffels, P., Müller, L., Schieber, A., and Weber, F. (2017). Profiling of iridoid glycoside in vaccinium species by UHPLC-MS. *Food Res. Int.* 100 (3), 462–468. doi: 10.1016/j.foodres.2016.11.018

Jaiswal, S. K., Dubey, M. K., Das, S., and Rao Ch, V. (2014). Gastroprotective effect of the iridoid fraction from *Barleria prionitis* leaves on experimentally-induced gastric ulceration. *Chin. J. Nat. Med.* 12 (10), 738–744. doi: 10.1016/S1875-5364(14)60113-8

Kanchanapooma, T., Kasaia, R., and Yamasakia, K. (2002). Iridoid glucosides from barleria lupulina. *Phytochem* 58, 337–341. doi: 10.1016/S0031-9422(01)00236-9

Kaur, B., and Kaur, N. (2016). Metabolic fingerprinting of different populations of phyllanthus niruri l. from punjab using electrospray ionization mass spectrometry (ESI–MS). *Med. Chem. Res.* 25, 2798–2821. doi: 10.1007/s00044-016-1674-z

Kite, G. C. (2020). Characterisation of phenylethanoid glycosides by multiple-stage mass spectrometry. *Rapid Commun. Mass Spectrom.* 34 (S4), 8563–8572. doi: 10.1002/rcm.8563

Kosmulalage, K. S., Zahid, S., Udenigwe, C. C., Akhtar, S., Ata, A., and Samarasekera, R. (2007). Glutathione s transferase, acetylcholinesterase inhibitory and antibacterial activities of chemical constituents of barleria prionitis. *Z. Naturforsch.* 62, 580–586. doi: 10.1515/znb-2007-0417

Kozyraa, M., Komstab, L., and Wojtanowskia, K. (2019). Analysis of phenolic compounds and antioxidant activity of methanolic extracts from inflorescences of *Carduus* sp. *Phytochem. Lett.* 31, 256–262. doi: 10.1016/j.phytol.2019.04.012

Leea, S. R., Clardyb, J., Sengerc, D. R., Caod, S., and Kima, K. H. (2016). Iridoid and phenylethanoid glycosides from the aerial part of barleria lupulina. *Rev. Bras. Farmacognosia.* 26, 281–284. doi: 10.1016/j.bjp.2016.01.002

Lever, J., Krzywinski, M., and Altman, N. (2017). Points of significance: Principal component analysis. *Nat. Methods* 14 (7), 641–643. doi: 10.1038/nmeth.4346

Liu, W., Cao, L., Jia, J., Li, H., Li, W., Li, J, et al. (2021). Rapid chemome profiling of artemisia capillaris thunb. using direct infusion-mass spectrometry. *J. Tradit. Chin. Med. Sci.* 8, 327–335. doi: 10.1016/j.jtcms.2021.10.003

Møller, J. K. S., Catharino, R. R., and Eberlin, M. N. (2007). Electrospray ionization mass spectrometry fingerprinting of essential oils: Spices from the labiatae family. *Food Chem.* 100, 1283–1288. doi: 10.1016/j.foodchem.2005.10.013

Mauri, P., and Pietta, P. (2000). Electrospray characterization of selected medicinal plant extracts. *J. Pharm. BioMed. Anal.* 23, 61–68. doi: 10.1016/S0731-7085(00)00264-8

Merlani, M., Barbakadze, V., Amiranashvili, L., Gogilashvili, L., Poroikov, V., Petrou, A, et al. (2019). New caffeic acid derivatives as antimicrobial agents: design, synthesis, evaluation and docking. *Curr. Top. Med. Chem.* 19 (4), 292–304. doi: 10.2174/1568026619666190122152957

Pannakal, S. T., Eilstein, J., and Prasad, A. (2022). Comprehensive characterization of naturally occurring antioxidants from the twigs of mulberry (Morus alba) using on-line high-performance liquid chromatography coupled with chemical detection and high-resolution mass spectrometry. *Phytochem. Anal.* 33, 105–114. doi: 10.1002/pca.3072

Sharma, P., Sharma, G. N., Shrivastava, B., and Jadhav, H. R. (2014). Evaluation of antioxidant potential of barleria prionitis leaf and stem. *Am. J. Phytomed. Clin. Therapeut.* 2 (11), 1177–1186.

Singh, S. (2016). Enhancing phytochemical levels, enzymatic and antioxidant activity of spinach leaves by chitosan treatment and an insight into the metabolic pathway using DART-MS technique. *Food Chem.* 199, 176–184. doi: 10.1016/j.foodchem.2015.11.127

Singh, S. (2023). Salicylic acid elicitation improves antioxidant activity of spinach leaves by increasing phenolic content and enzyme levels. *Food Chem. Adv.* 2, 1–11. doi: 10.1016/j.focha.2022.100156

Singh, B., Bani, S., Gupta, D. K., Chandan, B. K., and Kaul, A. (2003). Antiinflammatory activity of 'TAF' an active fraction from the plant *Barleria prionitis* Linn. *J. Ethnopharmacol.* 85 (2-3), 187–193. doi: 10.1016/S0378-8741(02)00358-6

Singh, B., Chandan, B., Prabhakar, A., Taneja, S., Singh, J., and Qazi, G. (2005). Chemistry and hepatoprotective activity of an active fraction from *Barleria prionitis* Linn. *In Exp. animals Phytother. Res.* 19 (5), 391–404. doi: 10.1002/ptr.1509

Singh, S., Nagappan, S., and Verma, S. K. (2020). Antimicrobial haplindole alkaloids as chemical marker for rapid identification of stigonematales (Cyanobacteria). *Anal. Chem. Lett.* 10 (5), 602–608. doi: 10.1080/22297928.2020.1852106

Sunil, K. J., Mukesh, K. D., Sanjeeb, D., Arti, R. V., and Rao Ch, V. (2010). A comparative study on total phenolic content, reducing power and free radical scavenging activity of aerial parts of barleria prionitis. *Int. J. Phytomed.* 2 (2), 155–159. doi: 10.5138/ijpm.2010.0975.0185.02024

Sun, Z., Zhan, H., Wang, C., and Guo, P. (2022). Shanzhiside methylester protects against depression by inhibiting inflammation *via* the miRNA-155-5p/SOCS1 axis. *Psycopharmacol* 239 (7), 2201–2213. doi: 10.1007/s00213-022-06107-7

Tanase, C., Cosarca, S., and Muntean, D. (2019). A critical review of phenolic compounds extracted from the bark of woody vascular plants and their potential biological activity. *Molecules* 24 (6), 1182–1200. doi: 10.3390/molecules24061182

Taneja, S. C., and Tiwari, H. P. (1975). Structures of two new iridoids from barleria prionitis linn. *Tetrahedron. Lett.* 16 (24), 1995–1998. doi: 10.1016/S0040-4039(00)72344-6

Verma, P. K., Sharma, A., Joshi, S. C., Gupta, R. S., and Dixit, V. P. (2005). Effect of isolated fractions of *Barleria prionitis* root methanolic extract on reproductive function of male rats: Preliminary study. *Fitoterapia* 76 (5), 428–432. doi: 10.1016/j.fitote.2005.03.007

Vertuani, S., Beghelli, E., Scalambra, E., Malisardi, G., Copetti, S., Toso, R.D, et al. (2011). Activity and stability studies of verbascoside, a novel antioxidant, in dermo-cosmetic and pharmaceutical topical formulations. *Molecules* 16 (8), 7068–7080. doi: 10.3390/molecules16087068

Whitman, J. D., and Lynch, K. L. (2019). Optimization and comparison of information-dependent acquisition (IDA) to sequential window acquisition of all theoretical fragment ion spectra (SWATH) for high-resolution mass spectrometry in clinical toxicology. *Clin. Chem.* 65 (7), 862–870. doi: 10.1373/clinchem.2018.300756

Widyowatia, R., Tezukaa, Y., Miyaharab, T., Awalea, S., and Kadotaa, S. (2010). Alkaline phosphatase (ALP) enhancing iridoid glucosides from the Indonesian medicinal plant barleria lupulina. *Nat. Prod. Commun.* 5 (11), 1711–1716. doi: 10.1177/1934578X1000501101

Zhang, W., Bai, Y., and Qiao, Y. (2018a). 8-O-Acetyl shanzhiside methylester from lamiophlomis rotata reduces neuropathic pain by inhibiting the ERK/TNF-α pathway in spinal astrocytes. *Front. Cell Neurosci.* 8, 12:54. doi: 10.3389/fncel.2018.00054

Zhang, D., Gao, Y., Jiang, S., Chen, Y., Zhang, Y., and Pan, Z. (2018b). The similarity and variability of the iridoid glycoside profile and antioxidant capacity of aerial and underground parts of lamiophlomis rotate according to UPLC-TOF-MS and multivariate analyses. *RSC* 8, 2459–2468. doi: 10.1039/c7ra10143k

Zhang, Y., Pan, J., Zhong, J., Wang, Y., Fan, X., and Cheng, Y. (2012). Virtual separation of phytochemical constituents by their adduct-ion patterns in full mass spectra. *J. Chromatogr. A.* 1227, 181–193. doi: 10.1016/j.chroma.2012.01.006

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 18 January 2023
DOI 10.3389/fpls.2023.1073848

OPEN ACCESS

EDITED BY
Ramesha Thimmappa,
Amity University, India

REVIEWED BY
Shihai Xing,
Anhui University of Chinese Medicine, China
Dongming Ma,
Guangzhou University of Chinese Medicine, China

*CORRESPONDENCE
Chuanguang Liu
guyliu@126.com
Jihua Wang
wangjihua@gdaas.cn

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

RECEIVED 19 October 2022
ACCEPTED 04 January 2023
PUBLISHED 18 January 2023

CITATION
Zhang W, Zeng Y, Jiao M, Ye C, Li Y, Liu C and Wang J (2023) Integration of high-throughput omics technologies in medicinal plant research: The new era of natural drug discovery.
Front. Plant Sci. 14:1073848.
doi: 10.3389/fpls.2023.1073848

Integration of high-throughput omics technologies in medicinal plant research: The new era of natural drug discovery

Wenting Zhang[1,2], Yuan Zeng[3,4], Meng Jiao[5], Chanjuan Ye[6], Yanrong Li[5], Chuanguang Liu[6]* and Jihua Wang[1,2]*

[1]Guangdong Provincial Key Laboratory of Crops Genetics & Improvement, Crops Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China, [2]Guangdong Provincial Engineering & Technology Research Center for Conservation and Utilization of the Genuine Southern Medicinal Resources, Guangzhou, China, [3]School of Plant and Environmental Sciences, Virginia Tech, VA, Blacksburg, United States, [4]Southern Piedmont Agricultural Research and Extension Center, Virginia Tech, VA, Blackstone, United States, [5]College of Life Sciences, South China Agricultural University, Guangzhou, China, [6]Rice Research Institute, Guangdong Rice Engineering Laboratory, Guangdong Key Laboratory of New Technology in Rice Breeding, Guangdong Academy of Agricultural Sciences, Guangzhou, China

Medicinal plants are natural sources to unravel novel bioactive compounds to satisfy human pharmacological potentials. The world's demand for herbal medicines is increasing year by year; however, large-scale production of medicinal plants and their derivatives is still limited. The rapid development of modern technology has stimulated multi-omics research in medicinal plants, leading to a series of breakthroughs on key genes, metabolites, enzymes involved in biosynthesis and regulation of active compounds. Here, we summarize the latest research progress on the molecular intricacy of medicinal plants, including the comparison of genomics to demonstrate variation and evolution among species, the application of transcriptomics, proteomics and metabolomics to explore dynamic changes of molecular compounds, and the utilization of potential resources for natural drug discovery. These multi-omics research provide the theoretical basis for environmental adaptation of medicinal plants and allow us to understand the chemical diversity and composition of bioactive compounds. Many medicinal herbs' phytochemical constituents and their potential health benefits are not fully explored. Given their large diversity and global distribution as well as the impacts of growth duration and environmental factors on bioactive phytochemicals in medicinal plants, it is crucial to emphasize the research needs of using multi-omics technologies to address basic and applied problems in medicinal plants to aid in developing new and improved medicinal plant resources and discovering novel medicinal ingredients.

KEYWORDS

medicinal plant, high-throughput omics, biosynthesis pathways, active ingredients, phytochemicals

Introduction

Many medicinal plant-derived alkaloid, terpene, polyphenol, coumarin, and saponin have received increasing attentions from the pharmaceutical industries due to their potent antioxidant, antibacterial, antiphlogistic, anticancer, and antidiabetic activities (Mumtaz et al., 2017; Pandita et al., 2021). However, large scale production of these active ingredients was limited due to the lack of medicinal plant genomic information. As the advancement of high-throughput sequencing technology, genomes of many medicinal plants are assembled, leading to functional characterization of genes that involve in specific secondary biosynthesis pathways. For example, genes in the biosynthesis pathway of heterologously producing saponin, a biological compound displaying antimicrobial and anti-inflammatory activities, were characterized following the publication of Chinese ginseng (*Panax notoginseng*) genomes (Fan et al., 2020; Jiang et al., 2021). The biosynthesis, regulation, and transportation of a bioactive component, benzylisoquinoline alkaloid (BIA) in landraces of Chinese opium poppy was revealed upon the availability of its genome (Hu et al., 2018).

Incorporating other high-throughput sequencing or analytical techniques in metabolomics, proteomics, and transcriptomics into medicinal plant research can aid in discovering functional genes, key metabolites, biological elements with pharmacological potential and molecular markers associated with phytochemical compounds. For example, mining transcriptome and metabolome profiles of mayapple (*Podophyllum hexandrum*), a plant with anticancer compounds, identified genes involved in biosynthetic pathway of podophyllotoxin (Lau and Sattely, 2015). *Abelmoschus esculentus* is a medicinal plant containing a large amount of active ingredients such as anthocyanins, flavonoids, polysaccharides, and terpenoids. Based on the transcriptome data of *A. esculentus*, one significant marker was detected by association analysis of fruit color (anthocyanin content) which may be used for the genetic improvement of *A. esculentus* (An et al., 2022). Analyzing proteomic datasets of *Dendrobium huoshanense* led to discovery of crotonylated proteins in the Calvin cycle, photosynthesis, and alkaloid and polysaccharide biosynthesis metabolic pathways (Wu et al., 2022a).

To date, multi-omics approaches have promoted the development of large databases on a specific medicinal plant species or multiple species spanning their genome, transcriptome, proteome, metabolome, or phytochemicals (Table 1). For example, Ginseng Genome Database is the first comprehensive database among other herbs due to its well-established genome information (Jayakodi et al., 2018). The Global Pharmacopoeia Genome Database (GPGD) is a database contains 2,203 organelle genomes from 674 species, 55 whole genomes from 49 species, and 9,682 transcriptome datasets from 350 species (Liao et al., 2021). MepmiRDB, the first MicroRNA (miRNA) database of medicinal plants, gathers miRNA information of 29 species such as ginseng, wolfberry and red sage (Yu et al., 2019). Moreover, phytochemical databases of medicinal plants, such as IMPPAT (Mohanraj et al., 2018), NPACT (Mangal et al., 2013), NuBBEDB (Pilon et al., 2017), and NANPDB (Ntie-Kang et al., 2017), are also publicly available. Each phytochemical database provides classic information (*e.g.*, common name, taxonomy, location, medicinal parts, and application) and chemical/structure information of medicinal properties. Together, these databases allow researchers to conduct deep data mining to explore gene annotation and expression profiles of pharmacological properties, study the roles of miRNA in regulating biosynthesis and accumulating secondary metabolites, and acquire DNA barcode data to facilitate identification of medicinal materials.

Although the rapid advancement of sequencing and spectrometry technologies and the availability of computational tools has stimulated medicinal plant research worldwide, there is still not a comprehensive review from a multi-omics aspect summarizing the current research progress in discovering genes, proteins, and key/secondary metabolites that involve in biosynthesis pathways in medicinal plants. Therefore, our goal for this review is to provide an update on application of omics technologies in medicinal plant research to explore compounds in biosynthesis pathways for natural drug discovery. We briefly discuss medicinal plant genomes at chromosome, chloroplast, and mitochondrial level, and the needs of conducting comparative genomics, epigenomics and pan-genomics research as well as constructing genetic mapping for the development and selection of medicinal plants with high bioactive compounds. Next, we emphasize the importance of applying other omics approaches (*i.e.*, transcriptome, proteome and metabolome) to identify key molecular products related to biosynthesis of active ingredients of medicinal plants, with a focus on phenolic acid, flavonoids, and alkaloid biosynthesis pathways. Under the current challenges in medicinal plant breeding, we hope that our review could inspire more efforts to integrate high-throughput omics technologies in medicinal plant research to facilitate natural drug discovery.

Updates in medicinal plant genomes

Medicinal plant genome assembly is challenging due to their large genome sizes and complicated polyploid chromosomes; however, long-read sequencing technologies have increased a significant number of assembled medicinal plant genomes at the chromosomal level (Supplementary Table 1; as the date of July 30th, 2022). According to Cheng et al. (2021), 161 reference genomes representing 126 medicinal plant species (in red in Supplementary Table 1) were published as the date of June 4, 2021. In this review, we summarized additional 118 genomes from 78 medicinal plant species, and they are presented in Supplementary Table 1 (in black).

As shown in Figure 1, the chromosome number of medicinal plants varies widely, ranging from 8 to 80, and the genome sizes are from 157 Mb (crown flower, *Calotropis gigantea*) (Hoopes et al., 2018) to 70.18 Gb. The largest genome is *Paris polyphylla* (Li et al., 2020b) followed by the well-known vegetable *Allium sativum* (Garlic) (Liu et al., 2021a) with a genome size of 16.24 Gb. There are 40 plants with genome sizes between 2 to 16 Gb, including opium poppy (*Papaver somniferum*-2,720 Mb) (Pei et al., 2021b), Chinese mugwort (*Artemisia argyi*-8,030 Mb) (Miao et al., 2022), *Cymbidium sinense* (3,520 Mb) (Yang et al., 2021), cultivated tobacco (*Nicotiana tabacum*-4,600 Mb) (Sierro et al., 2014) and rehmannia (*Rehmannia glutinosa*-2,490 Mb) (Ma et al., 2021). Most of medicinal plants (*i.e.*, 232 out of 279 reported medicinal plant genomes) have a genome size smaller than 2 Gb. Of these,

TABLE 1 Databases on medicinal plant -omics and phytochemicals.

Database name	Database type	URL	Genus/Species name	Reference
Specific genus/species				
LjaFGD	Genome, Transcriptome	http://www.gzybioinformatics.cn/LjaFGD/index.php	*Lonicera japonica*	(Xiao et al., 2021)
RPGD	Genome, Transcriptome	http://bioinfor.kib.ac.cn/RPGD/	*Rhododendron*	(Liu et al., 2021b)
AprGPD	Genome, Transcriptome	http://apricotgpd.com	*Prunus*	(Chen et al., 2021)
croFGD	Genome, Transcriptome	http://bioinformatics.cau.edu.cn/croFGD/	*Catharanthus roseus*	(She et al., 2019)
MGH	Genome, Transcriptome	http://maca.eplant.org	*Lepidium meyenii*	(Chen et al., 2018)
SmGDB	Genome, Transcriptome	http://8.140.162.85/	*Salvia miltiorrhiza*	(Zhou et al., 2022)
Ginseng Genome Database	Genome, Transcriptome, Proteome, Metabolome	http://ginsengdb.snu.ac.kr/	*Panax ginseng*	(Jayakodi et al., 2018)
Comprehensive species				
HMOD	Genome, Transcriptome, Metabolome	http://herbalplant.ynau.edu.cn/	Up to 138 species	(Wang et al., 2018)
BPGD	Genome, Transcriptome	http://www.bpgenome.com	Up to 34 species	(Zhou et al., 2021a)
GPGD	Genome, Transcriptome	http://www.gpgenome.com	Up to 350 species	(Liao et al., 2021)
MPOD	Genome, Transcriptome	http://medicinalplants.ynau.edu.cn/	Up to 187 species	(He et al., 2022b)
TCMPG	Genome	http://cbcb.cdutcm.edu.cn/TCMPG/	195 species	(Meng et al., 2022)
1 K-MPGD	Genome	http://www.herbgenome.com/	108 species	(Su et al., 2022)
MepmiRDB	miRNA	http://mepmirdb.cn/mepmirdb/index.html	29 species	(Yu et al., 2019)
IMPPAT	Phytochemicals	https://cb.imsc.res.in/imppat	–	(Mohanraj et al., 2018)
NPACT	Phytochemicals	http://crdd.osdd.net/raghava/npact/	–	(Mangal et al., 2013)
AromaDb	Phytochemicals	http://bioinfo.cimap.res.in/aromadb/	–	(Kumar et al., 2018)
NANPDB	Phytochemicals	http://african-compounds.org/anpdb/	–	(Ntie-Kang et al., 2017)
NuBBEDB	Phytochemicals	https://nubbe.iq.unesp.br/portal/nubbedb.html	–	(Pilon et al., 2017)
ETM-DB	Phytochemicals	http://biosoft.kaist.ac.kr/etm/home.php/	–	(Bultum et al., 2019)

-: not applicable.

liverworts (*Marchantia polymorpha*-226 Mb), green chiretta (*Andrographis paniculate*-284 Mb), and Australian dodder (*Cuscuta australis*-265 Mb) are common medicinal plants with small genomes.

Plant organelles such as plastid and mitochondrion retain their own genome architectures (Wu et al., 2020). These organelles also participate in biosynthesizing fatty acids, amino acids, hormones, vitamins, and nucleotides (Dobrogojski et al., 2020), and their slow evolution due to low recombination rates are ideal for studying phylogenetic relationships and important traits (*e.g.*, phytochemicals) among medicinal plant species (Drouin et al., 2008; Smith, 2015). Currently, there are more than 2,000 chloroplast and mitochondrial genomes of medicinal plants, with only a few complete mitochondrial genomes (Wu et al., 2020; Liao et al., 2021). Following the availability of mitochondrial and chloroplast genome sequences, many medicinal plant molecular markers have been developed to aid in breeding and examining the authenticity and quality of herbs. However, there is a significant gap of using these genomic resources to illustrate functional genes in medicinal plant phytochemical production.

Advancing comparative genomics, genetic mapping, pan-genome, and epigenome research to aid in production of secondary metabolites for pharmaceutical use

The number of comparative genomic studies to identify biosynthesis-related genes of medicinal plants has increased in the past few years. For example, genes that encode oligopeptide transporters (OPT) were shown to mediate the transportation of many bioactive chemical compounds in *Panax ginseng* and other flowering plants through conducting genomic comparisons (Su et al.,

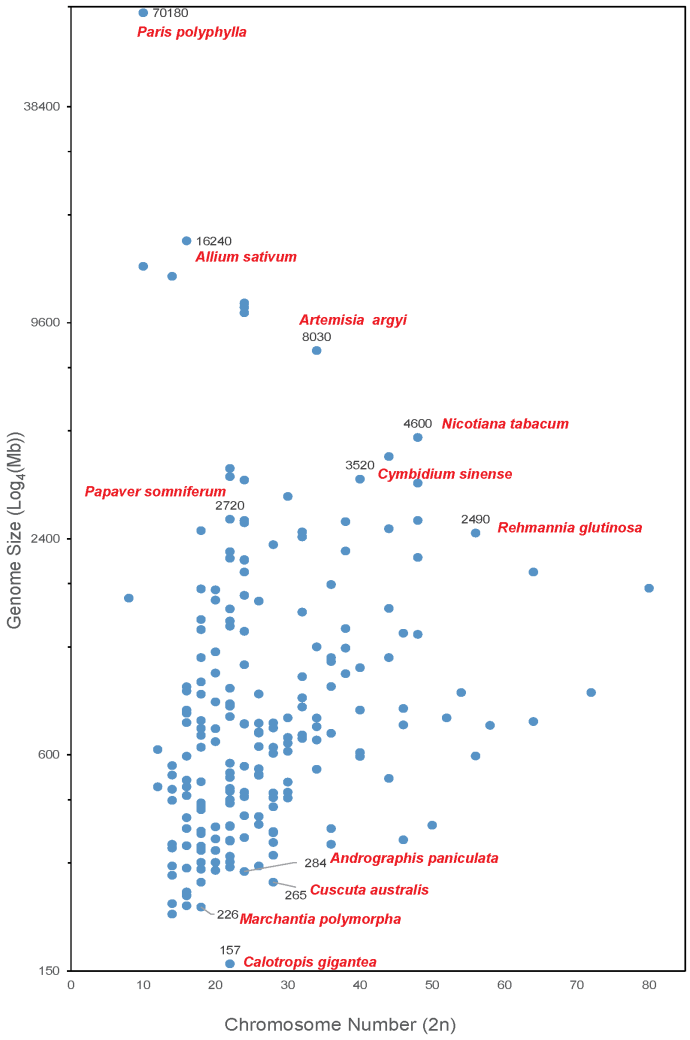


FIGURE 1
Medicinal plants with complete genome sequencing. The X-axis shows the chromosome number of a medicinal plant genome. The Y-axis represents the genome size (Log_4(Mb)) of each plant.

2019). Comparing genomes of *Senna tora* and other 14 plants in Fabaceae revealed that the *Senna tora* genome is enriched in biosynthesis-related genes of phenylpropane, isoflavones, and terpenes (Kang et al., 2020). Comparative genomic analysis of multiple *Cannabis sativa*, including the female strains PK with high delta-9-tetrahydrocannabinol (Δ9-THC, or simply THC), the CBDRx (cs10) strains with high cannabidiol (CBD) and female strains Cannbio-2 with balanced CBD : THC cannabinoid ratios suggested extensive copy number variation in cannabinoid synthesis (McKernan et al., 2020). These comparative genomic studies have set up good examples for researchers to explore subtle genetic variations on the plant-derived metabolite contents in other medicinal plant species, such as Chinese pistache (*Pistacia chinensis*), Ginger (*Zingiber officinale*), green chiretta (*Andrographis paniculata*) and passion fruit (*Passiflora edulia*) (Supplementary Table 1).

Constructing genetic linkage map provides the basis for gene mapping and cloning as well as studying the structure and function of a genome. However, the long growth cycle, short cultivation history, complex genetic background, and highly heterozygous genes of most medicinal plants make it difficult to establish their genetic maps. Traditional molecular markers like SSR, AFLP, RFLP, RAPD, and EST-SSR were used to construct genetic maps of medicinal plants such as *Trifolium pratense* (Isobe et al., 2003), artichoke (*Cynara scolymus*) (Lanteri et al., 2006), bladder campion (*Silene Vulgaris*) (Bratteler et al., 2006) and passion fruit (*Passiflora Edulis*) (Carneiro et al., 2002). Recently, genome-wide molecular markers are rapidly developed, leading to the development of a high-saturate and versatile molecular linkage maps of medicinal plants. With the increasing quantity and quality of medicinal plant genomes, we foresee that constructing genetic maps of phenotypes related to bioactive compounds in medicinal plants will bloom.

Current pan-genome and epigenome research of medicinal plants mainly focuses on determining phenotypic and environmental adaptability among subspecies. As an example, agronomic trait-associated SNPs identified by pan-genome analysis indicated that location and year significantly affected yield-related phenotypes in pigeon pea (*Cajanus cajan*), a plant species with rich vitamin B, protein, ascorbic acid and carotene (Zhao et al., 2020). Developmental methylome of the medicinal plant Cape periwinkle (*Catharanthus*

roseus) determined that cellular and physiological functions are likely to be impacted by tissue-specific covariations between context-dependent DNA methylation (Dugé de Bernonville et al., 2020). The cold environment affected ginsenosides accumulation in perennial American ginseng (*Panax quinquefolium*), and a cyclically reversible dynamism between methylation and demethylation of DNA was also reported in response to temperature seasonality (Hao et al., 2020). Further investigations of medicinal plant pan-genomes and epigenomes addressing bioactive compound metabolism are needed to improve the production of secondary metabolites for pharmaceutical applications.

Functional gene mining of medicinal plant transcriptome to characterize secondary metabolite biosynthesis pathways

Phenolic compounds, terpenoids and alkaloids are important secondary metabolites in medicinal plants (Mrudulakumari Vasudevan and Lee, 2020), but the pharmacological activity of medicinal plants is often estimated by the content of phenolic compounds. Phenolic compounds consist of (1) monophenols (*e.g.*, benzoic acid derivatives (hydroxybenzoic acids)) and cinnamic acid derivatives (*e.g.*, hydroxycinnamic acids); (2) oligophenols (*e.g.*, flavonoids, stilbenes and coumarins) and (3) polyphenols (*e.g.*, lignin and tannins) (Marchiosi et al., 2020). Due to the broad benefits of phenolic compounds to human health, current research aim at improving the production of phenolic compounds on a large scale.

Using RNA-seq technology can identify key functional genes or transcription factors involved in phenolic compound biosynthetic pathways of medicinal plants. With an improved understanding of phenolic compound biosynthesis pathways, pharmaceutical industries can achieve the goal of obtaining bioactive compounds more efficiently. Here, we summarize the synthetic metabolic pathways of three representative phenolic compounds: flavonoids, phenolic acids, and lignin (Figure 2) by showing key genes and metabolites as well as medicinal plants that have been reported to synthesize specific metabolites. As shown in Figure 2, biosynthesizing phenolic compounds starts from the phenylpropanoid pathway, where phenylalanine (or tyrosine) is subsequently converted to the intermediates cinnamic acid, coumaric acid and p-coumaroyl-CoA by phenylalanine ammonia lyase (PAL; or tyrosine ammonia lyase, TAL), cinnamate 4-hydroxylase (C4H), and 4-coumaroyl coenzyme A (CoA) ligase (4CL). These intermediates serve as a starting compound(s) to further produce other metabolites in each synthetic metabolic pathway.

Phenolic acid biosynthetic pathway

Phenolic acids, as representative compounds of monophenols, have gained a lot of attentions because of their antioxidant, antimicrobial, anti-inflammatory activities. In commercial Danshen (*Salvia*) decoctions, phenolic acids, especially salvianolic acid, are the major marker component used for quality assessment according to the official Chinese Pharmacopoeia. A proposed biosynthetic pathway of phenolic acids in *Salvia apiana* has been schemed based on full-length transcriptomic and metabolomic profiling (Shi et al., 2021; Hu et al., 2022). In addition, functional genes and transcription factors as regulators of phenolic acids biosynthesis in *Salvia miltiorrhiza* have been reported (Yu et al., 2018; Deng et al., 2020; Zhou et al., 2021c). For example, rosmarinic acid synthase (RAS) and a cytochrome P450-dependent monooxygenase (CYP98A), precursors responsible for rosmarinic acid biosynthesis, were mostly highly expressed in roots. MYB transcription factors have a positive effect on methyl jasmonate (MeJA)-induced phenolic acid biosynthesis in *S. miltiorrhiza*. Overexpressing *SmMYB2* in hairy roots significantly increased the levels of salvianolic acids by binging the MYB-binding motifs of CYP98A14 and upregulating CYP98A14 expression.

Lignin biosynthetic pathway

Medicinal plant lignin are reported as antioxidants (Karmanov et al., 2021; Lu et al., 2022). Lignin and their degradation products (*e.g.*, phenylpropanoids) have shown prominent anti-UVC (ultraviolet C) activities (Sakagami et al., 2022). In the terminal of lignin biosynthesis, peroxidase (POD) catalyzed individual monolignols to polymerize the complex lignin (Figure 2). The comparison between the CM (consecutive monoculture) and NG (normal growth for 1 year) root transcriptomes of *Achyranthes bidentata* revealed that genes encoding POD were mostly activated in the CM condition, suggesting the contribution of POD to lignin biosynthesis pathway (Yang et al., 2018). Comparing to the number of research in the area of phenolic acids in medicinal plants, research that focus on exploring lignin synthetic pathway from medicinal plants are limited.

Flavonoids biosynthetic pathway

Flavonoids, such as flavone, flavonol, flavanones, anthocyanin, chalcone, aurone, isoflavone and proanthocyanidin, are the most widely distributed phenolic compounds in plants (Liu et al., 2021c). Starting from coumaroyl-CoA, biosynthesis pathways of flavonoids are further divided into different branches responsible for the accumulation of various flavonoids, under the regulation of different enzymes and genes (Figure 2). For example, the expression of chalcone synthase (CHS), the key and first rate-limiting enzyme in the flavonoid biosynthetic pathway, decides the total content of flavonoids in medicinal plants (Yang et al., 2019; Ohta et al., 2021). In addition, two types of flavone synthase (FNS), FNSI (soluble 2-oxoglutarate-dependent dioxygenases) and FNSII (NADPH-dependent cytochrome P450 monoxygenases), play important roles in the accumulation of luteolin and apigenin in different medicinal plants (Zhao et al., 2016; Jiang et al., 2019; Li et al., 2020a; Tian et al., 2022).

Plant-derived alkaloids have been used as medicine for a long time. Alkaloids are a large and complex group of cyclic compounds that contain nitrogen. Medicinal plants from Ranunculales, such as opium poppy (*Papaver somniferum*) and *Rhizoma Coptidis*, are often used as model systems for studying benzylisoquinoline alkaloids

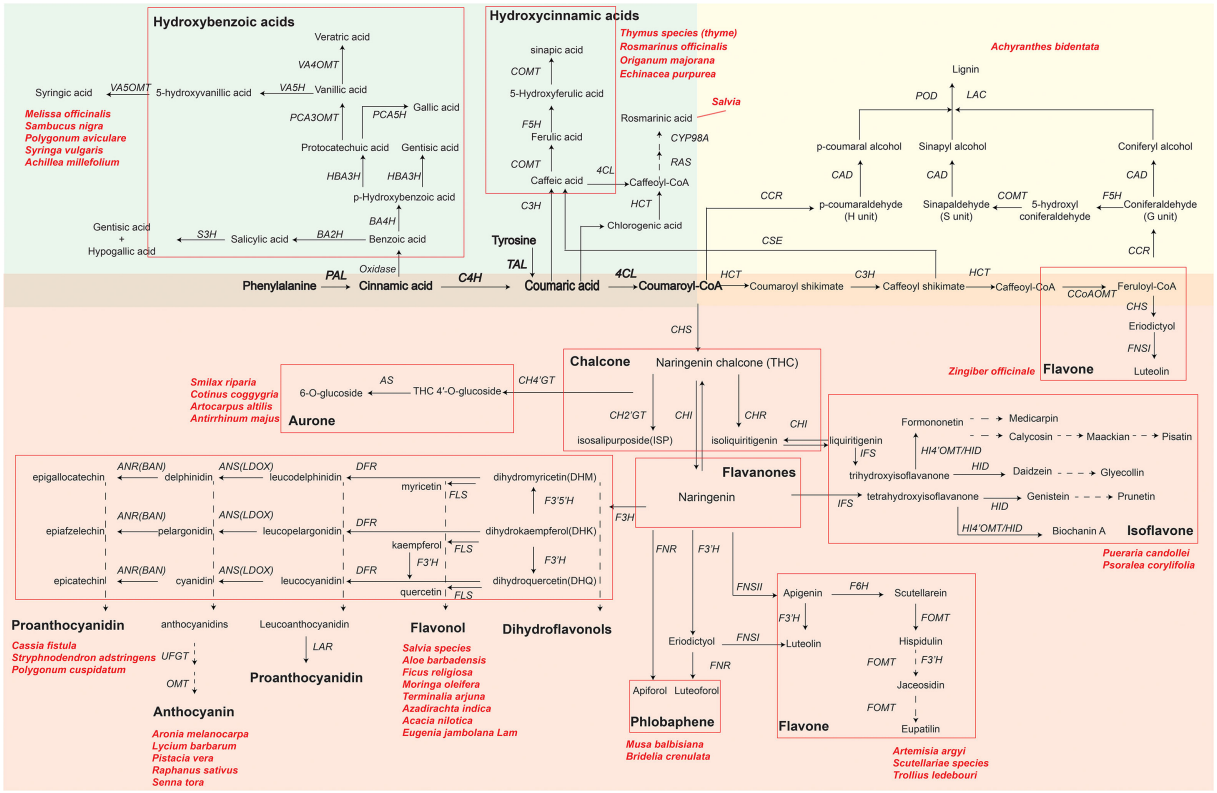


FIGURE 2
Key genes and metabolites of three phenolics compounds biosynthesis pathways in plants. The green, yellow and red background areas represent the biosynthetic pathways of phenolic acids, lignin and flavonoids, respectively. Bolded metabolites and genes represent the beginning of henylpropanoid pathway shared by phenolics compounds biosynthesis. Different flavonoids are indicated by black bold font and grouped by red rectangular boxes in each pathway. Key genes in biosynthesis pathways are in italics: *PAL*, phenylalanine ammonialyase; *C4H*, cinnamic acid 4-hydroxylase; *4CL*, 4-coumarate-CoA ligase; *HCT*, hydroxycinnamoyl CoA shikimate hydroxycinnamoyl transferase; *RAS*, rosmarinic acid synthase; *CYP98A*, cytochrome P450-dependent monooxygenase; *C3H*, *p*-coumarate 3-hydroxylase; *CHS*, chalcone synthase; *CH4'GT*, chalcone 4'-O-glucosyltransferase; *CH2'GT*, chalcone 2'-glucosyltransferase; *AS*, aureusidin synthase; *CHR*, chalcone reductase; *CHI*, chalcone flavanone isomerase; *IFS*, isoflavone synthase; *F3'H*, favonoid 3' hydroxylase; *FNS* (*FNSI* and *FNSII*), flavone synthase; *F6H*, flavanone-6-hydroxylase; *FOMT*, flavonoid O-methyltransferase; *HI4'OMT*, hydroxyisoflavone 4'-O-methyltransferase; *HID*, hydroxyisoflavanone dehydratase; *DFR*, dihydroflavonol 4-reductase; *F3H*, favanone 3-hydroxylase; *F3'5'H*, flavonoid-3',5'-hydroxylase; *FLS*, flavonol synthase; *ANS*, anthocyanidin synthase (*LDOX*, leucoanthocyanidin dioxygenase); *ANR*, anthocyanidin reductase; *LAR*, leucoanthocyanidin reductase; *UFGT*, UDP glucose:flavonoid 3-O-glycosyltranferase; *OMT*, O-methyltransferase; *CCoAOMT*, caffeoyl-CoA O-methyltransferase; *F5H*, ferulate 5-hydroxylase; *CSE*, caffeoyl shikimate esterase; *COMT*, caffeic acid O-methyltransferase; *CCR*, cinnamoyl-CoA reductase; *CAD*, cinnamyl alcohol dehydrogenase; *LAC*, laccase; *POD*, peroxidase; *TAL*, tyrosine ammonia lyase; *BA2H*, benzoic acid 2-hydroxylase; *S3H*, salicylic acid 3-hydroxylase; *BA4H*, benzoic acid 4-hydroxylase; *HBA3H*, p-hydroxybenzoic acid 3-hydroxylase; *PCA5H*, protocatechuic acid 5-hydroxylase; *PCA3OMT*, protocatechuic acid 3-O-methyltransferase; *VA4OMT*, vanillic acid 4-O-methyltransferase; *VA5H*, vanillic acid 5-hydroxylase; *VA5OMT*, vanillic acid 5-O-methyltransferase. Naringenin chalcone and naringenin in a larger font are key primarily intermediate metabolite in flavonoid biosynthesis. Dashed arrows indicate that some unknown enzymes are involved in these processes. Scientific names are representative medicinal plants, which are exploited by various pharmaceutical companies to produce phenolics compounds.

(BIAs) biosynthesis. Gene mining by functional transcriptomics can effectively promote the discovery of alkaloids biosynthetic pathway and facilitate their characterization (Li et al., 2020c; Hao et al., 2021; Xu et al., 2022). Multicopy genes in alkaloids pathway, such as BBE and 4OMT, were divergently expressed between copies in *P. somniferum*, being differentially expressed between either tissues or different stages, suggesting possible differences in their regulatory function (Pei et al., 2021). Wang et al. reported that the total alkaloids contents in Protocorm-like bodies (PLBs) was almost twice as high as that of plant organs of *Dendrobium officinale*, a species with high alkaloid content. Using RNA-seq technology, the authors identified putative genes that encode enzymes in the alkaloids biosynthetic pathway in PLBs and leaves of *D. officinale* (Wang et al., 2021).

The effects of plant growth duration on bioactive substance accumulation in the same medicinal plant species were explored (Lin et al., 2020; Liu et al., 2022). For example, comparing transcriptome profiles of one to four years old *D. officinale* stems revealed that key genes, such as *CHS* and *FLS* that are involved in flavonol synthesis, were highly expressed in the biennial samples, suggesting that the optimal harvesting period of *D. officinale* is 2-3 years (Yuan et al., 2022). Moreover, multiple DEGs involved in liquiritin biosynthesis (e.g., UDP-glucosyltransferase (UGTs)), displayed distinct expression patterns in *Glycyrrhiza uralensis* farmed between 1 year and 3 years (Zhong et al., 2022).

It is of a great significance to characterize genes related in secondary metabolite biosynthesis between closely related species or varieties in medicinal plants (Koo et al., 2022; Liu et al., 2022; Shafi et al., 2022), or between different tissues in same species (He et al., 2022a; Naika et al., 2022). The comparative transcriptome analysis of the marijuana strain Purple Kush and the 'Finola' cultivar of *Cannabis*

sativa demonstrated that the Δ9-tetrahydrocannabinolic acid synthase was detected in Purple Kush but replaced by cannabidiolic acid synthase in the 'Finola' cultivar, explaining the significant chemical difference between marijuana-derived Δ9-tetrahydrocannabinol (THC) and hemp-derived cannabidiol (CBD) (van Bakel et al., 2011). Moreover, transcriptional changes during tanshinones accumulation stage among different varieties of *Salvia miltiorrhiza* expanded our vision on intraspecific variation and gene regulation mechanism of secondary metabolite biosynthesis pathways (Zhou et al., 2021b). Interestingly, platycodin D content of calli was higher than that of leaves in *Platycodon grandiflorus*, while platycoside E content of calli was lower than that of leaves. Comparative analysis of the transcriptome data identified 54 and 23 specifically expressed transcription factors in leaf and calli respectively, providing valuable resources for researchers to study the conversion between platycoside E and platycodin D (Su et al., 2021). Li et al. used RNA-seq to study the molecular mechanisms of different tissues of Chinese sage (*Salvia miltiorrhiza*) in response to moderate drought stress. GO enrichment analysis showed that several transcription factors, such as AP2/ERF, bHLH and WRKY that regulate abiotic responses of *S. miltiorrhiza*, were significantly enriched in roots and leaves. Under moderate drought stress, genes encoding key enzymes in the biosynthesis of phenylpropane and terpenoids were also upregulated (Li et al., 2020d). Together, these results provide us a solid foundation to further investigate biosynthesis mechanisms of medicinal components in medicinal plants.

Proteomic dissection of medicinal plants in different manners for potential drug development

Over the past decades, considerable amounts of medicinal plants have been proved to exhibit potent effects on different human diseases based on their specific therapeutic compounds. Figure 3 summarizes the main research foci of medicinal plant proteomics that lead to the identification of proteinaceous compounds involved in active bioactive compound conversion in recent years. The contents of these four foci often complement each other and are not studied independently. For example, global proteome and phosphoproteome profilings of *Dendrobium huoshanense* under greenhouse planting (GP) and the cultivation modes of stone planting under the forest (SPUF) revealed that SPUF was more conducive to the accumulation of polysaccharides and alkaloids, and that there was a possible correlation between phosphorylation levels of different enzyme sites and the polysaccharide/alkaloid content (Wu et al., 2022b).

Most of proteomic studies in medicinal plants focus on investigating protein abundance changes under different environmental conditions (Figure 3, blue part). For instance, illumination (Zhang et al., 2021), temperature, drought (Xu et al., 2021; Zhang et al., 2022a), flood, saline (Fortini et al., 2022), exogenous hormone (Yang et al., 2022), air composition and planting style are the main factors affecting the content of enzymes or regulators involved in metabolites biosynthesis in the medicinal plants. Unlike abiotic conditions, studies that investigate the effects of biotic stresses on the proteins influencing accumulation of active compounds in traditional medicinal plants are relatively less. Through leaf and rhizome proteomic analysis of kutki (*Picrorhiza kurroa*), the abundance of proteins associated with carbon metabolism in CO_2 enhancement were upregulated (*i.e.*, glucose and fructose) in a tissue-specific manner (Kumar et al., 2020). Two-dimensional electrophoretic and MALDI-TOF/TOF mass spectrometry analyses identified 20 differentially expressed proteins of hoary cress (*Lepidium draba*) related to photosynthesis, energy metabolism and other functions to water stress (6% PEG) (Jamshidi Goharrizi et al., 2020). Many precious medicinal plants grow in high latitudes are cold tolerant. Exploring heat responses at the protein level between heat-tolerant and heat-sensitive strains of *Clematis florida* provided evidence on its adaptive mechanism of thermotolerance (Jiang et al., 2020). *Sophora alopecuroides* is a famous saline-alkali tolerant and drought-tolerant medicinal plant. Tandem mass tag (TMT) based proteomic profiling of *S.*

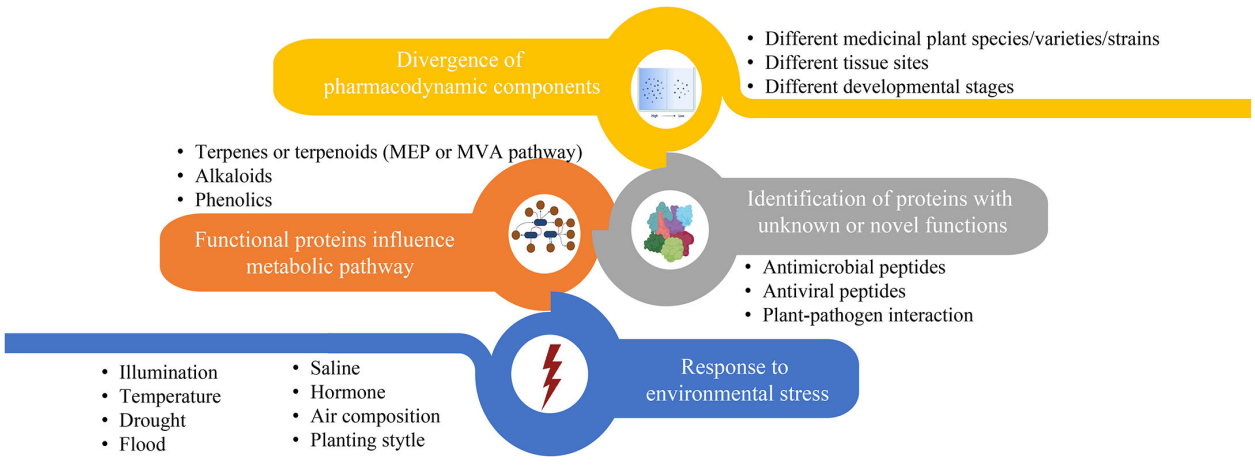


FIGURE 3
The main contents of proteome research related to biosynthesis or accumulation of metabolites in medicinal plant.

alopecuroides leaves confirmed that salt stress altered several transporter proteins related to the secondary metabolite's biosynthesis pathway in *S. alopecuroides* leaves (Ma et al., 2022).

Other proteomic studies identified novel proteins/peptides of medicinal plant origin with pharmacological interests (Figure 3, grey part) (Moyer et al., 2021b). For instance, seven novel peptides belonging to three antimicrobial peptide classes, including lipid transfer proteins, snakins and a defensin, were identified by MS-based peptidomics analysis from the aerial tissues of edible amaranth (*Amaranthus tricolor*) plants (Moyer et al., 2021a). Two novel bioactive peptides isolated from the Asian medicinal plant *Acacia catechu* are recommended for further investigation as antiviral peptides with their potent inhibition activities against dengue viruses (Panya et al., 2019).

The synthesis regulatory pathway of secondary metabolites involves many functional enzymes which provide guidance for drug discovery based on protein expression (Figure 3, orange part). In addition to phenolic compounds and alkaloids in medicinal plants and their corresponding biosynthesis pathways that we previously discussed (Figure 2), terpenoids are of economic interests for drug

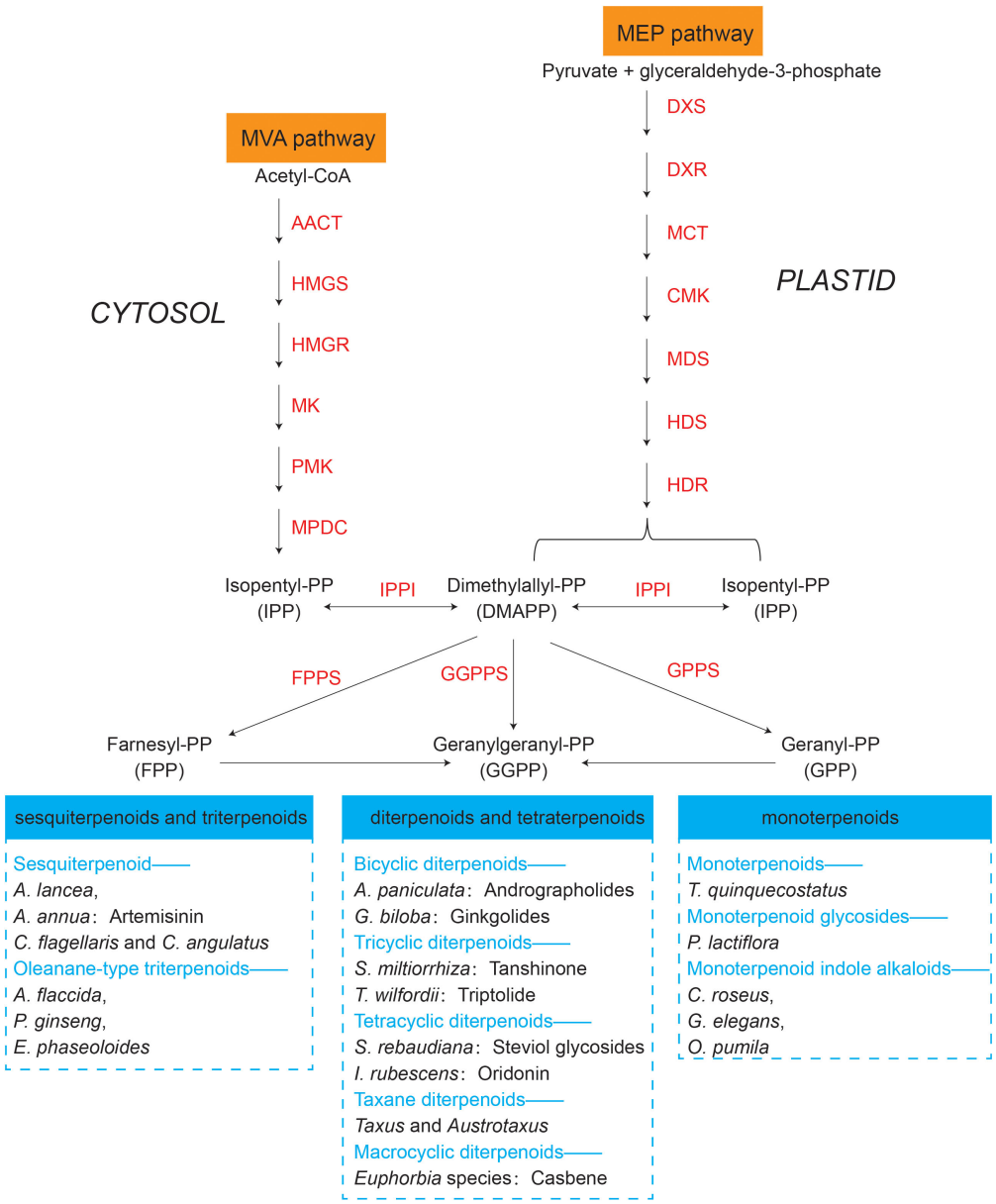


FIGURE 4
Terpenoids synthesized by representative medicinal plants through mevalonate (MVA) and methylerythritol phosphate (MEP) pathways. The key proteins involved in the two pathways are in orange. AACT, acetoacetyl-CoA thiolase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; MPDC, mevalonate diphosphate decarboxylase; DXS, 1-deoxy-d-xylulose 5-phosphate synthase; DXR, 1-deoxy-d-xylulose 5-phosphate reductoisomerase; MCT, 2C-methyl-d-erythritol 4-phosphate cytidyl transferase; CMK, 4-diphosphocytidyl-2C-methyl-d-erythritol kinase; MDS, 2C-methyl-d-erythritol 2,4-cyclodiphosphate synthase; HDS, 1-hydroxy-2-methyl-2-€-butenyl 4-diphosphate synthase; HDR, 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase; IPPI, Isopentenyldiphosphate-isomerase; FPPS, farnesyl diphosphate (FPP) synthase; GGPPS, geranylgeranyl diphosphate (GGPP) synthase; GPPS, geranyl diphosphate (GPP) synthase. Representative medicinal plants containing various terpenoids are listed in the blue dashed box.

development. The terpenoid metabolites are derived from the common precursor isopentenyl diphosphate (IPP), which can be synthesized *via* two different pathways: the mevalonate (MVA) pathway in the cytoplasm and the methylerythritol phosphate (MEP) pathway in plastids. Terpenoids can be classified into monoterpenes, sesquiterpenes, meroterpenes, triterpenes, diterpenoids and other terpenoids in medicinal plants (Awouafack et al., 2013; Sandjo and Kuete, 2013b; Sandjo and Kuete, 2013a; Tchimene et al., 2013; Nazir et al., 2021) (Figure 4). Among the differentially expressed proteins between leaves and rhizomes of soft windflower (*Anemone flaccida*), most proteins involved in the metabolic pathway of triterpenoid saponins biosynthesis were upregulated in rhizomes (Zhan et al., 2016). Comparative proteome analysis of the leaves, roots, shoots and fruits of Korean Ginseng by label-free quantitative proteomics identified that 67 out of 1,179 differentially regulated proteins were associated with ginsenoside biosynthesis pathways, including MEP pathway, MVA pathway, UDP-glycosyltransferase and oxidoreductase (Van Nguyen et al., 2021).

Divergence of pharmacodynamic components in different species/varieties/strains (Qin et al., 2014; Song et al., 2021), tissues (Guo et al., 2022; Pan et al., 2022) and development stages (Huang et al., 2021; Zhang et al., 2022b) of medicinal plants implicates expression level of key proteins involved in biosynthesis and metabolism of medicinal compounds (Figure 3, yellow part). Based on quantitatively targeted subproteomic analysis of the high and low artemisinin content of sweet wormwood (*Artemisia annua*), the increase expression of DBR2 accounts for the high artemisinin content (Chen et al., 2020). The genus *Paris* includes a variety of genotypes with different medicinal contents. Proteomic changes in rhizomes between closely related species of *Paris polyphylla* enhanced our understandings on the molecular basis of different medicinal properties. Such as, the higher efficiency of sucrose utilization in the sugar metabolic pathway of *P. polyphylla* var. *chinensis* fare probably related to the elevated protein abundance (Liu et al., 2019). Comparing protein contents of latex at different developmental stages revealed that stress- and defense-related proteins in green fruit phase were of higher abundances, but upregulated proteins in flowering phase were related to transcription, protein folding, and active transport of molecules, providing new insights into the biology and medicinal use of greater celandine (*Chelidonium majus*) (Nawrot et al., 2017).

Metabolomics profiles of medicinal plants revealed that both primary and secondary metabolites have pharmacological potential

Metabolomics studies in medicinal plants aim to provide comprehensive examination of metabolite profiles and quality

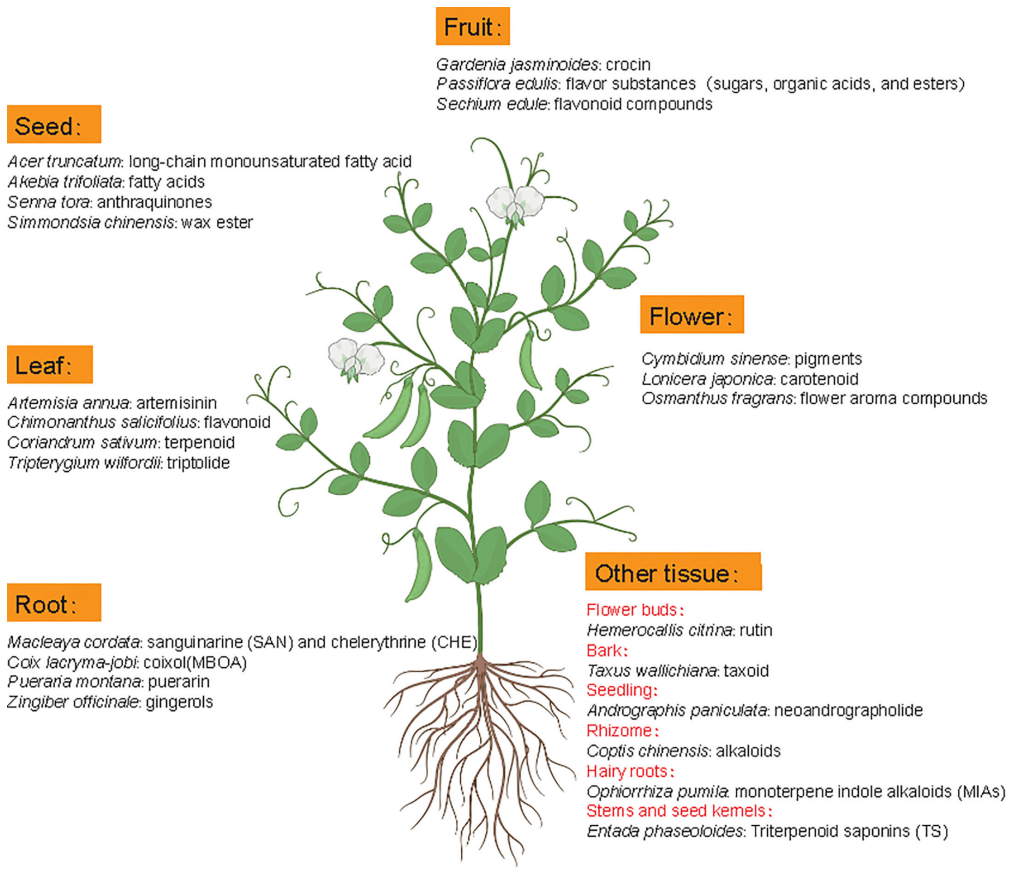


FIGURE 5
Metabolic components with main pharmacodynamic function from different tissues of representative medicinal plants. Latin names are the representative medicinal plants with specific tissue(s) as the main source of medicinal ingredients.

assessment of medicinal plants. Although the extraction method could influence the effective acquisition of metabolites, the difference of medicinal material quality is mainly attributed to metabolite diversity and composition among different tissues or species because roots, leaves, flowers, fruits, seeds, rhizomes, bark or whole plants of many natural plants can contain different active ingredients with nutritional or therapeutic function (Figure 5) (Rai et al., 2021). For example, artemisinin mainly accumulates in leaves of *Artemisia annua* but gingerols is mainly stored in *Zingiber officinale* roots. In some cases, active compounds accumulate in special tissues, like bark (*Axus wallichiana*: taxoid) and rhizome (*Coptis chinensis*: alkaloids).

To comprehensively characterize metabolites and low-molecular-weight molecules with therapeutic values, we often adapt two technology platforms: nuclear magnetic resonance (NMR) and/or gas/liquid chromatography-mass spectrometry (GC-MS or LC-MS/MS). For example, GC-MS and LC-MS/MS generated metabolic profiles of roots, stems, and leaves of *Panax notoginseng* from different geographical regions showed that the composition of saponins was similar between root and stem but was different than plant leaves. These results provide further evidence that different parts in addition to *P. notoginseng* roots, for example their stems, can be used in practice (Gao et al., 2022). Another investigation on metabolite profiles of twenty medicinal plants identified that two hydroxylated fatty acids (13S-Hydroxy-9Z,112,15Z-octadecatrienoic acid and 13-Hydroxy-9Z,11E-octadecadienoic acid) are potential targets for developing anti-viral drugs (More et al., 2022). Similarly, metabolome analysis of neuroactive plants (*e.g.*, *Hypericum perforatum* L. (St. John's wort), *Passiflora incarnate* L. (maypop), *Valeriana officinalis* L. (Valerian) and *Melissa officinalis* L. (Lemon balm)) revealed that primary metabolites in the tricarboxylic acid (TCA) cycle as well as secondary metabolites belong to flavonoids and terpenoids positively correlated with BDNF (*i.e.*, brain-derived neurotrophic factor; an important indicator of neurodegenerative diseases) expression level *in vitro* (Gonulalan et al., 2020).

Technological advances in mass spectrometry-based platforms have facilitated the separation and identification of multiple metabolomic profiles of medicinal plants. With the implementation of the Herbal Genome Project (Chen et al., 2011; Su et al., 2022) and the development of traditional Chinese medicine synthetic biology (Mortimer, 2019), the metabolomic research of medicinal plants will accelerate the discovery of novel bioactive compounds with pharmacological potential (Bhardwaj et al., 2022; Nisar et al., 2022).

Prospects

Medicinal plants have significant economic and social benefits due to their pharmacological activities. The increase of habitat destruction and human consumption of medicinal plants worldwide has increased their risk of extinction (Wang et al., 2020). However, unlike agricultural/horticultural crops, traditional breeding of medicinal plants is a more challenging task because factors such as medicinal parts, production of active ingredients with pharmacological potential, and growth cycle need to be considered. Multi-omics analysis coupled with bioinformatics and statistical analysis is a comprehensive approach to uncover the chemical diversity and the regulatory mechanisms and the formation of pharmacological properties of medicinal plants (Figure 6). Once genes, metabolites, peptides, or proteins involved in the biosynthetic pathways of active medicinal plant bio-compounds are elucidated, genome engineering or synthetic biology can be applied to produce them effectively and sustainably. Therefore, incorporating multi-omics technologies into medicinal plant research should be encouraged by organizations and research institutions across the

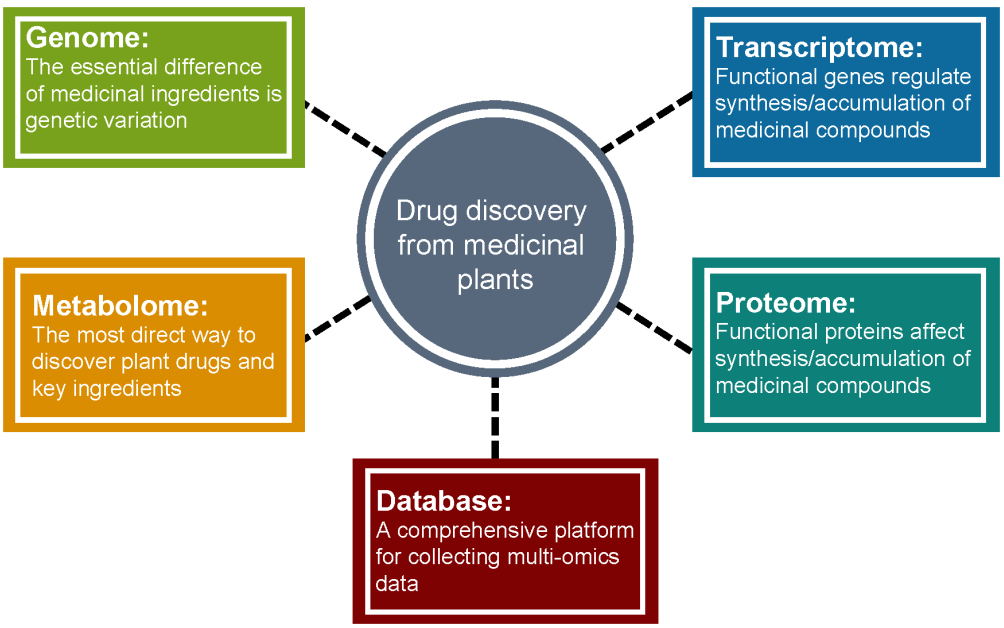


FIGURE 6
Medicinal plant multi-omics research to facilitate drug discovery.

globe to promote their cultivation in order to fulfill the needs of synthesizing bioactive components of medicinal plants for pharmaceutical applications.

Author contributions

WZ conceived the manuscript, and CL and JW supervised the study. WZ, YZ, MJ, and YL wrote the paper. WZ, MJ, and YL analyzed data, and generated figures and tables. All authors reviewed the manuscript and approved the submitted version.

Funding

The preparation of this article was supported by the National Natural Science Foundation of China (31960064); the Special Fund for Introducing Scientific and Technological Talents of Guangdong Academy of Agricultural Sciences (Grant # R2021YJ-YB2008). The funder was not involved in conceptualizing this article, collecting and interpreting the information, writing or submitting the article.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1073848/full#supplementary-material

References

An, X., Luo, X. H., Liu, T. T., Li, W. L., and Zou, L. N. (2022). Development and application of fruit color-related expressed sequence tag-simple sequence repeat markers in *Abelmoschus esculentus* on the basis of transcriptome sequencing. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.907895

Awouafack, M. D., Tane, P., Kuete, V., and Eloff, J. N. (2013). "2 - sesquiterpenes from the medicinal plants of Africa," in *Medicinal plant research in Africa*. Ed. V. Kuete. (Oxford: Elsevier), 33–103.

Bhardwaj, K., Sharma, R., Cruz-Martins, N., Valko, M., Upadhyay, N. K., Kuča, K., et al. (2022). Studies of phytochemicals, antioxidant, and antibacterial activities of *Pinus gerardiana* and *Pinus roxburghii* seed extracts. *BioMed. Res. Int.* 2022, 5938610. doi: 10.1155/2022/5938610

Bratteler, M., Lexer, C., and Widmer, A. (2006). A genetic linkage map of *Silene vulgaris* based on AFLP markers. *Genome* 49, 320–327. doi: 10.1139/g05-114

Bultum, L. E., Woyessa, A. M., and Lee, D. (2019). ETM-DB: integrated Ethiopian traditional herbal medicine and phytochemicals database. *BMC complementary Altern. Med.* 19, 212. doi: 10.1186/s12906-019-2634-1

Carneiro, M. S., Camargo, L. E., Coelho, A. S., Vencovsky, R., Rui, P. L., Stenzel, N. M., et al. (2002). RAPD-based genetic linkage maps of yellow passion fruit (*Passiflora edulis* sims. f. flavicarpa deg.). *Genome* 45, 670–678. doi: 10.1139/g02-035

Cheng, Q. Q., Ouyang, Y., Tang, Z. Y., Lao, C. C., Zhang, Y. Y., Cheng, C. S., et al. (2021). Review on the development and applications of medicinal plant genomes. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.791219

Chen, C., Liu, H., Gou, N., Huang, M., Xu, W., Zhu, X., et al. (2021). AprGPD: the apricot genomic and phenotypic database. *Plant Methods* 17, 98. doi: 10.1186/s13007-021-00797-4

Chen, S. L., Xiang, L., Guo, X., and Li, Q. S. (2011). An introduction to the medicinal plant genome project. *Front. Med.* 5, 178–184. doi: 10.1007/s11684-011-0131-0

Chen, M. H., Yan, T. X., Ji, L. Y., Dong, Y., Sidoli, S., Yuan, Z. F., et al. (2020). Comprehensive map of the *Artemisia annua* proteome and quantification of differential protein expression in chemotypes producing high versus low content of *Artemisinin*. *PROTEOMICS* 20, 1900310. doi: 10.1002/pmic.201900310

Chen, J. H., Zhang, J. W., Lin, M. G., Dong, W., Guo, X. Y., Dong, Y., et al. (2018). MGH: a genome hub for the medicinal plant maca (*Lepidium meyenii*). *Database* 2018, bay113. doi: 10.1093/database/bay113

Deng, C. P., Wang, Y., Huang, F. F., Lu, S. J., Zhao, L. M., Ma, X. Y., et al. (2020). SmMYB2 promotes salvianolic acid biosynthesis in the medicinal herb *Salvia miltiorrhiza*. *J. Integr. Plant Biol.* 62, 1688–1702. doi: 10.1111/jipb.12943

Dobrogojski, J., Adamiec, M., and Luciński, R. (2020). The chloroplast genome: a review. *Acta Physiologiae Plantarum* 42, 98. doi: 10.1007/s11738-020-03089-x

Drouin, G., Daoud, H., and Xia, J. N. (2008). Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol.* 49, 827–831. doi: 10.1016/j.ympev.2008.09.009

Dugé de Bernonville, T., Maury, S., Delaunay, A., Daviaud, C., Chaparro, C., Tost, J., et al. (2020). Developmental methylome of the medicinal plant *Catharanthus roseus* unravels the tissue-specific control of the monoterpene indole alkaloid pathway by DNA methylation. *Int. J. Mol. Sci.* 27 (19), 6685. doi: 10.3390/ijms21176028

Fan, G. Y., Liu, X. C., Sun, S., Shi, C. C., Du, X., Han, K., et al. (2020). The chromosome level genome and genome-wide association study for the agronomic traits of *Panax notoginseng*. *iScience* 23, 101538. doi: 10.1016/j.isci.2020.101538

Fortini, E. A., Batista, D. S., Felipe, S. H. S., Silva, T. D., Correia, L. N. F., Farias, L. M., et al. (2022). Physiological, epigenetic, and proteomic responses in *Pfaffia glomerata* growth *in vitro* under salt stress and 5-azacytidine. *Protoplasma*. doi: 10.1007/s00709-022-01789-4

Gao, M. Y., Cao, X. N., Wei, S. J., Huang, X. H., Ouyang, H. Z., Chang, Y. X., et al. (2022). Quantitative comparison and chemical profile of different botanical parts of *Panax notoginseng* from different regions. *Front. Nutr.* 9. doi: 10.3389/fnut.2022.841541

Gonulalan, E. M., Nemutlu, E., Bayazeid, O., Koçak, E., Yalçın, F. N., and Demirezer, L. O. (2020). Metabolomics and proteomics profiles of some medicinal plants and correlation with BDNF activity. *Phytomedicine* 74, 152920. doi: 10.1016/j.phymed.2019.152920

Guo, J., Wu, Y. Q., Guo, F. Y., and Wang, G. B. (2022). Proteomic and metabolomic analyses reveal stage- and tissue- specific flavonoid accumulation in *Ginkgo biloba*. *LWT* 171, 114111. doi: 10.1016/j.lwt.2022.114111

Hao, D.-C., Li, P., Xiao, P.-G., and He, C.-N. (2021). Dissection of full-length transcriptome and metabolome of dichocarpum (Ranunculaceae): implications in evolution of specialized metabolism of ranunculales medicinal plants. *PeerJ* 9, e12428. doi: 10.7717/peerj.12428

Hao, M. Z., Zhou, Y. H., Zhou, J. H., Zhang, M., Yan, K. J., Jiang, S., et al. (2020). Cold-induced ginsenosides accumulation is associated with the alteration in DNA methylation and relative gene expression in perennial American ginseng (*Panax quinquefolius* l.) along with its plant growth and development process. *J. Ginseng Res.* 44, 747–755. doi: 10.1016/j.jgr.2019.06.006

He, B., Han, X., Liu, H. L., Bu, M. J., Cui, P., and Xu, L. A. (2022a). Deciphering alternative splicing patterns in multiple tissues of ginkgo biloba important secondary metabolites. *Ind. Crops Products* 181, 114812. doi: 10.1016/j.indcrop.2022.114812

He, S., Yang, L., Ye, S., Lin, Y., Li, X., Wang, Y., et al. (2022b). MPOD: applications of integrated multi-omics database for medicinal plants. *Plant Biotechnol. J.* 20, 797–799. doi: 10.1111/pbi.13769

Hoopes, G. M., Hamilton, J. P., Kim, J., Zhao, D., Wiegert-Rininger, K., Crisovan, E., et al. (2018). Genome assembly and annotation of the medicinal plant *Calotropis gigantea*, a producer of anticancer and antimalarial cardenolides. *G3 Genes|Genomes|Genetics* 8, 385–391. doi: 10.1534/g3.117.300331

Huang, H., Liang, J., Tan, Q., Ou, L. F., Li, X. L., Zhong, C. H., et al. (2021). Insights into triterpene synthesis and unsaturated fatty-acid accumulation provided by chromosomal-

level genome analysis of *Akebia trifoliata* subsp. *australis*. *Horticulture Res.* 8, 33. doi: 10.1038/s41438-020-00458-y

Hu, J. D., Wang, F. Y., Liang, F. Y., Wu, Z. D., Jiang, R., Li, J. X., et al. (2022). Identification of abietane-type diterpenoids and phenolic acids biosynthesis genes in *Salvia apiana* jepson through full-length transcriptomic and metabolomic profiling. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.919025

Hu, Y., Zhao, R., Xu, P., and Jiao, Y. (2018). The genome of opium poppy reveals evolutionary history of morphinan pathway. *Genomics Proteomics Bioinf.* 16, 460–462. doi: 10.1016/j.gpb.2018.09.002

Isobe, S., Klimenko, I., Ivashuta, S., Gau, M., and Kozlov, N. N. (2003). First RFLP linkage map of red clover (*Trifolium pratense* l.) based on cDNA probes and its transferability to other red clover germplasm. *TAG. Theor. Appl. Genet. Theoretische und angewandte Genetik* 108, 105–112. doi: 10.1007/s00122-003-1412-z

Jamshidi Goharrizi, K., Fatehi, F., Nazari, M., Salehi, F., and Maleki, M. (2020). Assessment of changes in the content of sulforaphane and expression levels of CYP79F1 and myrosinase genes and proteomic profile of *Lepidium draba* plant under water-deficit stress induced by polyethylene glycol. *Acta Physiologiae Plantarum* 42, 101. doi: 10.1007/s11738-020-03085-1

Jayakodi, M., Choi, B. S., Lee, S. C., Kim, N. H., Park, J. Y., Jang, W., et al. (2018). Ginseng genome database: an open-access platform for genomics of *Panax ginseng*. *BMC Plant Biol.* 18, 62. doi: 10.1186/s12870-018-1282-9

Jiang, C. H., Bi, Y. K., Mo, J. B., Zhang, R. Y., Qu, M. N., Feng, S. C., et al. (2020). Proteome and transcriptome reveal the involvement of heat shock proteins and antioxidant system in thermotolerance of *Clematis florida*. *Sci. Rep.* 10, 8883. doi: 10.1038/s41598-020-65699-2

Jiang, Y. F. Y., Ji, X. Y., Duan, L. X., Ye, P., Yang, J. F., Zhan, R. T., et al. (2019). Gene mining and identification of a flavone synthase II involved in flavones biosynthesis by transcriptomic analysis and targeted flavonoid profiling in chrysanthemum indicum l. *Ind. Crops Products* 134, 244–256. doi: 10.1016/j.indcrop.2019.04.009

Jiang, Z. Q., Tu, L. C., Yang, W. F., Zhang, Y. F., Hu, T. Y., Ma, B. W., et al. (2021). 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

Kang, S. H., Pandey, R. P., Lee, C. M., Sim, J. S., Jeong, J. T., Choi, B. S., et al. (2020). Genome-enabled discovery of anthraquinone biosynthesis in *Senna tora*. *Nat. Commun.* 11, 5875. doi: 10.1038/s41467-020-19681-1

Karmanov, A. P., Kanarsky, A. V., Kocheva, L. S., Semenov, E. I., and Belyy, V. A. (2021). *In vitro* study of adsorption efficiency of natural lignins towards aflatoxin B2. *Reactive Funct. Polymers* 167, 105033. doi: 10.1016/j.reactfunctpolym.2021.105033

Koo, H., Lee, Y. S., Nguyen, V. B., Giang, V. N. L., Koo, H. J., Park, H. S., et al. (2022). Comparative transcriptome and metabolome analyses of four panax species explore the dynamics of metabolite biosynthesis. *J. Ginseng Res.* 47(1), 44–53. doi: 10.1016/j.jgr.2022.07.001

Kumar, R., Joshi, R., Kumari, M., Thakur, R., Kumar, D., and Kumar, S. (2020). Elevated CO2 and temperature influence key proteins and metabolites associated with photosynthesis, antioxidant and carbon metabolism in *Picrorhiza kurroa*. *J. Proteomics* 219, 103755. doi: 10.1016/j.jprot.2020.103755

Kumar, Y., Prakash, O., Tripathi, H., Tandon, S., Gupta, M. M., Rahman, L. U., et al. (2018). AromaDb: a database of medicinal and aromatic plant's aroma molecules with phytochemistry and therapeutic potentials. *Front. Plant Sci.* 9, 1081. doi: 10.3389/fpls.2018.01081

Lanteri, S., Acquadro, A., Comino, C., Mauro, R., Mauromicale, G., and Portis, E. (2006). A first linkage map of globe artichoke (*Cynara cardunculus* var. scolymus l.) based on AFLP, s-SAP, m-AFLP and microsatellite markers. *TAG. Theor. Appl. Genet. Theoretische und angewandte Genetik* 112, 1532–1542. doi: 10.1007/s00122-006-0256-8

Lau, W., and Sattely, E. S. (2015). Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science* 349, 1224–1228. doi: 10.1126/science.aac7202

Liao, B. S., Hu, H. Y., Xiao, S. M., Zhou, G. R., Sun, W., Chu, Y., et al. (2021). Global pharmacopoeia genome database is an integrated and mineable genomic database for traditional medicines derived from eight international pharmacopoeias. *Sci. China Life Sci.* 65(4), 809–819. doi: 10.1007/s11427-021-1968-7

Li, H., Li, D., Yang, Z., Zeng, Q. W., Luo, Y. W., and He, N. J. (2020a). Flavones produced by mulberry flavone synthase type I constitute a defense line against the ultraviolet-b stress. *Plants* 9, 215. doi: 10.3390/plants9020215

Li, J., Lv, M. Q., Du, L., Yunga, A., Hao, S. J., Zhang, Y. L., et al. (2020b). An enormous *Paris polyphylla* genome sheds light on genome size evolution and polyphyllin biogenesis. *bioRxiv* 2020, 2006.2001.126920. doi: 10.1101/2020.06.01.126920

Lin, W. D., Li, Y. L., Lu, Q. W., Lu, H. F., and Li, J. M. (2020). Combined analysis of the metabolome and transcriptome identified candidate genes involved in phenolic acid biosynthesis in the leaves of *Cyclocarya paliurus*. *Int. J. Mol. Sci.* 21(4), 1337. doi: 10.3390/ijms21041337

Li, Q. S., Ramasamy, S., Singh, P., Hagel, J. M., Dunemann, S. M., Chen, X., et al. (2020c). Gene clustering and copy number variation in alkaloid metabolic pathways of opium poppy. *Nat. Commun.* 11, 1190. doi: 10.1038/s41467-020-15040-2

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

Liu, F., Meng, Y. Y., He, K., Song, F. J., Cheng, J. H., Wang, H. X., et al. (2019). Comparative analysis of proteomic and metabolomic profiles of different species of *Paris*. *J. Proteomics* 200, 11–27. doi: 10.1016/j.jprot.2019.02.003

Liu, H. L., Wang, X. B., Wang, G. B., Cui, P., Wu, S. G., Ai, C., et al. (2021a). The nearly complete genome of *Ginkgo biloba* illuminates gymnosperm evolution. *Nat. Plants* 7, 748–756. doi: 10.1038/s41477-021-00933-x

Liu, S. A., Zhang, H. Y., and Yuan, Y. D. (2022). A comparison of the flavonoid biosynthesis mechanisms of *Dendrobium* species by analyzing the transcriptome and metabolome. *Int. J. Mol. Sci.* 23, 11980. doi: 10.3390/ijms231911980

Liu, N., Zhang, L., Zhou, Y., Tu, M., Wu, Z., Gui, D., et al. (2021b). The rhododendron plant genome database (RPGD): a comprehensive online omics database for *Rhododendron*. *BMC Genomics* 22, 376. doi: 10.1186/s12864-021-07704-0

Li, X. Y., Zhou, J. W., Yan, Z. Y., and Chen, X. (2020d). Sequencing and analysis of transcriptome to reveal regulation of gene expression in *Salvia miltiorrhiza* under moderate drought stress. *Chin. Traditional Herbal Drugs* 51, 1600–1608. doi: 10.7501/j.issn.0253-2670.2020.06.029

Lu, X. Y., Gu, X. L., and Shi, Y. J. (2022). A review on lignin antioxidants: Their sources, isolations, antioxidant activities and various applications. *Int. J. Biol. Macromolecules* 210, 716–741. doi: 10.1016/j.ijbiomac.2022.04.228

Ma, L. G., Dong, C. M., Song, C., Wang, X. L., Zheng, X. K., Niu, Y., et al. (2021). *De novo* genome assembly of the potent medicinal plant *Rehmannia glutinosa* using nanopore technology. *Comput. Struct. Biotechnol. J.* 19, 3954–3963. doi: 10.1016/j.csbj.2021.07.006

Ma, T. L., Li, W. J., Hong, Y. S., Zhou, Y. M., Tian, L., Zhang, X. G., et al. (2022). TMT based proteomic profiling of *Sophora alopecuroides* leaves reveal flavonoid biosynthesis processes in response to salt stress. *J. Proteomics* 253, 104457. doi: 10.1016/j.jprot.2021.104457

Mangal, M., Sagar, P., Singh, H., Raghava, G. P., and Agarwal, S. M. (2013). NPACT: naturally occurring plant-based anti-cancer compound-activity-target database. *Nucleic Acids Res.* 41, D1124–D1129. doi: 10.1093/nar/gks1047

Marchiosi, R., Dos Santos, W. D., Constantin, R. P., De Lima, R. B., Soares, A. R., Finger-Teixeira, A., et al. (2020). Biosynthesis and metabolic actions of simple phenolic acids in plants. *Phytochem. Rev.* 19, 865–906. doi: 10.1007/s11101-020-09689-2

McKernan, K. J., Helbert, Y., Kane, L. T., Ebling, H., Zhang, L., Liu, B., et al. (2020). Sequence and annotation of 42 *Cannabis* genomes reveals extensive copy number variation in cannabinoid synthesis and pathogen resistance genes. *bioRxiv* 2020, 2001.2003.894428. doi: 10.1101/2020.01.03.894428

Meng, F. B., Tang, Q., Chu, T. Z., Li, X. H., Lin, Y., Song, X. M., et al. (2022). TCMPG: an integrative database for traditional Chinese medicine plant genome. *Horticulture Res.* 9, uhac060. doi: 10.1093/hr/uhac060

Miao, Y., Luo, D., Zhao, T., Du, H., Liu, Z., Xu, Z., et al. (2022). Genome sequencing reveals chromosome fusion and extensive expansion of genes related to secondary metabolism in *Artemisia argyi*. *Plant Biotechnol. J.* 20, 1902–1915. doi: 10.1111/pbi.13870

Mohanraj, K., Karthikeyan, B. S., Vivek-Ananth, R. P., Chand, R. P. B., Aparna, S. R., Mangalapandi, P., et al. (2018). IMPPAT: a curated database of Indian medicinal plants, phytochemistry and therapeutics. *Sci. Rep.* 8, 4329. doi: 10.1038/s41598-018-22631-z

More, G. K., Vervoort, J., Steenkamp, P. A., and Prinsloo, G. (2022). Metabolomic profile of medicinal plants with anti-RVFV activity. *Heliyon* 8, e08936. doi: 10.1016/j.heliyon.2022.e08936

Mortimer, J. C. (2019). Plant synthetic biology could drive a revolution in biofuels and medicine. *Exp. Biol. Med.* 244, 323–331. doi: 10.1177/1535370218793890

Moyer, T. B., Allen, J. L., Shaw, L. N., and Hicks, L. M. (2021a). Multiple classes of antimicrobial peptides in *Amaranthus tricolor* revealed by prediction, proteomics, and mass spectrometric characterization. *J. Natural products* 84, 444–452. doi: 10.1021/acs.jnatprod.0c01203

Moyer, T. B., Brechbill, A. M., and Hicks, L. M. (2021b). Mass spectrometric identification of antimicrobial peptides from medicinal seeds. *Molecules* 26 (23), 7304. doi: 10.3390/molecules26237304

Mrudulakumari Vasudevan, U., and Lee, E. Y. (2020). Flavonoids, terpenoids, and polyketide antibiotics: Role of glycosylation and biocatalytic tactics in engineering glycosylation. *Biotechnol. Adv.* 41, 107550. doi: 10.1016/j.biotechadv.2020.107550

Mumtaz, A., Ashfaq, U. A., Ul Qamar, M. T., Anwar, F., Gulzar, F., Ali, M. A., et al. (2017). MPD3: a useful medicinal plants database for drug designing. *Natural Product Res.* 31, 1228–1236. doi: 10.1080/14786419.2016.1233409

Naika, M. B. N., Sathyanarayanan, N., Sajeevan, R. S., Bhattacharyya, T., Ghosh, P., Iyer, M. S., et al. (2022). Exploring the medicinally important secondary metabolites landscape through the lens of transcriptome data in fenugreek (Trigonella foenum graecum l.). *Sci. Rep.* 12, 13534. doi: 10.1038/s41598-022-17779-8

Nawrot, R., Lippmann, R., Matros, A., Musidlak, O., Nowicki, G., and Mock, H. P. (2017). Proteomic comparison of *Chelidonium majus* l. latex in different phases of plant development. *Plant Physiol. Biochem.* 112, 312–325. doi: 10.1016/j.plaphy.2017.01.010

Nazir, M., Saleem, M., Tousif, M. I., Anwar, M. A., Surup, F., Ali, I., et al. (2021). Meroterpenoids: A comprehensive update insight on structural diversity and biology. *Biomolecules* 11, 957. doi: 10.3390/biom11070957

Nisar, R., Ahmad, S., Khan, K.-U.-R., Sherif, A. E., Alasmari, F., Almuqati, A. F., et al. (2022). Metabolic profiling by GC-MS, *in vitro* biological potential, and in silico molecular docking studies of *Verbena officinalis*. *Molecules* 27 (19), 6685. doi: 10.3390/molecules27196685

Ntie-Kang, F., Telukunta, K. K., Döring, K., Simoben, C. V., A. Moumbock, A. F., Malange, Y. I., et al. (2017). NANPDB: A resource for natural products from northern African sources. *J. Natural products* 80, 2067–2076. doi: 10.1021/acs.jnatprod.7b00283

Ohta, Y., Atsumi, G., Yoshida, C., Takahashi, S., Shimizu, M., Nishihara, M., et al. (2021). Post-transcriptional gene silencing of the chalcone synthase gene CHS causes corolla lobe-specific whiting of Japanese gentian. *Planta* 255, 29. doi: 10.1007/s00425-021-03815-w

Pandita, D., Pandita, A., Wani, S. H., Abdelmohsen, S., Alyousef, H. A., Abdelbacki, A. M. M., et al. (2021). Crosstalk of multi-omics platforms with plants of therapeutic importance. *Cells* 10 (6), 1296. doi: 10.3390/cells10061296

Pan, L. M., Wan, L. Y., Song, L. S., He, L. L., Jiang, N., Long, H. R., et al. (2022). Comparative proteomic analysis provides new insights into the development of haustorium in *Taxillus chinensis* (DC.) danser. *BioMed. Res. Int.* 2022, 9567647. doi: 10.1155/2022/9567647

Panya, A., Yongpitakwattana, P., Budchart, P., Sawasdee, N., Krobthong, S., Paemanee, A., et al. (2019). Novel bioactive peptides demonstrating anti-dengue virus activity isolated from the Asian medicinal plant *Acacia catechu*. *Chem. Biol. Drug Design* 93, 100–109. doi: 10.1111/cbdd.13400

Pei, L., Wang, B., Ye, J., Hu, X., Fu, L., Li, K., et al. (2021). Genome and transcriptome of papaver somniferum Chinese landrace CHM indicates that massive genome expansion contributes to high benzylisoquinoline alkaloid biosynthesis. *Horticulture Res.* 8, 5. doi: 10.1038/s41438-020-00435-5

Pilon, A. C., Valli, M., Dametto, A. C., Pinto, M. E. F., Freire, R. T., Castro-Gamboa, I., et al. (2017). NuBBEDB: an updated database to uncover chemical and biological information from Brazilian biodiversity. *Sci. Rep.* 7, 7215. doi: 10.1038/s41598-017-07451-x

Qin, C., Yu, C. S., Shen, Y., Fang, X. D., Chen, L., Min, J. M., et al. (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. *Proc. Natl. Acad. Sci.* 111, 5135–5140. doi: 10.1073/pnas.1400975111

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, 405. doi: 10.1038/s41467-020-20508-2

Sakagami, H., Amano, S., Uota, S., Tanuma, S. I., Inomata, M., Shindo, A., et al. (2022). Prominent anti-UVC activity of lignin degradation products. *In Vivo* 36, 2689–2699. doi: 10.21873/invivo.13004

Sandjo, L. P., and Kuete, V. (2013a). "3 - diterpenoids from the medicinal plants of Africa," in *Medicinal plant research in Africa*. Ed. V. Kuete. (Oxford: Elsevier), 105–133.

Sandjo, L. P., and Kuete, V. (2013b). "4 - triterpenes and steroids from the medicinal plants of Africa," in *Medicinal plant research in Africa*. Ed. V. Kuete. (Oxford: Elsevier), 135–202.

Shafi, K. M., Sajeevan, R. S., Kouser, S., Vishnuprasad, C. N., and Sowdhamini, R. (2022). Transcriptome profiling of two moringa species and insights into their antihyperglycemic activity. *BMC Plant Biol.* 22, 561. doi: 10.1186/s12870-022-03938-6

She, J., Yan, H., Yang, J., Xu, W., and Su, Z. (2019). croFGD: *Catharanthus roseus* functional genomics database. *Front. Genet.* 10. doi: 10.3389/fgene.2019.00238

Shi, M., Hua, Q., and Kai, G. Y. (2021). Comprehensive transcriptomic analysis in response to abscisic acid in *Salvia miltiorrhiza*. *Plant Cell Tissue Organ Culture (PCTOC)* 147, 389–404. doi: 10.1007/s11240-021-02135-x

Sierro, N., Battey, J. N. D., Ouadi, S., Bakaher, N., Bovet, L., Willig, A., et al. (2014). The tobacco genome sequence and its comparison with those of tomato and potato. *Nat. Commun.* 5, 3833. doi: 10.1038/ncomms4833

Smith, D. R. (2015). Mutation aates in plastid genomes: They are lower than you might think. *Genome Biol. Evol.* 7, 1227–1234. doi: 10.1093/gbe/evv069

Song, Y., Liu, J. Y., Wang, J. Z., and Liu, F. D. (2021). Growth, stoichiometry, and palatability of *Suaeda salsa* from different habitats are demonstrated by differentially expressed proteins and their enriched pathways. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.733882

Su, H., Chu, Y., Bai, J., Gong, L., Huang, J., Xu, W., et al. (2019). Genome-wide identification and comparative analysis for OPT family genes in *Panax ginseng* and eleven flowering plants. *Molecules* 24, 15. doi: 10.3390/molecules24010015

Su, X. L., Liu, Y. Y., Han, L., Wang, Z. J., Cao, M. Y., Wu, L. P., et al. (2021). A candidate gene identified in converting platycoside e to platycodin d from platycodon grandiflorus by transcriptome and main metabolites analysis. *Sci. Rep.* 11, 9810. doi: 10.1038/s41598-021-89294-1

Su, X. J., Yang, L. L., Wang, D. L., Shu, Z. Q., Yang, Y. C., Chen, S. L., et al. (2022). 1 K medicinal plant genome database: an integrated database combining genomes and metabolites of medicinal plants. *Horticulture Res.* 9, uhac075. doi: 10.1093/hr/uhac075

Tchimene, M. K., Okunji, C. O., Iwu, M. M., and Kuete, V. (2013). "1 - monoterpenes and related compounds from the medicinal plants of Africa," in *Medicinal plant research in Africa*. Ed. V. Kuete. (Oxford: Elsevier), 1–32.

Tian, S. L., Yang, Y. Y., Wu, T., Luo, C., Li, X., Zhao, X. J., et al. (2022). Functional characterization of a flavone synthase that participates in a kumquat flavone metabolon. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.826780

van Bakel, H., Stout, J. M., Cote, A. G., Tallon, C. M., Sharpe, A. G., Hughes, T. R., et al. (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 12, R102. doi: 10.1186/gb-2011-12-10-r102

Van Nguyen, T., Kim, S. W., Min, C. W., Gupta, R., Lee, G. H., Jang, J. W., et al. (2021). Optimization of protein isolation and label-free quantitative proteomic analysis in four different tissues of Korean ginseng. *Plants* 10 (7), 1409. doi: 10.3390/plants10071409

Wang, Z. J., Jiang, W. M., Liu, Y. Y., Meng, X. X., Su, X. L., Cao, M. Y., et al. (2021). Putative genes in alkaloid biosynthesis identified in *Dendrobium officinale* by correlating the contents of major bioactive metabolites with genes expression between protocorm-like bodies and leaves. *BMC Genomics* 22, 579. doi: 10.1186/s12864-021-07887-6

Wang, W. L., Xu, J. F., Fang, H. Y., Li, Z. J., and Li, M. H. (2020). Advances and challenges in medicinal plant breeding. *Plant Sci.* 298, 110573. doi: 10.1016/j.plantsci.2020.110573

Wang, X., Zhang, J., He, S., Gao, Y., Ma, X., Gao, Y., et al. (2018). HMOD: An omics database for herbal medicine plants. *Mol. Plant* 11, 757–759. doi: 10.1016/j.molp.2018.03.002

Wu, Z. Q., Liao, X. Z., Zhang, X. N., Tembrock, L. R., and Broz, A. (2020). Genomic architectural variation of plant mitochondria–a review of multichromosomal structuring. *J. Systematics Evol.* doi: 10.1111/jse.12655

Wu, L. P., Meng, X. X., Huang, H. Z., Liu, Y. Y., Jiang, W. M., Su, X. L., et al. (2022b). Comparative proteome and phosphoproteome analyses reveal different molecular mechanism between stone planting under the forest and greenhouse planting of *Dendrobium huoshanense*. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.937392

Wu, J., Meng, X. X., Jiang, W. M., Wang, Z. J., Zhang, J., Meng, F., et al. (2022a). Qualitative proteome-wide analysis reveals the diverse functions of lysine crotonylation in *Dendrobium huoshanense*. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.822374

Xiao, Q., Li, Z., Qu, M., Xu, W., Su, Z., and Yang, J. (2021). LjaFGD: *Lonicera japonica* functional genomics database. *J. Integr. Plant Biol.* 63, 1422–1436. doi: 10.1111/jipb.13112

Xu, L. P., Hu, Y. B., Jin, G. Z., Lei, P., Sang, L. Q., Luo, Q. X., et al. (2021). Physiological and proteomic responses to drought in leaves of *Amygdalus mira* (Koehne) yü et Lu. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.620499

Xu, Z. C., Li, Z., Ren, F. M., Gao, R. R., Wang, Z., Zhang, J. L., et al. (2022). The genome of corydalis reveals the evolution of benzylisoquinoline alkaloid biosynthesis in ranunculales. *Plant J.* 111, 217–230. doi: 10.1111/tpj.15788

Yang, F. X., Gao, J., Wei, Y. L., Ren, R., Zhang, G. Q., Lu, C. Q., et al. (2021). The genome of *Cymbidium sinense* revealed the evolution of orchid traits. *Plant Biotechnol. J.* 19, 2501–2516. doi: 10.1111/pbi.13676

Yang, Y. H., Li, M. J., Yi, Y. J., Li, R. F., Dong, C., and Zhang, Z. Y. (2018). The root transcriptome of achyranthes bidentata and the identification of the genes involved in the replanting benefit. *Plant Cell Rep.* 37 (6), 611–625. doi: 10.1007/s00299-018-2255-z

Yang, L. L., Yan, Y. C., Zhao, B. Y., Xu, H. M., Su, X. H., and Dong, C. M. (2022). Study on the regulation of exogenous hormones on the absorption of elements and the accumulation of secondary metabolites in the medicinal plant *Artemisia argyi* leaves. *Metabolites* 12, 984. doi: 10.3390/metabo12100984

Yang, L., Zhang, J. C., Qu, J. T., He, G., Yu, H. Q., Li, W. C., et al. (2019). Expression response of chalcone synthase gene to inducing conditions and its effect on flavonoids accumulation in two medicinal species of anoectochilus. *Sci. Rep.* 9, 20171. doi: 10.1038/s41598-019-56821-0

Yuan, Y. D., Zuo, J. J., Zhang, H. Y., Zu, M. T., Yu, M. Y., and Liu, S. (2022). Transcriptome and metabolome profiling unveil the accumulation of flavonoids in *Dendrobium officinale*. *Genomics* 114, 110324. doi: 10.1016/j.ygeno.2022.110324

Yu, H. Z., Guo, W. L., Yang, D. F., Hou, Z. N., and Liang, Z. S. (2018). Transcriptional profiles of SmWRKY family genes and their putative roles in the biosynthesis of tanshinone and phenolic acids in *Salvia miltiorrhiza*. *Int. J. Mol. Sci.* 19, 1593. doi: 10.3390/ijms19061593

Yu, D. L., Lu, J. J., Shao, W. S., Ma, X. X., Xie, T., Ito, H., et al. (2019). MepmiRDB: a medicinal plant microRNA database. *Database* 2019, baz070. doi: 10.1093/database/baz070

Zhang, Z. M., Su, Q., Xia, B. H., Li, Y. M., Qin, X. Y., Luo, H. S., et al. (2022b). Integrative transcriptomic, proteomic and metabolomic analysis reveals the dynamic regulation of secondary metabolism upon development of *Prunella vulgaris l. Fitoterapia* 163, 105334. doi: 10.1016/j.fitote.2022.105334

Zhang, D., Yang, Z. R., Song, X. Q., Zhang, F. L., and Liu, Y. (2022a). TMT-based proteomic analysis of liquorice root in response to drought stress. *BMC Genomics* 23, 524. doi: 10.1186/s12864-022-08733-z

Zhang, S. C., Zhang, L., Zou, H. Y., Qiu, L., Zheng, Y. W., Yang, D. F., et al. (2021). Effects of light on secondary metabolite biosynthesis in medicinal plants. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.781236

Zhan, C. S., Li, X. H., Zhao, Z. Y., Yang, T. W., Wang, X. K., Luo, B. B., et al. (2016). Comprehensive analysis of the triterpenoid saponins biosynthetic pathway in *Anemone flaccida* by transcriptome and proteome profiling. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01094

Zhao, J. L., Bayer, P. E., Ruperao, P., Saxena, R. K., Khan, A. W., Golicz, A. A., et al. (2020). Trait associations in the pangenome of pigeon pea (*Cajanus cajan*). *Plant Biotechnol. J.* 18, 1946–1954. doi: 10.1111/pbi.13354

Zhao, Q., Zhang, Y., Wang, G., Hill, L., Weng, J. K., Chen, X. Y., et al. (2016). A specialized flavone biosynthetic pathway has evolved in the medicinal plant, *Scutellaria baicalensis*. *Sci. Adv.* 2, e1501780. doi: 10.1126/sciadv.1501780

Zhong, C., Chen, C., Gao, X., Tan, C., Bai, H., and Ning, K. (2022). Multi-omics profiling reveals comprehensive microbe–plant–metabolite regulation patterns for medicinal plant glycyrrhiza uralensis fisch. *Plant Biotechnol. J.* 20, 1874–1887. doi: 10.1111/pbi.13868

Zhou, G. R., Liao, B. S., Li, Q. S., Xu, J., and Chen, S. L. (2021a). Establishing a genomic database for the medicinal plants in the Brazilian pharmacopoeia. *Chin. Med.* 16, 71. doi: 10.1186/s13020-021-00484-5

Zhou, C. H., Lin, C. C., Xing, P. Y., Li, X. F., and Song, Z. Q. (2022). SmGDB: genome database of *Salvia miltiorrhiza*, an important TCM plant. *Genes Genomics* 44, 699–707. doi: 10.1007/s13258-022-01251-y

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 March 2023
DOI 10.3389/fpls.2023.1131173

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Archana Panche,
Mahatma Gandhi Mission University, Maharashtra, India
Makhlouf Chaalal,
Université Frères Mentouri Constantine 1, Algeria

*CORRESPONDENCE
Pankaj Kumar Sharma
pankajsharma@pilani.bits-pilani.ac.in

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

RECEIVED 24 December 2022
ACCEPTED 22 February 2023
PUBLISHED 08 March 2023

CITATION
Joshi T, Mandal SK, Puri S, Asati V, Deepa PR and Sharma PK (2023) Investigating the antioxidant activity enhancer effect of *Cyamopsis tetragonoloba* seed extract on phenolic phytochemicals.
Front. Plant Sci. 14:1131173.
doi: 10.3389/fpls.2023.1131173

Investigating the antioxidant activity enhancer effect of *Cyamopsis tetragonoloba* seed extract on phenolic phytochemicals

Tripti Joshi, Sumit Kumar Mandal, Sonakshi Puri, Vidushi Asati, P. R. Deepa and Pankaj Kumar Sharma*

Department of Biological Sciences, Birla Institute of Technology and Science (BITS), Rajasthan, India

Introduction: Phenolic phytochemicals are known for antioxidant-mediated pharmacological effects in various diseases (diabetes, cancer, CVDs, obesity, inflammatory and neurodegenerative disorders). However, individual compounds may not exert the same biological potency as in combination with other phytochemicals. *Cyamopsis tetragonoloba* (Guar), an underutilized semi-arid legume which has been used as a traditional food in Rajasthan (India), is also a source of the important industrial product guar gum. However, studies on its biological activity, like antioxidant, are limited.

Methods: We tested the effect of *C. tetragonoloba* seed extract to enhance the antioxidant activity of well-known dietary flavonoids (quercetin, kaempferol, luteolin, myricetin, and catechin) and non-flavonoid phenolics (caffeic acid, ellagic acid, taxifolin, epigallocatechin gallate (EGCG), and chlorogenic acid) using DPPH radical scavenging assay. The most synergistic combination was further validated for its cytoprotective and anti-lipid peroxidative effects in *in vitro* cell culture system, at different concentrations of the extract. LC-MS analysis of purified guar extract was also performed.

Results and discussion: In most cases, we observed synergy at lower concentrations of the seed extract (0.5-1 mg/ml). The extract concentration of 0.5 mg/ml enhanced the antioxidant activity of Epigallocatechin gallate (20 µg/ml) by 2.07-folds, implicating its potential to act as an antioxidant activity enhancer. This synergistic seed extract-EGCG combination diminished the oxidative stress nearly by double-fold when compared with individual phytochemical treatments in *in vitro* cell culture. LC-MS analysis of the purified guar extract revealed some

previously unreported metabolites, including catechin hydrate, myricetin-3-galactoside, gossypetin-8-glucoside, and puerarin (daidzein-8-C-glucoside) which possibly explains its antioxidant enhancer effect. The outcomes of this study could be used for development of effective nutraceutical/dietary supplements.

KEYWORDS

Cyamopsis tetragonoloba, phenolics, antioxidant activity enhancer, cytoprotective effect, lipid peroxidation, nutraceutical

Introduction

Legumes, or dry beans and pulses, are members of the Fabaceae family that grow in pods of annual, biennial, and perennial plants. They are not only one of the largest but also among the economically most significant families of flowering plants due to their nitrogen-fixing capacity and restoration of nitrogen-depleted soil by crop rotation (Mbagwu et al., 2011). Legumes are generally recognized for their high concentrations of bioactive components, including phenolics, phytosterols, carbohydrates, and saponins, which help lower the risk of oxidizing substances, bacteria, diabetes, inflammatory disease, and cancer (Ayilara et al., 2022; Jat et al., 2023). Numerous research endeavours have discovered the antioxidant properties of different species of legumes, and they have found a strong correlation between antioxidant potential and total phenolic content (Asati et al., 2022; Timoracká et al., 2022). At a time when one out of every five children under the age of five is chronically malnourished, legumes are now considered a future superfood capable of eradicating hunger and contributing to health (Martín-Cabrejas, 2018; Chaudhary et al., 2022). Due to their low cost and positive environmental impact, their natural bioactive compounds are currently a trend in the food processing industry (Riaz et al., 2022). As consumers become more aware of the nutritional and nutraceutical composition of legumes, their global demand continues to rise (Pham and Luan, 2021).

Cyamopsis tetragonoloba (Guar), has been used traditionally for food and fodder purposes (Pankaj and Dhankar, 2023) (Figures 1A, B). It is economically important and also known as the heart of the farmer fields, as India contributes 80% of the global guar gum production. Guar can be used as a laxative, digestive aid, appetizer, or cooling agent (Mukhtar et al., 2004). Potentially, guar gum can help hypercholesterolemic insulin-dependent diabetic patients with improved glycemic control and lower serum LDL-cholesterol concentrations (Vuorinen-Markkola et al., 1992). Potent phytochemicals including phenolics and flavonoids are found in the seeds (Kuravadi et al., 2012). Owing to the presence of multiple therapeutically active molecules, like quercetin, daidzein, and kaempferol, it is used as a complementary medicinal plant (Jain and Rijhwani, 2018). For example, Kaushik et al. (2020) reported that *C. tetragonoloba* could play an important role in developing inexpensive and effective anti-dengue medicine.

Plant secondary metabolites are multifunctional metabolites produced from various biochemical pathways. The biosynthesis of aromatic amino acids, tryptophan, tyrosine, and phenylalanine, which are common precursors for phenolics and nitrogen-containing compounds is initiated by the shikimate pathway (Jan et al., 2021; Elhamouly et al., 2022). As per García-Calderón et al. (2020), the secondary metabolites originating from the phenylpropanoid metabolism include monolignols, flavonoids and isoflavonoids, various phenolic acids, and stilbenes, which function to: protect the plants against oxidative stress and pathogens; also as chemical signals in symbiotic nitrogen fixation with rhizobia. Further, despite the important role played by the *Lotus japonicus* (a model legume) in elucidating the molecular genetics of legume–rhizobia symbiosis, the class of phenolic compounds used by this species in order to attract its chosen symbiont is still unknown. Species-specific differences in flavonoid accumulation have also been observed. For example, in *L. japonicas*, different types of abiotic stress situations (such as UV-B irradiation) resulted in an accumulation of isoflavonoids as a possible alternative to accumulation of flavonols.

Phenolics are a broad class of bioactive compounds that contain at least one benzene ring and one or more hydroxyl groups. The complexity of phenolic compounds ranges from simple phenols to highly polymerized compounds (Lin et al., 2016). These compounds are differentially distributed in the cotyledon (mainly non-flavonoid phenolics) and the seed coat (flavonoids) of legumes (Amarowicz, 2020). The distinctive bioactive potential, color and flavor of legumes are due to the most abundant phenolic compound, flavonoids, which are composed of two aromatic rings linked by a 3-C bridge, in the form of heterocyclic C ring (Pham and Luan, 2021).

In humans, oxidative stress is caused by an imbalance between the formation of reactive oxygen species (ROS) and the endogenous antioxidants, leading to a reaction cascade that can damage lipids, proteins, and DNA (Rudrapal et al., 2022). Antioxidants operate as scavengers of reactive free radicals, inhibiting lipid peroxidation and other related processes, and thereby protect the body from resulting diseases (Pande and Srinivasan, 2013b). The antioxidant properties of phenolic compounds are influenced by their chemical makeup. The most important aspects of flavonoids' activity as main antioxidants are the position and degree of hydroxylation on the B ring (Kerwin, 2004).

According to the reports by World Health Organisation (WHO), nearly 80% of the global population depends on plant based medicines owing to their positive impact on health and lower side-effects (Adki et al., 2020; Singh and Gaikwad, 2020). There is a

10.3389/fpls.2023.1131173

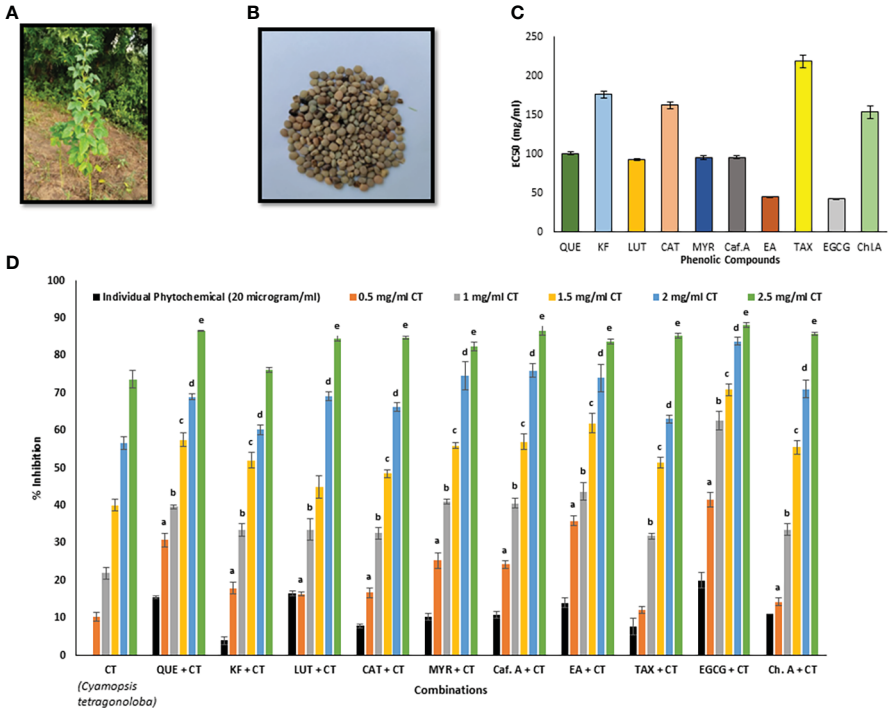


FIGURE 1
(A) *C. tetragonoloba* plant, **(B)** *C. tetragonoloba* seeds, **(C)** EC_{50} values of individual phenolic compounds, **(D)** DPPH inhibition% of phenolic compounds with and without *C. tetragonoloba* seed extract. CT, *C. tetragonoloba*; QUE, quercetin; KF, kaempferol; LUT, luteolin; CAT, catechin; MYR, myricetin; Caf. A, caffeic acid; EA, ellagic acid; TAX, taxifolin; EGCG, epigallocatechin gallate; Ch. A, chlorogenic acid. Values denote mean ± SD (n=3). a, b, c, d, and e represent statistically significant different values ($P \leq 0.05$) with respect to 0.5, 1, 1.5, 2, & 2.5 mg/ml concentration of CT, respectively.

plethora of research stating the antioxidant potential of individual pure phytochemicals or plant extracts. However, despite knowing their excellent antioxidant activity and natural origin, studies on the biological activities of their combinations are surprisingly limited. It is reasonable to hypothesize that edible plant extracts can be used to enhance the antioxidant activity of known phytochemicals, thereby increasing their bioactive potential. The current research focuses on investigating the antioxidant activity enhancer (AAE) effect of *Cyamopsis tetragonoloba* seed extract on dietary flavonoids (quercetin, kaempferol, luteolin, myricetin, and catechin) and non-flavonoid phenolics (caffeic acid, ellagic acid, taxifolin, epigallocatechin gallate (EGCG), and chlorogenic acid) using DPPH radical scavenging assay and lipid peroxidation assessment in *in vitro* cultured cells. The total phenolic content and total flavonoid content were calculated. LC-MS analysis was performed to identify possibly novel and unreported compounds (also as potential contributors to antioxidant activity enhancement of standard phenolic phytochemicals) from the seed extract of *C. tetragonoloba*.

Materials and methods

Chemicals and reagents

Quercetin, kaempferol methanol, dimethyl sulfoxide (DMSO), DPPH, Folin Ciocalteu reagent, sodium carbonate, aluminum chloride, sodium nitrite, sodium hydroxide, Amberlite XAD7HP, Sephadex LH-20 were procured from Sigma-Aldrich Chemicals Company (United States). Luteolin, catechin, myricetin, caffeic acid, ellagic acid, taxifolin, epigallocatechin gallate (EGCG), chlorogenic acid, gallic acid were obtained from Yucca Enterprises (Mumbai, India), Mouse embryonic fibroblast cells (3T3-L1) were obtained from National Centre for Cell Science (NCCS, Pune, India), Dulbecco's Modified Eagle's Medium (DMEM) and fetal bovine serum (FBS) were purchased from Gibco Life Technologies (Carlsbad, CA, USA), 2-Thiobarbituric acid (TBA), Trichloroacetic acid (TCA) were purchased from Sigma Aldrich.

Plant collection and extraction

Commercially available seeds of *Cyamopsis tetragonoloba* (guar) were purchased from a local grocery store in Pilani market (Jhunjhunu district, Rajasthan, India).

Preparation of seed extract

The seeds of *Cyamopsis tetragonoloba* were ground to a fine powder using a Waring blender. The fine powder was defatted with hexane in a 1:5 (w/v) ratio at room temperature for 1 hr in an orbital shaker incubator. This was followed by centrifugation at 3000xg for 10 min. The supernatant was decanted and the pellet was extracted

two more times. The hydrophobic compounds were separated in the hexane extract and the pellet was dried at room temperature. This dried pellet was extracted thrice with 5 volumes of 80% methanol by the same process mentioned above. The supernatant was filtered using a Whatman filter paper (No. 1). This step was repeated twice, and all the supernatants were pooled and concentrated to dryness using a rotary evaporator (Aditya Scientific, Hyderabad, India). The concentrated extracts were stored at 4°C for further analysis.

Total phenolic content

The TPC was determined by the Folin-Ciocalteu method described by (Slinkard and Singleton, 1977; Tsao et al., 2003; Wang et al., 2011) with slight modifications. Briefly, 200 µl of the seed extract was mixed with 800 µl of 7.5% sodium carbonate and 1ml of the FCR (Folin-Ciocalteu Reagent). The mixture was shaken gently and incubated at room temperature for 30 min, and the absorbance was read at 765 nm. A gallic acid standard curve was prepared with different concentrations (50-250 µg/ml), and the TPC values were expressed as micrograms of gallic acid equivalents (GAE) per gram of sample. All tests were performed in triplicates.

Total flavonoid content

The TFC was calculated by the Aluminum chloride colorimetric assay described by John et al. (2014). In brief, an aliquot (1ml) of extracts or standard solutions of quercetin (200-1000 µg/ml) was added in a flask containing 4 ml distilled water. To the flask was added 300 µl 5% $NaNO_2$, followed by 300 µl 10% $AlCl_3$ after five minutes. This was followed by addition of 2ml 1M NaOH, and the volume was made up to 10 ml with distilled water. The solution was shaken, and absorbance was read at 510 nm. The TFC values were expressed as mg of quercetin equivalents per g of sample. All tests were performed in triplicates.

Estimation of antioxidant potential

The antioxidant activity of plant seed extract and pure phytochemicals was tested by DPPH (2,2-diphenylpicrylhydrazyl) assay (after 2-fold dilution). For binary combinations (to test the potential enhancement of the antioxidant activity of pure phytochemicals by *C. tetragonoloba* seed extract), the seed extract was used in varying concentrations (0.5-2.5 mg/ml) whereas the concentration of pure compounds was kept constant (20 µg/ml). These were mixed in a 1:1 (v/v) ratio.

DPPH free radical scavenging assay

The DPPH assay was performed as described by Hidalgo et al. (2010). The methanolic solution of DPPH is purple/violet colored, which fades to pale yellow in the presence of antioxidants, and the loss in absorbance is measured at 517 nm. A 100 µM DPPH solution was prepared in methanol, and 290 µl of this solution was mixed with 10 µl of individual compound/seed extract or their combinations. The reaction was carried out in a 96-well microplate, incubated in the dark at room temperature for 1hr, and absorbance was measured at 517 nm using a microplate reader (ThermoScientific Multiskan G0). The percentage DPPH radical scavenging activity was calculated by the following equation:

$$Inhibition\ \ \% = \frac{A_c - A_s}{A_c}$$

Where A_c is the absorbance of the control and A_s is the absorbance of the sample. Solution without the sample (seed extract or phytochemical) was taken as control. The results were expressed as EC_{50} (µM) obtained by plotting a curve between concentration and inhibition percentage. EC_{50} is the effective concentration necessary to get 50% inhibition. The lower the EC_{50} value, higher will be the antioxidant activity.

Cell viability assay

The non-toxic dosage of test phytochemicals - CT and EGCG, was determined by MTT (3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) based cell viability assay, which shows the metabolic activity of cells. Murine fibroblast NIH-3T3 cells (NCCS, Pune, India) were cultured at 37°C in humidified 5% CO_2 in Dulbecco's modified eagle medium (DMEM) supplemented with 10% (vol/vol) fetal bovine serum (FBS), penicillin G and streptomycin (100 mg/l). After the cells were confluent, cells were trypsinized from the surface of the culture flask by using a 0.25% trypsin solution. The cells were plated on the cultivation flask (surface 25 cm^2) at a density of 6 X 10^4 NIH-3T3 cells per ml medium and incubated for 24 h prior to the experiments. The NIH/3T3 cells were seeded in a 96-well (8×10^3 cells per well) and were incubated for 24 h. The test phytochemicals, CT and EGCG, were dissolved in DMSO (stock solution 10 mM), and diluted in media to a final concentration of 5 µM to 50 µM (concentration of DMSO 0.5%). After 24 h of incubation, 90 µl of growth medium and 10 µl MTT dye (5 mg/ml) were added to each well and incubated for an additional three hours. The MTT solution-containing media was then removed. After adding DMSO the plate was shaken gently to dissolve the formazan crystals. The absorbance was measured at 570 nm (Multiskan FC, Thermo Scientific, DE) (Ahmad et al., 2019; Mandal et al., 2022). The percentage of cytotoxicity was determined as follows:

$$\%\ \text{cell viability} = (\text{Mean absorbance of treated group}) / (\text{Mean absorbance of control group}) \times 100$$

Lipid peroxidation assay

To evaluate the effect of phytochemicals (individual and in combination) on oxidative stress, levels of malondialdehyde

(MDA), a stable end product of lipid peroxidation was estimated by TBARS (thiobarbituric acid reactive substances) assay. NIH/3T3 cells without any treatment with H_2O_2 or phytochemicals served as the control, while the fibroblast cells exposed to 100 µM H_2O_2 for 6 hrs served as the oxidative stress induced model. NIH/3T3 cells were initially treated with the phytochemicals for 3h, followed by exposure to 100 µM H_2O_2. After 3h, the cells were treated with lysis buffer, homogenized, centrifuged at 13000 x g at 4°C for 15 minutes and the supernatant was collected. Cell lysates of each experimental group were normalized on the basis of equal amount of protein (100 µg), and incubated with 500 µl of 10% trichloroacetic acid (TCA). This was followed by reacting with 750 µl of thiobarbituric acid (TBA, 1% w/v) in an acidic condition, and the solution was heated in a boiling water bath for 15 minutes to generate a pink colour adduct which was measured spectrophotometrically (Multiskan FC, Thermo Scientific, DE) at 530 nm (Wenz et al., 2019). Values were expressed as µM of malondialdehyde/mg protein.

Purification of phenolics and flavonoids from *C. tetragonoloba* seed extract

In order to identify potential phytocompounds in the *C. tetragonoloba* seed extract leading to antioxidant activity enhancement of the pure phytochemical, the seed extract was subjected to column chromatography. Purification was done by previously reported protocol with slight modifications (Asati et al., 2022). In brief, Amberlite XAD7HP chromatography followed by Sephadex LH-20 were used to purify the defatted methanol extract. 5 gm of the extract was loaded on the matrix packed in a glass column (50 cm x 1.8 cm) and equilibrated using 100% methanol. Initially, 50% methanol was used to elute the column, and 20 fractions of 5 mL each were collected. This was followed by elution using 100% methanol. All fractions were analyzed by TLC (Thin layer chromatography) on silica gel F254 plates (Merck, USA) using Toluene: Acetic Acid: Acetone: Formic Acid::20:4:2:1 as solvent system and plates were visualized at 254 and 366 nm. The fraction exhibiting the maximum bands (C5) was further characterised using HPLC and LC-MS.

HPLC and LC-MS analysis of the purified fraction

The purified fractions were filtered by 0.45 µm Puradisc filters and HPLC (Shimadzu Corporation, Tokyo, Japan) analysis was performed for the purpose of optimization. Photo diode array detector was used; absorbance was monitored from 200 to 365 nm. The protocol reported by Song et al. (1998) for chromatographic separation was standardized by slight modifications. In brief, separation was performed on a C18 column (SpherisorbR, 250 mm x 4.6 mm, particle size 5 µm, Waters) with optimized mobile phase A (pH 3.0 Milli Q water) and B (Acetonitrile). Glacial acetic acid was used as pH modifier. The system was equilibrated for an hour at 1 ml/min flow rate. Initially, phase A and acetonitrile concentrations were 10% and 90%, respectively, for the purpose of washing. After sample injection, 0.1% glacial acetic acid in acetonitrile was maintained at 15% for 5 minutes. Solvent B reached 35% in 33 minutes and 10% in 40 minutes. The solvent flow rate was l mL/min for the first 5 min, increased to 1.5 mL/min over 0.5 min and maintained for 39 min, and then returned to 1 mL/min.

The C5 fraction obtained through Sephadex LH-20 chromatography (obtained as mentioned in the previous section on purification) was further analysed by LC-MS. This was performed as per Chaudhary et al. (2020) and Venuprasad et al. (2014). Q-TOF Micromass spectrometer (Waters Corporation, Milford, MA, USA) was used. Chromatographic separation was done using Spherisorb 5 µm ODS2 column with the help of auto sampler (flow rate of 0.2 mL/min, 280 nm wavelength and 20 µL injection volume). Solvents were: (A) Formic acid (0.1% v/v) and 10 mM ammonium fluoride and (B) acetonitrile + 0.1% Formic acid. Gradient (in solvent B) was: (i) 30%, from 0 to 15 min, (ii) 55%, from 15 min, (iii) 95%, from 25 to 45 min, and (iv) 35%, at 45–48 min; spray voltage 4 KV; gas temperature 325°C; gas flow 10 L/min; and nebulizer 40 psi. Electrospray mass spectra data were recorded on positive and negative ionization mode for a mass range *m/z* 50–*m/z* 1000. The instrument's MassLynx database was used to examine the products. RIKEN-RESPECT was used to evaluate mass spectrum fragments (Sawada et al., 2012).

Statistical analysis

Experiments were done in triplicate, and the values were calculated as mean ± standard deviation. One-way analysis of variance (ANOVA) was performed to assess the statistically significant difference between the mean values. P-value ≤ 0.05 was considered statistically significant.

Results and discussion

Total phenolic content and total flavonoid content

The antioxidant activity of plants is directly proportional to theirphenolic/flavonoid content (Joshi et al., 2022). It has been reported that phenolic compounds are best extracted with methanol (80% concentration) as compared to other solvents due to its polarity and solubility of phenolics (Moteriya, 2015). The TPC and TFC of the defatted methanolic extract of *C. tetragonoloba* seeds were 280 ± 9.5 mg GAE/g and 496 ± 15.2 mg QE/g of the extract, respectively. Various factors, including growth and storage conditions (climate, soil, water), and time of harvest are responsible for different phytochemical composition (Wright et al., 2001). Our research group has earlier reported the variation in total flavonoid content of *P. cineraria* pod extracts obtained from trees in different geographical regions (Asati et al., 2022). Sharma et al. (2017) reported that the TPC and TFC of different guar cultivars

collected from various states of India ranges between 60.03 to 204.67 mg GAE/g and 4.26 to 12.43 mg QE/g, respectively. Thus, cultivar selection is important for functional food development from traditional plants.

Antioxidant activity of standard phenolic compounds

The antioxidant activity of standard phenolics were tested by DPPH free radical scavenging assay. EGCG (EC_{50}- 42.69 ± 0.16 µg/ml) and ellagic acid (EC_{50}- 45.08 ± 0.40 µg/ml) showed the maximum antioxidant activity among the 10 tested compounds (Figure 1C). The order of antioxidant activity from lowest to highest was as follows- taxifolin< kaempferol< catechin ≤ chlorogenic acid< quercetin< myricetin = caffeic acid< luteolin< ellagic acid ≤ EGCG. This result is in accordance with the trend of antioxidant activity reported by other researchers (Hirano et al., 2001; Hidalgo et al., 2010). It is believed that the gallate group at position 3 plays the most crucial role in their ability to scavenge free radicals, with an extra hydroxyl group inserted at position 5' in the B ring also contributing to their scavenging capabilities (Heim et al., 2002; Braicu et al., 2011).

Arrangement and number of hydroxyl moieties on the ring, presence of catechol group in the B ring, and 2, 3 double bonds in the C ring, are some characteristics that strongly correlate with antioxidant potential. According to Freeman et al. (2010), these groups can also be used to find the reduction potentials, as a molecule with lower reduction potential has more tendency to donate its electron and act as a strong antioxidant. The results are in accordance with these reports; Quercetin, myricetin, and luteolin showed almost similar antioxidant activities because of almost similar structure (a catechol group in the B ring and a 2, 3 double bond in the C ring (Freeman et al., 2010)) while EGCG has the lowest reduction potential, thereby showing maximum activity.

Antioxidant activity enhancer effect of *C. tetragonoloba* seed extract on standard phenolic phytochemicals

C. tetragonoloba seed extract was used in combination with 10 phenolic compounds to test for possible synergism (antioxidant activity enhancement). The DPPH radical scavenging activity of the pure compounds with and without the seed extract were compared (Figure 1D). It was observed that with increase in the concentration of extract from 0.5-2.5 mg/ml, the DPPH % inhibition also increases. Table 1 shows the percentage inhibition of the combinations at different concentrations and the types of interaction. An interaction can be said to be synergistic when the experimental value is greater than the theoretical value (calculated by summing up the inhibition percentage pertaining to antioxidant activity of individual phytochemicals and seed extracts), additive when the experimental and theoretical values are equal; and when experimental value is less than the theoretical value, it is an antagonistic interaction.

As seen from Figure 1D, when the seed extract (varying concentrations) was added to the phenolic compounds (20 µg/ml), most of the combinations showed synergistic antioxidant effect at lower concentrations of extract (0.5-1 mg/ml), while on further increasing the extract concentration, additive (at 1.5 mg/ml) and antagonistic (at 2-2.5 mg/ml) interactions were observed. The results agreed with the previous report that phytochemicals need to be combined in specific ratios to show synergistic effect (Joshi et al., 2022). When the extract concentration was 0.5 mg/ml, the DPPH % inhibition of extract with EGCG was 41.58%, which was 4-folds and 2.07-folds higher than that of seed extract (10.37%) and EGCG (20%). All the concentrations of extract showed synergistic interaction with EGCG. Therefore, this synergistic seed extract-EGCG combination was further validated for its cytoprotective and anti-lipid peroxidative effects in *in vitro* cell culture system. Similar results were reported by Zhang et al. (2016), that mulberry leaf polysaccharides (MLPs) can be used as antioxidant activity enhancers of flavonoids. It was reported that despite having low antioxidant activity themselves, MLPs showed synergistic interaction with flavonoids. In another study, combination of *C. tetragonoloba* with garlic and capsaicin (responsible for pungent flavor of red pepper), offer a significant increase in the antioxidant status (Pande and Srinivasan, 2013a; Pande and Srinivasan, 2013b).

Romano et al. (2009) reported the antioxidant activity enhancing effect of rosemary extract on synthetic antioxidants (butylated hydroxyanisole (BHA) and butylated hydroxytoluene (BHT)). This could give the food sector a strong reason to combine natural and synthetic antioxidants in processed food products to increase storage stability and prevent any potential hazardous effects from using excessive levels of antioxidants. Flavonoids (quercetin, kaempferol, and isorhamnetin) present in the almond skin have been shown to act synergistically with vitamin E and C (Chen et al., 2005). In another study, the combination of longan peel extract (LP), vitamin E, and ascorbyl palmitate (derivative of ascorbic acid) lowered the free radicals in tuna oil, contributing to the antioxidant effect; thus, LP could have an application as a food additive against lipid oxidation in oils (Rakariyatham et al., 2021). This study also offered mechanistic insights into antioxidant synergy among phytochemicals.

As *C. tetragonoloba* and the tested phytochemicals (polyphenolic compounds) have been used for their antioxidant property for many years, the current study suggests potential usage of their combination in order to achieve a greater therapeutic effect. These findings may be helpful for people who want to increase their antioxidant intake – without compromising on safety - as well as for the development of novel medications and functional foods with higher antioxidant potential.

There are different hypotheses for possible mechanisms responsible for the above-mentioned interactions. Synergistic interactions could be due to- a) the regeneration of strong antioxidants by the weaker ones (Marinova et al., 2008), b) formation of stable intermolecular adducts with strong antioxidant activity (Olszowy, 2020), c) the type and

10.3389/fpls.2023.1131173

TABLE 1 DPPH inhibition % of all phytochemicals (each at 20 µg/ml) in combination with different concentrations of the *C. tetragonoloba* (CT) seed extract (0.5-2.5 mg/ml).

Combinations		DPPH Inhibition % at different concentrations (mg/ml) ± STDEV									
		0.5		1		1.5		2		2.5	
QUE + CT	E	30.78 ± 1.81*		39.67 ± 0.46*		57.55 ± 1.84		69.02 ± 0.76*		86.55 ± 0.15*	
	T	25.94 ± 0.39	Syn	37.53 ± 1.33	Syn	55.68 ± 2.11	Ad	72.19 ± 2.57	An	89.28 ± 2.40	An
KF + CT	E	17.90 ± 1.55*		33.49 ± 1.67*		52.06 ± 2.04*		60.22 ± 1.36		73.25 ± 0.67*	
	T	14.30 ± 1.19	Syn	25.89 ± 1.25	Syn	44.04 ± 1.93	Syn	60.55 ± 1.32	Ad	77.64 ± 2.71	An
LUT + CT	E	16.40 ± 0.55*		33.59 ± 2.91*		44.94 ± 2.98*		69.20 ± 1.14*		84.64 ± 0.77*	
	T	26.99 ± 0.24	An	38.58 ± 0.99	An	56.73 ± 2.15	An	73.24 ± 2.17	An	90.33 ± 2.61	An
CAT + CT	E	16.74 ± 1.31		32.59 ± 1.67		48.52 ± 1.10		66.26 ± 1.12		84.87 ± 0.38*	
	T	18.24 ± 1.43	Ad	29.83 ± 2.04	Ad	47.98 ± 0.96	Ad	64.49 ± 2.25	Ad	81.58 ± 1.66	Syn
MYR + CT	E	25.33 ± 2.16*		41.14 ± 0.64*		55.98 ± 0.69*		71.68 ± 3.77		80.51 ± 1.09*	
	T	20.75 ± 0.94	Syn	32.34 ± 1.58	Syn	50.49 ± 1.45	Syn	67.00 ± 2.12	Ad	84.09 ± 2.02	An
Caf. A + CT	E	24.31 ± 1.04*		40.61 ± 1.36*		57.01 ± 2.07*		76.04 ± 1.90*		86.72 ± 1.20	
	T	21.18 ± 0.51	Syn	32.77 ± 0.69	Syn	50.92 ± 2.32	Syn	67.43 ± 1.82	Syn	84.52 ± 2.88	Ad
EA + CT	E	35.89 ± 1.33*		43.74 ± 2.37*		61.92 ± 2.57*		73.97 ± 3.71		86.15 ± 0.70*	
	T	24.39 ± 1.94	Syn	35.98 ± 2.39	Syn	54.13 ± 0.81	Syn	70.64 ± 2.14	Ad	87.73 ± 1.77	An
TAX + CT	E	12.06 ± 0.88*		31.88 ± 0.80		51.49 ± 1.42*		63.07 ± 0.97		85.36 ± 0.62*	
	T	18.01 ± 2.05	An	29.60 ± 2.75	Ad	47.75 ± 0.35	Syn	64.26 ± 2.88	Ad	81.35 ± 0.99	Syn
EGCG + CT	E	41.58 ± 1.89*		62.67 ± 2.50*		70.91 ± 1.60*		83.82 ± 1.03*		88.20 ± 0.69*	
	T	22.37 ± 2.74	Syn	33.96 ± 3.54	Syn	52.11 ± 0.83	Syn	68.62 ± 3.74	Syn	85.71 ± 0.20	Syn
Ch. A + CT	E	14.28 ± 1.09*		33.57 ± 1.51		55.50 ± 1.93*		71.09 ± 2.38		85.77 ± 0.44	
	T	21.46 ± 1.84	An	33.05 ± 2.78	Ad	51.20 ± 1.83	Syn	67.71 ± 3.76	Ad	84.80 ± 1.44	Ad

Expanded forms of the abbreviations for the phytochemicals are given in the legend for Figure 1.
E, Experimental value; T, Theoretical value; Syn, Synergistic; Ad, Additive; An, Antagonistic interaction. All experiments were done in triplicates. *P-value ≤ 0.05 was considered statistically significant.

concentration of antioxidant (Shi et al., 2007). Hypotheses for antagonistic interactions are- a) regeneration of weaker antioxidants by stronger antioxidants, b) polymerization of antioxidants decreases their activity, c) disappearance of free antioxidant radicals due to irreversible reactions. Further studies need to be conducted to validate the specific mechanism for antioxidant synergism observed between the tested phytochemicals and *C. tetragonoloba* seed extract in the current study.

Validation of antioxidant activity enhancer effect of CT seed extract towards EGCG in cultured fibroblast cells

Non-cytotoxic dosage of phytochemicals determined in NIH/3T3 cells

The viability of the normal fibroblast cells (NIH/3T3) was assessed by MTT assay to evaluate the non-cytotoxic dosage range for CT and EGCG, at various concentrations. The safe dosage of CT, and EGCG was considered to be the concentration at which at least 80% of cells were viable (non-toxic dosage). After 24 hours incubation, the test phytochemicals showed dose-dependent decrease in cell viability, wherein CT and EGCG revealed 80% cell viability upto a dosage of 20 µg/ml (Figures 2A, B).

Lipid peroxidation assay

Following H_2O_2 treatment, a significant increment in the concentration of MDA ($P \leq 0.05$) was observed in the normal fibroblast cells, indicating that it served as an oxidative stress induced cell culture model. In this induced model, the phytochemicals afforded antioxidant protection to the cells and reduced the MDA levels that were closer to the normal control fibroblast cells. The levels of the biochemical marker of lipid peroxidation, MDA, was decreased by 6.54%, 7.17%, 12.32%, respectively in the H_2O_2-induced cells, treated by CT, EGCG, and their combination, respectively. It was interesting to note that the combination of CT and EGCG markedly diminished the oxidative stress nearly by a double-fold when compared with individual phytochemical treatments. This clearly indicated the anti-oxidant enhancer effect of CT on EGCG (Figure 2C).

10.3389/fpls.2023.1131173

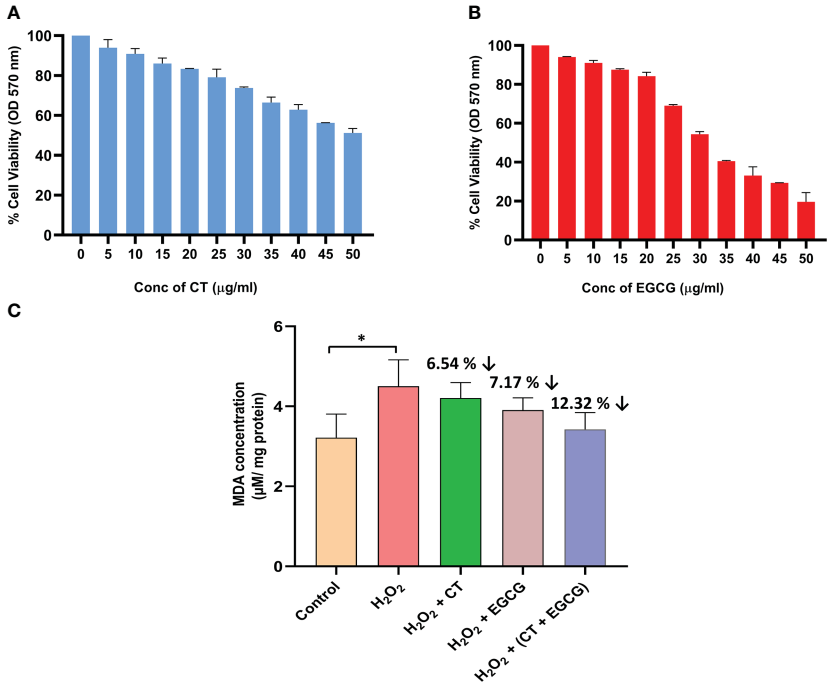


FIGURE 2
(A, B) represent the dose dependent changes in cell viability of normal fibroblast cells exposed to different concentrations of CT and EGCG, respectively. Values denote mean ± SD of three experiments performed in triplicate. **(C)** shows the reduction in oxidative stress (lipid peroxidation) in the H2O2 induced fibroblast cells treated with phytochemicals (20 µg/ml of both CT and EGCG). Values denote mean ± SD of two experiments done in triplicate. The oxidative stress mimic (H2O2 group) and the treated groups were compared with the normal control group where statistically significant difference was expressed at $P \leq 0.05$ (denoted by *) (GraphPad Prism v8.0.2).

LC-MS analysis of the enriched and purified fraction C5 obtained from *C. tetragonoloba*

The fractions eluted out of Sephadex LH-20 column were analysed by TLC on Silica gel F254 plates (Supplementary Material, Figure S2). One fraction, i.e. C5, was chosen for mass spectrometric analysis. LC-MS analysis was carried out in order to identify the phytochemicals in CT seed extract with a possible role in interaction with the standard phenolic phytochemicals. In order to retain maximum structural information, a constant value of collision energy was given to each compound for obtaining mass spectra with different fragmentation patterns. The different m/z values were analysed using the RIKEN ReSPect database which is extensively designed to study the mass spectra of the plant based secondary metabolites.

The gradient flow method was chosen to separate the flavonoids in liquid chromatography. The significant peaks eluted out had the retention times of 3.98 min, 18.0 min, and 19.65 min respectively (Figure 3). These peaks were then subjected to ESI-MS full scan mode analyses in order to identify the protonated ions. The individual m/z spectra are given in Supplementary Material (Figure S3).

Major peak 1 eluted out at 3.98 minutes showed fragments with m/z of values 305.1310 indicating the presence of dihydroflavonol catechin hydrate. The second major peak, which was eluted out at 18.00 minutes showed the fragment with m/z observed at 481.472 indicated the presence of flavonols, Myricetin-3-Galactoside and

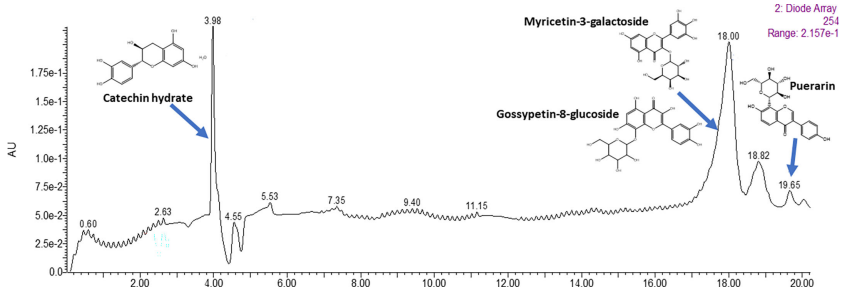


FIGURE 3
LC-MS chromatogram of purified fraction (C5) of *C. tetragonoloba* seeds obtained after Sephadex LH-20 chromatography.

Gossypetin-8-glucoside. The third major peak at 19.65 minutes indicated the presence of Puerarin which is daidzein 8-C-glucoside belonging to isoflavonoid class and was identified *via* the m/z fragments of 268.0568 and 297.0857. The retention time, chemical formula and mass of individual compounds are given in the Supplementary Material (Table S1).

In the present work, it has been seen that in comparison to the extract alone, both the *C. tetragonoloba* seed extract and kaempferol showed high fold increase in the antioxidant activity when combined at lower concentrations (Figure 1). This could be due to the presence of catechin and myricetin in the seed extract. We assume that the aglycone version of the identified compounds could be involved in synergistic interactions due to available and reactive hydroxyl groups. Many flavonoids are present as aglycones and convert to glycosides as the fruit matures on the plant. Again, upon consumption, the glycosides in food are cleaved into aglycones which are often more bioactive than the glycosylated versions. However, antioxidant synergy with flavonoid glycosides, such as quercetin-3-glucoside, has also been reported (Hidalgo et al., 2010). The authors mentioned that kaempferol showed synergistic interaction with catechin and myricetin.

Here, the chromatogram obtained was according to the polarity of the compounds eventually detected by MS analysis. The analysis showed that the methanolic extract contains only one isoflavonoid viz. puerarin, which was present in its glycosylated form. At relatively less abundance, as observed from the MS count ion, of Vitexin, which is a flavone, was also detected in the extract. The neutral ion losses from the different compounds indicated that the phytochemical constituents in the pods were present in glycosylated as were their aglycone forms. Kobeasy and El-salam (2011) investigated the major flavonoids present in the *C. tetragonoloba* collected from the Egypt region and showed the presence of luteolin and quercetin in aqueous extracts of seeds. Additionally, Morris and Wang (2017) explored the potential of *C. tetragonoloba L.* beans as a good source of kaempferol and quercetin in different cultivars grown in Georgia, USA under laboratory conditions.

To summarize, in the current study, a novel attempt at using *C. tetragonoloba* seed extract to enhance antioxidant activity of polyphenolic phytochemicals (of dietary importance) was made using both cell-free and cell culture systems, followed by purification of phenolics/flavonoids from the seeds (along with detailed phytochemical characterization using LC-MS technique).

A representative graphical abstract of the overall work done is given in Supplementary Material (Figure S1).

Conclusion

The current study sought to investigate the antioxidant activity enhancer effect of an edible desert legume, *Cyamopsis tetragonoloba*. Although the seeds of this plant have been used as a commercial source of guar gum, they remain relatively underutilized as sources of nutraceuticals. Furthermore, negligible work has been done towards sourcing antioxidant activity enhancers (AAE) from such edible legumes. The post-COVID era has witnessed a boost in the global nutraceutical industry. Consumers have developed preference towards nutraceuticals and dietary supplements obtained from plants due to lower toxicity. On the basis of the results of this study, it could be concluded that phenolic compounds present in *C. tetragonoloba* seed extract can interact with other compounds (standard phytochemicals) and act as antioxidant activity enhancers. These interactions can be synergistic, additive, or antagonistic, based on various characteristics, like chemical structures, availability of hydrogen ions, type of antioxidant assay used, concentrations, and combination ratios. The results obtained support our hypothesis of edible legumes as a host for a variety of natural antioxidant activity enhancers. Furthermore, the use of legumes as food ingredients and nutraceuticals is extremely promising for developing functional foods with positive health effects, often attributed to the antioxidant potential. The use of edible legume plants growing in the wild in Indian (semi) arid regions in formulating these nutraceuticals can be beneficial from both economic and environmental aspects as these plants are capable of growing on marginal and less fertile lands, and do not need heavy application of water or fertilizers. Being edible and safe, the seed extract of *C. tetragonoloba* can be used in food industry as an antioxidant activity enhancer. Those plant cultivars which are less useful as sources of guar gum (yield- or quality-wise), or even degummed seeds or guar gum industrial waste, could be attractive candidates for the same. The current study would potentially pave the way for more such research towards desert plants as sources of antioxidants/antioxidant activity enhancers, as well as mechanistic elucidation of biological activities of specific 'plant extract-phytochemical' combinations.

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

PD and PS contributed to conception and design of the study. TJ, SM, SP and VA performed the experiments. TJ wrote the first draft of the manuscript. SM, SP and VA wrote sections of the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

Authors are grateful to the administration of Birla Institute of Technology and Science (BITS), Pilani—Pilani Campus (Rajasthan, India) for infrastructural support and encouragement. Tripti Joshi is grateful to University Grants Commission, New Delhi, India, for Senior Research Fellowship. Sumit Kumar Mandal and Sonakshi Puri are grateful to BITS, Pilani for providing Institute Fellowship. Authors are thankful to SAIF (Sophisticated Analytical Instrumentation Facility), Panjab University, Chandigarh (India) for LC–MS data.

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.2023.1131173/full#supplementary-material

References

Adki, K., Laddha, A., Gaikwad, A., and Kulkarni, Y. (2020), 365–391. doi: 10.1016/B978-0-12-818553-7.00026-7

Ahmad, A., Ullah, F., Sadiq, A., Ayaz, M., Rahim, H., Rashid, U., et al. (2019). Pharmacological evaluation of aldehydic-pyrrolidinedione against HCT-116, MDA-MB231, NIH/3T3, MCF-7 cancer cell lines, antioxidant and enzyme inhibition studies. *Drug Des. Devel. Ther.* 13, 4185. doi: 10.2147/DDDT.S226080

Amarowicz, R. (2020). Legume seeds as an important component of human diet. *Foods (Basel Switzerland)* 9, 1812. doi: 10.3390/foods9121812

Asati, V., Deepa, P. R., and Sharma, P. K. (2022). Desert legume *Prosopis cineraria* as a novel source of antioxidant flavonoids/isoflavonoids: Biochemical characterization of edible pods for potential functional food development. *Biochem. Biophys. Rep.* 29, 101210. doi: 10.1016/j.bbrep.2022.101210

Ayilara, M. S., Abberton, M., Oyatomi, O. A., Babalola, O. O., and Odeyemi, O. (2022). Potentials of underutilized legumes in food security. *Front. Soil Sci.* 60. doi: 10.3389/fsoil.2022.1020193

Braicu, C., Pilecki, V., Balacescu, O., Irimie, A., and Neagoe, I. B. (2011). The relationships between biological activities and structure of flavan-3-ols. *Int. J. Mol. Sci.* 12, 9342–9353. doi: 10.3390/ijms12129342

Chaudhary, S., Dhanker, R., Kumar, R., and Goyal, S. (2022). Importance of legumes and role of sulphur oxidizing bacteria for their production: A review. *Legume Res. Int. J.* 45, 275–284. doi: 10.18805/LR-4415

Chaudhary, A., Sharma, S., Mittal, A., Gupta, S., and Dua, A. (2020). Phytochemical and antioxidant profiling of Ocimum sanctum. *J. Food Sci. Technol.* 57, 3852–3863. doi: 10.1007/s13197-020-04417-2

Chen, C.-Y., Milbury, P. E., Lapsley, K., and Blumberg, J. B. (2005). Flavonoids from almond skins are bioavailable and act synergistically with vitamins c and e to enhance hamster and human LDL resistance to oxidation. *J. Nutr.* 135, 1366–1373. doi: 10.1093/jn/135.6.1366

Elhamouly, N. A., Hewedy, O. A., Zaitoon, A., Miraples, A., Elshorbagy, O. T., Hussien, S., et al. (2022). The hidden power of secondary metabolites in plant-fungi interactions and sustainable phytoremediation. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1044896

Freeman, B. L., Eggett, D. L., and Parker, T. L. (2010). Synergistic and antagonistic interactions of phenolic compounds found in navel oranges. *J. Food Sci.* 75, C570–C576. doi: 10.1111/j.1750-3841.2010.01717.x

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

Heim, K. E., Tagliaferro, A. R., and Bobilya, D. J. (2002). Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships. *J. Nutr. Biochem.* 13, 572–584. doi: 10.1016/s0955-2863(02)00208-5

Hidalgo, M., Sánchez-Moreno, C., and de Pascual-Teresa, S. (2010). Flavonoid-flavonoid interaction and its effect on their antioxidant activity. *Food Chem.* 121, 691–696. doi: 10.1016/j.foodchem.2009.12.097

Hirano, R., Sasamoto, W., Matsumoto, A., Itakura, H., Igarashi, O., and Kondo, K. (2001). Antioxidant ability of various flavonoids against DPPH radicals and LDL oxidation. *J. Nutr. Sci. Vitaminol. (Tokyo).* 47, 357–362. doi: 10.3177/jnsv.47.357

Jain, P. K., and Rijhwani, S. (2018). Comparative gc-ms analysis of cyamopsis tetragonoloba fruit extracts. *Int. J. Pharm. Sci. Res.* 9, 4236–4242. doi: 10.13040/IJPSR.0975-8232.9(10).4236-42

Jan, R., Asaf, S., Numan, M., Lubna,, and Kim, K.-M. (2021). Plant secondary metabolite biosynthesis and transcriptional regulation in response to biotic and abiotic stress conditions. *Agronomy* 11, 968. doi: 10.3390/agronomy11050968

Jat, B. L., Sharma, H. C., Pagaria, P., Meena, A. K., Mali, G. R., and Khan, T. (2023). Legumes: Source of bioactive compounds and their potential use in legume crops improvement: A review. *Legume Res. Int. J.* 1, 8. doi: 10.18805/LR-4945

John, B., George, S., and Reddy, V. R. K. (2014). Total phenolics and flavonoids in selected medicinal plants from kerala. *Int. J. Pharm. Pharm. Sci.* 6, 406–408.

Joshi, T., Deepa, P. R., and Sharma, P. K. (2022). Effect of different proportions of phenolics on antioxidant potential: Pointers for bioactive Synergy/Antagonism in foods and nutraceuticals. *Proc. Natl. Acad. Sci. India. Sect. B.* 92, 939–946. doi: 10.1007/s40011-022-01396-6

Kaushik, S., Kaushik, S., Kumar, R., Dar, L., and Yadav, J. P. (2020). *In-vitro* and in silico activity of cyamopsis tetragonoloba (Gaur) l. supercritical extract against the dengue-2 virus. *Virusdisease* 31, 470–478. doi: 10.1007/s13337-020-00624-9

Kerwin, S. M. (2004). Soy saponins and the anticancer effects of soybeans and soy-based foods. *Curr. Med. Chem. Anticancer. Agents* 4, 263–272. doi: 10.2174/1568011043352993

Kobeasy, I., and El-salam, S. M. A. (2011). Biochemical studies on plantago major l. and cyamopsis tetragonoloba l. *Int. J. Biodivers. Conserv.* 3, 83–91. doi: 10.5897/IJBC.9000014

Kuravadi, N., Verma, S., Pareek, S., Gahlot, P., Kumari, S., Tanwar, U., et al. (2012). "Guar: An industrial crop from marginal farms," in *Agricultural sustainability*. Eds. G. S. Bhullar and N. K. Bhullar (Zurich, SW: Academic Press – Elsevier), 47–60. doi: 10.1016/B978-0-12-404560-6.00003-4

Lin, D., Xiao, M., Zhao, J., Li, Z., Xing, B., Li, X., et al. (2016). An overview of plant phenolic compounds and their importance in human nutrition and management of type 2 diabetes. *Molecules* 21, 1374. doi: 10.3390/molecules21101374

Mandal, S. K., Kumar, B. K., Sharma, P. K., Murugesan, S., and Deepa, P. R. (2022). In silico and *in vitro* analysis of PPAR - α/γ dual agonists: Comparative evaluation of potential phytochemicals with anti-obesity drug orlistat. *Comput. Biol. Med.* 147, 105796. doi: 10.1016/j.compbiomed.2022.105796

Marinova, E., Toneva, A., and Yanishlieva, N. (2008). Synergistic antioxidant effect of α-tocopherol and myricetin on the autoxidation of triacylglycerols of sunflower oil. *Food Chem.* 106, 628–633. doi: 10.1016/j.foodchem.2007.06.022

Martín-Cabrejas, M. (2018), 1–18. doi: 10.1039/9781788015721-00001

Mbagwu, F. N., Okafor, V. U., and Ekeanyanwu, J. (2011). Phytochemical screening on four edible legumes (Vigna subterranean, glycine max, arachis hypogea, and vigna uniguiculata) found in eastern Nigeria. *Afr. J. Plant Sci.* 5, 370–372. doi: 10.5897/AJPS.9000189

Morris, J. B., and Wang, M. L. (2017). Functional vegetable guar (Cyamopsis tetragonoloba l. taub.) accessions for improving flavonoid concentrations in immature pods. *J. Diet. Suppl.* 14, 146–157. doi: 10.1080/19390211.2016.1207002

Moteriya, P. (2015). *In vitro* free radical scavenging and antimicrobial activity of cyamopsis tetragonoloba l. *J. Pharmacogn. Phytochem.* 4, 102–106.

Mukhtar, H., Ansari, S. H., Naved, T., and Bhat, Z. (2004). Hypoglycemic activity of psidium guajava linn. leaf extract. *J. Nat. Remedies* 4, 186–189. doi: 10.18311/jnr/2004/186

Olszowy, M. (2020). Synergistic, antagonistic and additive antioxidant effects in the binary mixtures. *Phytochem. Rev.* 19, 63–103. doi: 10.1007/s11101-019-09658-4

Pande, S., and Srinivasan, K. (2013a). Potentiation of antioxidant effect of dietary tender cluster beans (Cyamopsis tetragonoloba) by garlic (Allium sativum) in high-cholesterol-fed rats. *Can. J. Physiol. Pharmacol.* 91, 818–822. doi: 10.1139/cjpp-2013-0082

Pande, S., and Srinivasan, K. (2013b). Protective effect of dietary tender cluster beans (Cyamopsis tetragonoloba) in the gastrointestinal tract of experimental rats. *Appl. Physiol. Nutr. Metab.* 38, 169–176. doi: 10.1139/apnm-2012-0091

Pankaj, V. T., and Dhankar, A. (2023). A review: Promising forage crops grown in India and their quality importance. *Pharma Innov. J.* 12 (1), 2238–2244.

Pham, H. H., and Luan, N. (2021). doi: 10.5772/intechopen.100171

Rakariyatham, K., Zhou, D., Lu, T., Yin, F., Yu, Z., Li, D., et al. (2021). Synergistic effects of longan (Dimocarpus longan) peel extracts and food additives on oxidative stability of tuna oil. *Lwt* 152, 112275. doi: 10.1016/j.lwt.2021.112275

Riaz, S., Hussain, I., Ibrahim, M., Ishtiaq, M., Ali, Q., Muazzam Ali, M., et al. (2022). Extraction and optimization of active metabolites from cluster bean: An *In vitro* biological and phytochemical investigation. *Dose. Response.* 20, 15593258221098992. doi: 10.1177/15593258221098992

Romano, C. S., Abadi, K., Repetto, V., Vojnov, A. A., and Moreno, S. (2009). Synergistic antioxidant and antibacterial activity of rosemary plus butylated derivatives. *Food Chem.* 115, 456–461. doi: 10.1016/j.foodchem.2008.12.029

Rudrapal, M., Khairnar, S. J., Khan, J., Dukhyil, A., Ansari, M. A., Alomary, M. N., et al. (2022). Dietary polyphenols and their role in oxidative stress-induced human diseases: insights into protective effects, antioxidant potentials and mechanism (s) of action. *Front. Pharmacol.* 13, 283. doi: 10.3389/fphar.2022.806470

Sawada, Y., Nakabayashi, R., Yamada, Y., Suzuki, M., Sato, M., Sakata, A., et al. (2012). RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. *Phytochemistry* 82, 38–45. doi: 10.1016/j.phytochem.2012.07.007

Sharma, A., Joshi, N., Kumar, R. A., Agrawal, P. K., and Prasad, S. (2017). High performance liquid chromatographic analysis of phenolic compounds and their antioxidant properties from different cultivars of cyamopsis tetragonaloba (L.) taub. *Microchem. J.* 133, 622–628. doi: 10.1016/j.microc.2017.04.020

Shi, J., Qu, Q., Kakuda, Y., Xue, S., Jiang, Y., Koide, S., et al. (2007). Investigation of the antioxidant and synergistic activity of lycopene and other natural antioxidants using LAME and AMVN model systems. *J. Food Compos. Anal.* 20, 603–608. doi: 10.1016/j.jfca.2007.03.004

Singh, J., and Gaikwad, D.. (2020). "Phytogenic feed additives in animal nutrition." in *Natural bioactive products in sustainable agriculture*. Eds. J. Singh and A. Yadav (Singapore: Springer), 273–289. doi: 10.1007/978-981-15-3024-1_13

Slinkard, K., and Singleton, V. L. (1977). Total phenol analysis: Automation and comparison with manual methods. *Am. J. Enol. Vitic.* 28, 49LP–4955. doi: 10.5344/ajev.1974.28.1.49

Song, T., Barua, K., Buseman, G., and Murphy, P. A. (1998). Soy isoflavone analysis: quality control and a new internal standard. *Am. J. Clin. Nutr.* 68, 1474S–1479S. doi: 10.1093/ajcn/68.6.1474S

Timoracká, M., Šnirc, M., Musilová, J., and Čičová, I. (2022). The variability of the total polyphenols content and the antioxidant activity in the varieties of selected legumes. *J. Microbiol. Biotechnol. Food Sci.* 12, 1–5. doi: 10.55251/jmbfs.9399

Tsao, R., Yang, R., Young, J. C., and Zhu, H. (2003). Polyphenolic profiles in eight apple cultivars using high-performance liquid chromatography (HPLC). *J. Agric. Food Chem.* 51, 6347–6353. doi: 10.1021/jf0346298

Venuprasad, M. P., Kandikattu, H. K., Razack, S., and Khanum, F. (2014). Phytochemical analysis of Ocimum gratissimum by LC-ESI–MS/MS and its antioxidant and anxiolytic effects. *South African J. Bot.* 92, 151–158. doi: 10.1016/j.sajb.2014.02.010

Vuorinen-Markkola, H., Sinisalo, M., and Koivisto, V. A. (1992). Guar gum in insulin-dependent diabetes: Effects on glycemic control and serum lipoproteins. *Am. J. Clin. Nutr.* 56, 1056–1060. doi: 10.1093/ajcn/56.6.1056

Wang, S., Meckling, K. A., Marcone, M. F., Kakuda, Y., and Tsao, R. (2011). Synergistic, additive, and antagonistic effects of food mixtures on total antioxidant capacities. *J. Agric. Food Chem.* 59, 960–968. doi: 10.1021/jf1040977

Wenz, C., Faust, D., Linz, B., Turmann, C., Nikolova, T., and Dietrich, C. (2019). Cell–cell contacts protect against t-BuOOH-induced cellular damage and ferroptosis *in vitro*. *Arch. Toxicol.* 93, 1265–1279. doi: 10.1007/s00204-019-02413-w

Wright, J. S., Johnson, E. R., and DiLabio, G. A. (2001). Predicting the activity of phenolic antioxidants: Theoretical method, analysis of substituent effects, and application to major families of antioxidants. *J. Am. Chem. Soc* 123, 1173–1183. doi: 10.1021/ja002455u

Zhang, D.-Y., Wan, Y., Xu, J.-Y., Wu, G.-H., Li, L., and Yao, X.-H. (2016). Ultrasound extraction of polysaccharides from mulberry leaves and their effect on enhancing antioxidant activity. *Carbohydr. Polym.* 137, 473–479. doi: 10.1016/j.carbpol.2015.11.016

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 April 2023
DOI 10.3389/fpls.2023.1136961

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Kanika Sharma,
Central Institute for Research on Cotton Technology (CIRCOT), India
Arif Jamal Siddiqui,
University of Hail, Saudi Arabia

*CORRESPONDENCE
Mohammed H. Alruhaili
malruhaili@kau.edu.sa
Samy Selim
sabdulsalam@ju.edu.sa
Hamada AbdElgawad
hamada.abdElgawad@uantwerpen.be

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

RECEIVED 03 January 2023
ACCEPTED 23 March 2023
PUBLISHED 20 April 2023

CITATION
Alruhaili MH, Almuhayawi MS, Gattan HS, Alharbi MT, Nagshabandi MK, Jaouni SKA, Selim S and AbdElgawad H (2023) Insight into the phytochemical profile and antimicrobial activities of *Amomum subulatum* and *Amomum xanthioides*: an *in vitro* and *in silico* study. *Front. Plant Sci.* 14:1136961. doi: 10.3389/fpls.2023.1136961

Insight into the phytochemical profile and antimicrobial activities of *Amomum subulatum* and *Amomum xanthioides*: an *in vitro* and *in silico* study

Mohammed H. Alruhaili[1,2]*, Mohammed S. Almuhayawi[1], Hattan S. Gattan[2,3], Mohanned Talal Alharbi[4], Mohammed K. Nagshabandi[4], Soad K. Al Jaouni[5], Samy Selim[6]* and Hamada AbdElgawad[7]*

[1]Department of Clinical Microbiology and Immunology Faculty of Medicine, King AbdulAziz University, Jeddah, Saudi Arabia, [2]Special Infectious Agents Unit, King Fahad Medical Research Center, King AbdulAziz University, Jeddah, Saudi Arabia, [3]Department of Medical Laboratory Sciences, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, [4]Department of Medical Microbiology and Parasitology, Faculty of Medicine, University of Jeddah, Jeddah, Saudi Arabia, [5]Department of Hematology/Oncology, Yousef Abdulatif Jameel Scientific Chair of Prophetic Medicine Application, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia, [6]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Jouf University, Sakaka, Saudi Arabia, [7]Department of Botany and Microbiology, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt

Introduction: Medicinal plants have been considered as potential source of therapeutics or as starting materials in drugs formulation.

Methods: The current study aims to shed light on the therapeutic potential of the *Amomum subulatom* and *Amomum xanthioides* Fruits by analyzing the phytochemical composition of their seeds and fruits using gas chromatography-mass spectrometry (GC-MS) and high-performance liquid chromatography (HPLC) techniques to determine the presence of bioactive components such as flavonoids, phenols, vitamins, steroids, and essential oils.

Results and Discussion: The protein content is usually higher than the total lipids in both species except the fruit of *A. subulatum* which contain more lipids than proteins. The total protein contents for *A. subulatum* were 235.03 ± 21.49 and 227.49 ± 25.82 mg/g dry weight while for *A. xanthioides* were 201.9 ± 37.79 and 294.99 ± 37.93 mg/g dry weight for seeds and fruit, respectively. The Carvacrol levels in *A. subulatum* is 20 times higher than that in *A. xanthioides*. Lower levels of α-Thujene, Phyllanderenes, Ascaridole, and Pinocarvone were also observed in both species. According to DPPH (2,2-diphenylpicrylhydrazyl) assay, seed the extract of *A. subulatum* exhibited the highest antioxidant activity (78.26±9.27 %) followed by the seed extract of *A. xanthioides* (68.21±2.56 %). Similarly, FRAP (Ferric Reducing Antioxidant Power) assay showed that the highest antioxidant activity was exhibited by the seed extract of the two species; 20.14±1.11 and 21.18±1.04 µmol trolox g–1 DW for *A. subulatum* and *A. xanthioides*, respectively. In terms of anti-lipid peroxidation, relatively higher values were obtained for the fruit extract of *A. subulatum* (6.08±0.35) and the seed extract of *A. xanthioides* (6.11±0.55).

Ethanolic seed extracts of *A. subulatum* had the highest efficiency against four Gram-negative bacterial species which causes serious human diseases, namely *Pseudomonas aeruginosa*, *Proteus vulgaris*, *Enterobacter aerogenes*, and *Salmonella typhimurium*. In addition, *P. aeruginosa* was also inhibited by the fruit extract of both *A. subulatum* and *A. xanthioides*. For the seed extract of *A. xanthioides*, large inhibition zones were formed against *P. vulgaris* and the fungus *Candida albicans*. Finally, we have *in silico* explored the mode of action of these plants by performing detailed molecular modeling studies and showed that the antimicrobial activities of these plants could be attributed to the high binding affinity of their bioactive compounds to bind to the active sites of the sterol 14-alpha demethylase and the transcriptional regulator MvfR.

Conclusion: These findings demonstrate the two species extracts possess high biological activities and therapeutical values, which increases their potential value in a number of therapeutic applications.

KEYWORDS

phytochemical analysis, natural products, medicinal plants, *Amomum subulatom*, *Amomum xanthioides*, antioxidant activity, antimicrobial activities, molecular docking

1 Introduction

Historically, medicinal plants were used worldwide either directly as therapeutics or as starting materials in drug formulation (Grover et al., 2002; Bauer and Brönstrup, 2014; Pešić and Stanković, 2015). The World Health Organization (WHO) evaluates that approximately 80% of the people worldwide now depend on medicinal plants for basic medical requirements (WHO establishes the Global Centre for Traditional Medicine in India (n.d.)). Medicinal plants are used to treat numerous illnesses, including diabetes, cardiovascular diseases, nervous system disorders, asthma, hypertension, and cancer (Prasathkumar et al., 2021). Natural antimicrobials found in medicinal plants are highly effective against newly emerging microbial strains (Cowan, 1999; Tijjani et al., 2011). Due to the inappropriate and excessive use of antimicrobial medications, microbes have become resistant to many antibiotics, which poses a significant challenge for the treatment of infectious illnesses (Davies, 1994; McGregor et al., 2014; Saleem et al., 2019; Thompson, 2022). Antibiotic resistance occurs due to the constant generation of resistant strains to drugs and the adaptation of microorganisms to frequently used antibiotics (Ventola, 2015; Aslam et al., 2018; Saleem et al., 2019). Lately, the interest in medicinal plants has increased because of the great potential of plant-based medicines. Consequently, one of the main sources of commercial pharmaceuticals is still the useful compounds extracted from medicinal plants. In addition to being medicinal, numerous plant extracts have been widely employed in fragrances, food flavoring, and food preservation (Pandey et al., 2017; Delesa, 2018).

Medicinal plants are distributed and considered a valuable source of novel medications worldwide (Balunas and Kinghorn, 2005; Rafieian-Kopaei, 2012). The usage of medicinal plants is expanding quickly globally due to the rising need for herbal medicines, healthcare products, and plant secondary metabolites. Therefore, isolation and purification of extracts from medicinal plants are important for discovering and developing new drugs. *A. subulatum*, generally recognized as black cardamom, is a herbaceous plant used as a medicinal spice in India (Gautam et al., 2016). It has historically been used to treat vomiting, abdominal pain, gastrointestinal (GI) infections, and rectal diseases (Kumar et al., 2015; Alam and Singh, 2021; Dhakal et al., 2022). The therapeutic capabilities of this medicinal herb have gained great attention in the past years since it is found to be a source of antimicrobials, antioxidants, anti-inflammatory, and cardio-protective compounds (Bhaswant et al., 2015; Gautam et al., 2016). The essential oil components of *A. subulatum* are primarily responsible for its therapeutic capability (Dhakal et al., 2022). Essential oils from *A. subulatum* seeds are found to have an antimicrobial effect against various pathogens including *Staphylococcus aureus*, *Escherichia coli*, *Bacillus pumilus*, *Pseudomonas aeruginosa*, and *Aspergillus niger* (Agnihotri and Wakode, 2010; Satyal et al., 2012). Hexane seed extract of *A. subulatum* exhibited high cellular toxicity on HeLa and MCF-7 cell lines, which are human cancer cell lines, indicating its anticancer potential, while the ethyl acetate extract possessed a considerable antioxidant activity based on the free radical scavenging assay (Sharma et al., 2017). In addition, fruit extracts of *A. subulatum* have shown high anti-inflammatory effects in rats with carrageenan-induced paw edema when compared with diclofenac, the main drug for this disease (Alam et al., 2011). The fruit extracts also exhibited anticancer activity in mice through the control of cytokines that cause inflammation as well as the NF-κB signaling (Sudarsanan et al., 2021). A recent study discovered that the dichloromethane extract of *A. subulatum* had an apoptotic effect against lung cancer cells (Makhija et al., 2022).

Amomum xanthioides is a vigorous herb that is used for food and medicinal purposes in southern China, India, Thailand, Vietnam, Laos, and Cambodia (Lamxay and Newman, 2012).

Ethyl acetate extracts from *A. xanthioides* seeds are found to have therapeutic potential against liver fibrosis in rat models (Kim et al., 2015). When administered to high-fat-diet mice, ethyl acetate extract from *A. xanthioides* showed an anti-fatty liver effect, indicating its therapeutic potential in managing non-alcoholic fatty liver disorders (Im et al., 2020). Moreover, anti-inflammatory effects are associated with *A. xanthioides* extracts in atopic dermatitis, a chronic relapsing skin inflammation (Choi et al., 2017). Essential oils extracted from *A. xanthioides* fruits were found to have growth inhibitory effects on *Enterococcus faecalis*, *Bacillus cereus*, *P. aeruginosa*, and *S. aureus*, with minimum inhibitory concentration values between 100 and 200 g/ml (Thinh et al., 2022).

In spite of all the previous data, *A. subulatum* and *A. xanthioides* are not being given much consideration for therapeutic development as a result of a lack of comprehensive chemical analysis and pharmacological research (Thinh et al., 2021; Dhakal et al., 2022). In this manner, the current study aimed to provide insight into the phytochemical composition of both species by investigating the bioactive components in their seeds and fruits, including flavonoids, phenols, vitamins, steroids, and essential oils. To this end, techniques such as gas chromatography–mass spectrometry (GC-MS) and high-performance liquid chromatography (HPLC) were employed. In addition, the antimicrobial and antioxidant potentials of the ethanol extracts were assessed through multiple assays. Finally, computational docking was performed to investigate the molecular interactions of the proposed compounds against certain microbial proteins determined from prior analysis to confirm the antimicrobial activity observed from these compounds.

2 Materials and methods

2.1 Plant sample collection and identification

The seeds and fruits of *A. subulatum* and *A. xanthioides* were obtained at a local Pakistani store in Antwerp, Belgium. The plant specimen was identified and authenticated by a plant taxonomist at the Botany Department of Beni Suef University, Egypt.

2.2 Total nutrients

Sugars were extracted in 0.2 g dry weight of *Amomum* seeds and fruit samples in 2 ml of boiled distilled water for 60 min at 100°C. After being cooled, the extract was centrifuged at 6,000 ×*g* for 15 min. The pellet was re-extracted using 2 ml of boiled distilled water. After centrifugation, the two supernatants were combined for further analysis. The concentration of total sugars in the supernatant was assessed following Nelson's method as defined by Clark and Switzer (1977). Sugar extract was added to freshly prepared Nelson's alkaline copper reagent (1:1 v/v), and the mixture was boiled for 20 min. Subsequently, 1 ml of arseno-molybdate reagent was added to the reaction mixture with shaking to dissolve Cu_2O. When the effervescence stopped, the change in color intensity was measured at 540 nm. The reduced sugar content was determined by means of a glucose standard curve. The protein concentration was assessed for *Amomum* seeds and fruits (0.2 g DW) extracted twice in NaOH (0.4% w/v) at room temperature. Extracts were shaken by using an orbital shaker at 220 rpm for 45 min. The total protein content was measured by using Lowry's method at 660 nm (Oh et al., 1951). Bovine serum albumin was used as a standard reagent. Total lipid concentrations were determined, where *Amomum* seeds and fruit samples were homogenized twice in a 2:1 mixture of chloroform/methanol (v/v). Then plants were centrifuged for 15 min at 3,000 ×*g*. In a 4:1 ratio of toluene:ethanol (v/v), the pellets were re-dissolved. The total lipid content was determined after concentration. Gravimetric analysis was used to determine the extracted lipids, which were represented as weight (g) per fresh weight (g) of the plant. To eliminate unwanted protein and starch, crude fibers were gelatinized by a heat-stable alpha-amylase at pH 6 and 100°C for 25 min and then enzymatically digested by a combination of protease (pH 7.5, 60°C, 25 min) and amyloglucosidase (pH 6, 0°C, 30 min). Fibers were allowed to precipitate in ethanol for washing, and the residues were weighed after washing.

2.3 Mineral quantification

The recognition of mineral elements was achieved according to AbdElgawad et al. (2014), where 200 mg dry weight of *Amomum* seeds and fruit samples was digested in 5:1 (v:v) HNO_3/H_2O solution for 30 min. Thereafter, macroelements and microelements were estimated (inductively coupled plasma–mass spectrometry (ICP-MS), Finnigan Element XR, and Scientific, Bremen, Germany). Nitric acid (1%) was used as blank.

2.4 Total phenolics and flavonoids

Total polyphenols and flavonoids were obtained by homogenizing 200 mg dry weight of *Amomum* seeds and fruit samples in 2 ml of 80% ethanol (v/v). The contents of phenolic and flavonoid were assessed using the Folin–Ciocalteu and aluminum chloride colorimetric assays, respectively (Saleh and Madany, 2015), with gallic acid and quercetin as standards, respectively (Sigma-Aldrich Co., St. Louis, MO, USA). For the determination of total phenolics, 1 ml of the phenolic extract was mixed with 1 ml of 10% Folin–Ciocalteu phenol reagent and 1 ml of 20% anhydrous Na_2CO_3 and afterward filled to a given volume with distilled water. The absorbance of the resulting blue color was observed after 30 min at 650 nm against a water-reagent blank (samples extracted and replaced by distilled water). The total phenolic content was determined from a catechol (Sigma-Aldrich Co., St. Louis, MO, USA) standard curve and reported as mg gallic/g dry weight. Total flavonoid content was determined by mixing 0.25 ml of the extract with 1.25 ml of distilled water in a test tube, followed by the addition of 75 µl of 5% (w/v) Na nitrite solution. After 6 min, 150 µl of 10% (w/v) $AlCl_3$ was added, and the mixture was left for another 5 min prior to adding 0.5 ml of 1 M of sodium hydroxide. The solution

was filled and mixed with distilled water up to 2.5 ml, and the absorbance was recorded at 510 nm. The total flavonoid content was measured using a quercetin standard curve and reported as mg quercetin/g dry weight.

2.5 Total alkaloid estimation

The total alkaloid was estimated following the protocol described by Sreevidya et al (Sreevidya and Mehrotra, 2003). along with the bismuth nitrate pentahydrate ($Bi(NO_3)_3{\cdot}5H_2O$) calibration curve. Samples were extracted in methanol-HCl at pH 2–2.5. The extract was mixed with Dragendorff's reagent (bismuth nitrate pentahydrate, glacial acetic acid, and 8.0 g of potassium iodide) and centrifuged for 10 min at 5,000 rpm, 25°C. The precipitate was further washed twice with methanol, and the residue was added to a disodium sulfide solution. The resulting brownish-black precipitate was afterward centrifuged for 10 min at 5,000 rpm. The residue was dissolved in concentrated nitric acid by mild warming. This solution was diluted with distilled water and mixed with a 3% thiourea solution. At 435 nm, the absorbance was recorded in comparison to a blank containing HNO_3 and thiourea by using a spectrophotometer (Perkin Elmer Lambda 25).

2.6 Saponins

The extraction and quantification of saponin in *Amomum* seeds and fruit samples were performed (Lai et al., 2013). Ground dried seeds and fruits were extracted in petroleum ether and shaken for 4 h at 1,000 rpm and 25°C ± 3°C. The solvent was then removed using a vacuum rotary evaporator at 60°C. Saponin in residues was extracted in 80% aqueous methanol and shaken for another 4 h. The extract was filtered and preserved at 4°C in the dark. For quantification, the spectrophotometric technique was applied to estimate the total saponin content of the samples. In a cold-water bath (0°C), 0.1 ml of the extract attained above was added to 0.4 ml of methanol solution (80%), 0.5 ml of freshly prepared vanillin solution (8% (w/v; prepared in ethanol), and 5.0 ml of sulfuric acid (72%). After that, the mixture was put in a water bath at 60°C for 10 min and lastly cooled in ice-cold water. The absorbance was measured at 544 nm against a reagent blank with a UV–Vis spectrophotometer (Shimadzu UV-160A PC, Shimadzu Corporation, Kyoto, Japan). The reagent blank was prepared by applying the same technique, but the extract was exchanged with an equivalent volume of 80% methanol. The results were estimated from a standard curve plotted with various crude soya saponin concentrations (0, 1,000, 2,000, and 3,000 ppm) containing at least 80% saponin (Waki, Osaki, Japan) in 80% aqueous methanol and expressed as mg soya saponin/100 g sample.

2.7 Fatty acid profile

The fatty acid profile was quantified according to Hassan et al. (2018). To obtain lipophilic fraction, 0.2 g dry weight of *Amomum* seeds and fruit samples was extracted in chloroform/methanol (2:1, v/v) at 25°C in the presence of the internal standard tripentadecanoic acid triglyceride. Fatty acids were derivatized with 1% pentafluorobenzyl (PFB) bromide in acetonitrile at room temperature for 20 min. The fatty acid PFB esters dissolved in 50 µl of *iso*-octane are injected. The investigation was applied by means of an Agilent single-quadrupole mass spectrometer with an inert mass selective detector (MSD-5975C detector, Agilent Technologies, Santa Clara, CA, USA) coupled directly to an Agilent 7890A gas chromatograph that was equipped with a split–splitless injector, a quick-swap assembly, an Agilent model 7693 autosampler, and an HP-5MS fused silica capillary column (5% phenyl/95% dimethylpolysiloxane, 30 m × 0.25 mm i.d., film thickness 0.25 µm, Agilent Technologies, USA). The temperature of the oven was kept at 80°C for 2 min and then elevated up to 200°C at 5°C/min (1 min hold) and then to 280°C at 20°C/min (3 min hold). A 1.0-µl sample was injected by means of a split mode (split ratio, 1:10). At a flow rate of 1.5 ml/min, helium gas was employed as a carrier gas. For MS detection, an electron ionization technique with an ionization energy of 70 eV was applied. The temperatures of the injector and MS transfer line were set at 220°C and 290°C, respectively. The mass scan ranged from 50 to 550 *m/z* with an E_m voltage of 1,035 V. The quantitative examination of fatty acids is performed by comparing the target molecule's mass spectrometric ion signal to that of an equivalent standard. Fatty acid standard curves were obtained by serially diluting a standard mixture of unlabeled quantitative fatty acid standards at specified quantities. Each fatty acid is given in quantitative standard dilution sets in the 0.15–500-ng range. A standard curve is created by doing a linear regression model on the ratio of the quantitative standard and internal standard ion yields plotted *vs.* the quantitative standard absolute quantities. The fatty acid content of the sample is then determined from the standard curve using analyte/internal standard ion yield ratios. In parallel, software freely available for the deconvolution of fatty acids spectrum was applied; i.e., the metabolite libraries available for metabolite identification were obtained using NIST08, a generalized chemical library (http://chemdata.nist.gov/mass-spc/ms-search/), and those specifically for metabolites were obtained using the Golm Metabolome (http://gmd.mpimp-golm.mpg.de).

2.8 Essential oil analysis

Plant samples were placed in 1 L of distilled water and subjected to hydrodistillation for 3 h, using a Clevenger-type apparatus (1.5% yield). To remove water remnants from essential oil, anhydrous Na_2SO_4 was utilized. The essential oil was stored at +4°C until tested and analyzed.

2.9 Gas chromatography and gas chromatography–mass spectrometry

Gas chromatography–flame ionization detector (GC-FID) and GC-MS were used to determine the essential oil both quantitatively

and qualitatively. GC analyses were performed on a Varian (Les Ulis, France) Star 3400 Cx chromatograph fitted with a fused silica capillary DB-5MS (5% phenyl methylpolysiloxane; 30 m/0.25 mm; film thickness, 0.25 mm) column. Chromatographic conditions were 60° C to 260°C temperature increases with a gradient of 5°C/min and 15 min isothermal at 260°C. A second increase was applied, reaching 340°C at 40°C/min. The total time of the analysis was 57 min. Petroleum ether was used for dissolving the oils to avoid saturating the column. Injection of the sample was performed at a split mode ratio of 1:10. Helium (purity 99.999%) was utilized as the carrier gas at 1 ml/min. The injector was operated at 200°C. The mass spectrometer (Varian Saturn GC/MS/MS 4D) was set at an electron multiplier voltage between 1,400 and 1,500 V and an emission current of 10 mA. The transfer line's temperature was 170°C, whereas the trap's temperature was 150°C. A total of 40 to 650 atomic mass units was covered by the mass scanning. The components were recognized by means of Wiley 2001 library data (NIST 02 version 2.62) of the GC-MS system, literature data, and comparison of the components' Kovats indices (KIs) and mass spectra with those of standards. In order to calculate KI, alkanes (C5–C24) were employed as reference points. Every determination was made twice and then averaged.

2.10 Vitamins

Carotene and β-cryptoxanthin contents were obtained in acetone and examined by a reversed-phase HPLC conducted with a diode array detector (Sarungallo et al., 2015). To extract carotene and β-cryptoxanthin, seed samples were shaken in a MagNa Lyser (3 × 10 s, 6,000 rpm) using acetone as a solvent. After centrifugation, the sample was brought in the autosampler of the HPLC (Shimadzu SIL10-ADvp) and kept at 4°C. The separation of the carotenoids and β-cryptoxanthin was performed using a reversed-phase method with a low-pressure gradient and was performed on a silica-based C18 column (Waters Spherisorb 5-μm ODS1 4.6 × 250 mm). Solvent A (acetonitrile:methanol:water, 1:9:10) and solvent B (methanol:ethyl acetate, 68:32) act as the mobile phase. The detection of the carotene and β-cryptoxanthin was performed by a diode array detector (Shimadzu SPD-M10Avp) at a wavelength range of 446–470 nm and integrated *via* the software program (Shimadzu Lab Solutions Lite).

Phylloquinone was detected according to the methods of Jakob et al (Jakob and Elmadfa, 1996). A reversed-phase HPLC system and samples were separated on analytical column Gynkotek ODS Hypersil (250 · 4.6 mm i.d., 5 lm), a guard-column Gynkotek ODS Hypersil (20 · 4.6 mm i.d., 5 lm, Bischoff, Leonberg, Germany). The mobile phase contained 1 L of dichloromethane–methanol mixture (ratio 1:9). This solvent was then combined with 5 ml of a methanolic solution containing l.37 g of $ZnCl_2$, 0.41 g of CH_3COONa, and 0.30 g of CH_3COOH. Recognition was performed at 243-nm excitation and 430-nm emission. The concentrations were calculated using a linear regression curve from standard solutions.

Tocopherols were obtained with hexane by means of the MagNa Lyser. The dried extract (CentriVap concentrator, Labconco, KS, USA) was resuspended in hexane, and tocopherols were isolated and estimated by HPLC (Shimadzu, 's Hertogenbosch, The Netherlands) (normal phase conditions, Particil Pac 5 μm column material, length 250 mm, i.d. 4.6 mm). Dimethyl tocol (DMT) was used as an internal standard (5 ppm). Data were investigated with Shimadzu Class VP 6.14 software.

2.11 UHPLC–MS/MS determination of phenolics and flavonoids

For quantification of individual phenolic acids and flavonoids, 0.3 g of dried powdered seeds or fruits was extracted in 2 ml of 80% (v/v) ethanol in a water bath at 70°C for 0.5 h. After centrifugation (12,000 rpm for 30 min), the supernatant was concentrated by a rotary evaporator (IKA-WERKE-RV06ML; Staufen, Germany). The obtained residue was dissolved in HPLC-grade methanol (final concentration = 1,000 ppm). All solutions were filtered through a 0.45-μm membrane filter (Iwaki Glass) before analysis. An Acquity UPLC System from Waters (Milford, CT, USA) was used for the extract's chromatographic analysis. This system is supplied with a binary solvent supply system, degasser, autosampler, and column heater. Utilizing a 100 mm × 2.1 mm Acquity BEH C18 column from Waters with 1.7-m particle size, chromatographic separation was carried out. A Waters (Manchester, UK) Xevo TQD tandem quadrupole mass spectrometer was used for MS/MS detection together with an electrospray ionization interface (ESI) that worked in the negative ion mode. The capillary voltage was 4.5 kV, the source temperature was 120°C, the desolvation gas temperature was 400°C, and the nitrogen flow rates for the cone and desolvation gases were 30 and 600 L/h, respectively. Eluent A, ultrapure water containing 0.1% formic acid, and Eluent B, acetonitrile, were the components of the mobile phase. Samples measuring 2 μl were introduced at a flow rate of 0.2 ml/min with a linear gradient starting at 3% B and escalating to 100% B in 10 min. The presence of phenolic chemicals in the sample was confirmed using the multiple-reaction monitoring (MRM) mode, together with *m*/*z* transitions of the precursor ions and product ions.

3,5-Dichloro-4-hydroxybenzoic acid (Sigma-Aldrich Co., St. Louis, MO, USA) was utilized as an internal standard to take into consideration recovery losses and ionization efficiencies. The concentrations were calculated based on the relative response of the internal analyte in relation to this internal standard added. The amount of internal tracer is always chosen in relation to the internal concentration of the analyte, taking into account that the concentration difference does not exceed a factor of 100, the range in which linearity is guaranteed.

2.12 DPPH antioxidant assay

Seven samples (300 μl of methanolic extract) were mixed with 900 μl of DPPH solution (4×10^{-5} M). After incubation at 60 nm in the dark at 37°C, the absorbance was measured at 517 nm with a spectrophotometer UV–visible ($A_{(sample)}$). A blank solution was also measured at the same wavelength ($A_{(blank)}$). The free radical-scavenging activity of each solution was then calculated as percent inhibition according to the following equation: (radical scavenging activity (%) = $((A_{(blank)} - A_{(sample)})/A_{(blank)}) \times 100$).

2.13 FRAP assay

The ferric ion-reducing antioxidant power (FRAP) activity was determined in the alcoholic (80% ethanol) seed and fruit extracts, which showed the highest activity by FRAP methods (AbdElgawad et al., 2019). FRAP assay was performed by adding 20 µl of each extract to a micro-titer plate and filling up to 200 µl of freshly prepared pre-warmed FRAP reagent. The mixture was incubated for 30 min at 37°C. The absorbance was determined at 593 nm. The antioxidant capability of the extracts was measured using the Trolox calibration curve.

2.14 TBARS lipid peroxidation assay

The level of lipid peroxidation was assessed by the thiobarbituric acid reactive substance (TBARS) method, using an egg yolk homogenate as the lipid-rich medium (Ohkawa et al., 1979). The seed extract and egg homogenate (0.5 ml of 10% v/v) were mixed with 15 mM of ferrous sulfate (to induce lipid peroxidation), and after 0.5 h, 1.5 ml of 10% trichloroacetic acid (TCA) was added. The mixture was then added to a 1.5-ml solution of 0.67% TBA and then boiled for 0.5 h. The chromogen obtained was measured at 535 nm.

2.15 Antimicrobial activity

The antibacterial activities of the seed and fruit ethanolic extracts were verified using the disc diffusion method (bacterial suspension containing 10^6 CFU/ml of the bacterial test strain spread on Mueller–Hinton agar). Each extract was put on sterilized filter paper discs (5 µg/disc). Absolute ethanol was used as a negative control. These discs were placed on the agar plates to be incubated at 37°C for 24 h. The inhibition zones were measured by Vernier caliper. Regarding antifungal activity, a well diffusion assay was used to determine the antifungal properties of seed and fruit ethanolic extracts. Inoculum measuring 0.2 ml (fungal strain in saline) was spread on an agar plate. Five ditches of 4 mm were made on each plate. Absolute ethanolic plant extracts (50 mg/ml) were prepared, and each well was filled with 50 µl of the methanolic extracts. Fluconazole (5 mg/ml) was used as a positive control, while absolute ethanolic was used as a negative control. The plates were incubated at 37°C for 24 h. Strains of *Candida* species were obtained from the Research Laboratory, Clinical Laboratory Sciences Department, Juef University, Saudi Arabia.

2.16 *In silico* studies and molecular docking

The 3D structures of the proposed molecules were retrieved from the PubChem database and energy-minimized to investigate their interactions with microbial proteins through molecular docking analyses. The proposed molecules were five essential oils (cadinol, cryptone, elemol, thyme, and pinene), tocopherol (vitamin E), kaempferol (flavonoid), and glutamine (amino acid). Two microbial receptors were targeted: sterol 14-alpha demethylase from *Candida albicans* (Protein Data Bank (PDB): 5FSA (Hargrove et al., 2017)) and the transcriptional regulator MvfR of *P. aeruginosa* (PDB: 6Q7U (Zender et al., 2020)). The receptors were retrieved from the PDB database and prepared for docking through the removal of associated ligands and solvent molecules and the addition of polar hydrogen atoms. In AutoDock, the grid box was adjusted to target the entire protein target with a grid spacing of 1 Å between the set points (Morris et al., 2009). Docking simulations and binding affinity were performed and calculated using AutoDock Vina (Trott and Olson, 2010). The molecular interactions were further investigated by visualizing the formed chemical bonds in each ligand–receptor complex as well as the 3D surface structures showing the aromatic interactions, H-bond formation, ionizability, solvent-accessible surface (SAS), and other properties by using BIOVIA Discovery Studio (Gaber et al., 2020; Samaha et al., 2020; El Azab et al., 2021; Gaber et al., 2021; Mohamed et al., 2021; Saied et al., 2021; Healey et al., 2022; Khirallah et al., 2022a; Khirallah et al., 2022b).

2.17 Statistical investigation

Our tests were accomplished as quadruple. The data are shown as mean ± standard error of the mean. Results were investigated by using one-way ANOVA and subsequently Tukey's *post hoc* test. The significance of the data was defined by p-value; $p > 0.05$ is regarded as non-significant and $p < 0.05$ as significant. The statistical investigation was performed using the software program GraphPad Prism (GraphPad Software, San Diego, CA, USA, 2007) (Abdel-Wahab BA et al., 2022; Mohamed et al., 2022b; Mohamed et al., 2022a).

3 Results and discussion

3.1 Proximate composition analysis

For *A. subulatum*, sugar has the largest total content among other nutrients. For seeds, it was estimated to be approximately 33% of the total nutrient content in seeds, while that in fruits was approximately 25.5% of the nutrient content. Higher sugar content was found in *A. xanthioides* than *A. subulatum*. After sugars, both species contained considerable amounts of proteins and lipids. Notably, the protein content is usually higher than the total lipids in both species except the fruit of *A. subulatum*, which contains more lipids than proteins. However, lower amounts of steroids, tannins, and alkaloids were also observed. The investigation of alcoholic extracts of some *Amomum* species revealed the presence of various phytoconstituents, including flavonoids, carbohydrates, anthocyanin, tannins, phenols, alkaloids, and steroids (Konappa et al., 2020). Li et al. reported that *Amomum tsao-ko* is rich in a variety of chemical components of proteins, phenolic compounds, tannins, organic acids, saponins, flavonoids, anthraquinone, coumarin, lactones, steroids, terpenoids, volatile oil, anthocyanins, and so on (Li et al., 2021). As tannin, a wide group of water-soluble polyphenolics, and alkaloid possess anti-nutritional and anti-feed

properties (Mueller-Harvey, 2006; Itkin et al., 2013), low availability of these compounds indicated high tissue quality of both *Amomum* species. However, tannins possessed antioxidant characteristics and have been shown to reduce total cholesterol, blood pressure, and immune system stimulation (Tong et al., 2021). Other nutrients are present in much lower concentrations including flavonoids, crude fiber, and ash (Figures 1A, B). The content of ash often indicated high levels of inorganic compounds and essential macroelements (Alzahrani et al., 2017).

This is consistent with the antioxidant activity associated with the ethyl acetate seed extract and the immune response activated in mice upon administration of fruit extracts (Sharma et al., 2017; Sudarsanan et al., 2021). Moreover, an anti-inflammatory function associated with tannins from various sources was reported in other studies (Park et al., 2014; Ambreen and Mirza, 2020; Sharma et al., 2021), which explains the beneficial use of *A. xanthioides* in the management of atopic dermatitis (Choi et al., 2017).

3.2 Essential oil isolation

Our data revealed that there are several essential oils found in both *A. subulatum* and *A. xanthioides* with different concentrations in both seeds and fruits (Table 1). Some of them showed significant differences between the two species in either seed or fruits, while the others showed no significant changes between the two species. High levels of thyme oil may promote the antimicrobial activity associated with *A. subulatum* and *A. xanthioides* (Satyal et al., 2012; Choi et al., 2017; Thinh et al., 2022). Thyme oil is used in food preservation and cosmetics due to its antibacterial, antifungal, and anti-inflammatory characteristics (Lorenzo et al., 2019). Thyme oil aids in promoting blood flow to the skin, which improves the healing process and removal of scars and imperfections, leaving the skin even and healthy (Zarzuelo and Crespo, 2002; Salehi et al., 2018). Reports of antibacterial activity of thyme oil have been found against *Clostridium botulinum*, a gram-positive bacterium that produces one of the most lethal neurotoxins known (De and De, 2019; Corsalini et al., 2021). In addition, thyme oil significantly reduced the growth of the fungus *Pyrenochaeta terrestris* during soil treatment, which suggests the potential use of thyme oil in the biological control of plant diseases (Hussien and Ibrahim, 2020). Due to their wide range of applications, *A. subulatum* and *A. xanthioides* can be used as sources for the extraction and production of thyme oil.

After thyme, elemol oil is the next-most prevalent oil in both species. The percent of elemol oil in seed was approximately 19%, while that in fruits was 34% for the two species. It is commonly used in fine fragrances as well as other applications (Bhatia et al., 2008). Elemol oil isolated from essential oils of *Rhynchanthus beesianus* rhizomes exhibited *insecticidal* activities against the adults of *Liposcelis entomophila* and *Tribolium castaneum* insects (Pan et al., 2022). Essential oils of *Cymbopogon schoenanthus* contained 22.8% elemol oil and exhibited larvicidal effects against *Anopheles funestus* and *Culex quinquefasciatus* larvae with lethal concentrations of 120.5 and 23.32 ppm, respectively (Wangrawa et al., 2022). In addition, elemol was the most abundant constituent (13.54%) in the essential oils of *Drimys winteri*, which showed insecticidal activity against the pests *Acanthoscelides obtectus* and *Aegorhinus superciliosus* (Tampe et al., 2020). The reported insecticidal activity of essential oils with elemol in multiple studies suggests its role as a natural eco-friendly insecticide for the biocontrol of herbivorous arthropods.

In addition to thyme and elemol components, considerable levels of the terpenoid α-pinene were also noticed in the two species. 1,8-Cineole, α-pinene, α-terpinene, and β-pinene were found to be the major components in *Amomum kravanh* (Diao et al., 2014). The oil has demonstrated a strong anticancer effect on various cell lines in addition to anti-inflammatory effects against multiple autoimmune disorders (Salehi et al., 2019; Prado-Audelo et al., 2021), which agrees with the high therapeutic activity of *A. subulatum*. Essential oils from *Magnolia candollei*, which contained 29.7% α-pinene and 10.2% elemol components, exhibited inhibitory against lipoxygenase and acetylcholinesterase, suggesting the potential therapeutic capabilities of these oils in diseases related to lipid metabolism as well as neurological disorders (Yahaya et al., 2022). Noticeably, the carvacrol levels in *A. subulatum* were 10 times higher than those in *A. xanthioides*.

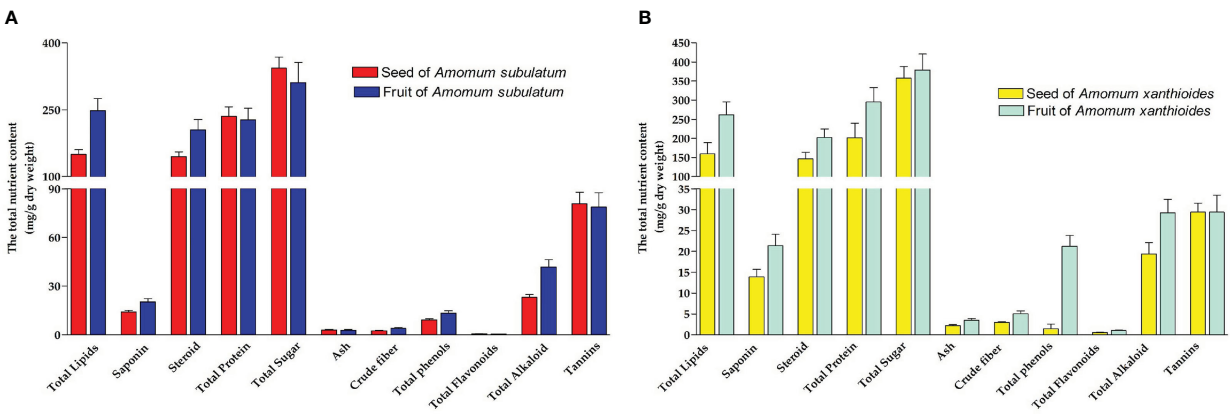


FIGURE 1
The total nutrient content (mg/g dry weight) in **(A)** seed and fruit of *Amomum subulatum* and **(B)** seed and fruit of *Amomum xanthioides*. The data are shown as the mean ± SEM of four plants per group.

TABLE 1 Essential oil percentages (v/g dry weight) in the seed and fruit of *Amomum subulatum* and *Amomum xanthioides*.

Essential oil	*Amomum subulatum*		*Amomum xanthioides*	
	Seed	Fruit	Seed	Fruit
γ-Terpinene	5.63 ± 0.41	7.72 ± 0.85	7.46 ± 0.59	6.01 ± 0.66
α-Thujene	0.41 ± 0.04	0.54 ± 0.06	0.29 ± 0.09	0.45 ± 0.05
α-Pinene	17.01 ± 1.81	19.88 ± 2.58	12.82 ± 0.94	22.82 ± 2.9
α-Phellandrene	0.94 ± 0.09	1.44 ± 0.28	0.5 ± 0.04	1.59 ± 0.17[b]
Myrcene	4.29 ± 0.31	4.73 ± 0.74	3.17 ± 0.34	5.49 ± 0.78
Sabinene	1.07 ± 0.08	3.91 ± 0.88[a]	2.74 ± 0.29	3.38 ± 0.61
β-Pinene	4.54 ± 0.34	7.9 ± 0.96	2.3 ± 0.69	1.26 ± 0.49
Carvacrol	13.01 ± 1.63	11.08 ± 1.6	0.86 ± 0.08[a]	0.25 ± 0.03
***p*-Cymene**	5.34 ± 0.49	5.27 ± 0.61	3.48 ± 0.25	1.88 ± 0.05[b]
β-Phellandrene	0.16 ± 0.02	0.16 ± 0.02	0.14 ± 0.02	0.3 ± 0.04
Thyme	50.96 ± 5.91	78.75 ± 8.61	48.98 ± 5.67	48.65 ± 5.43
Elemol	18.82 ± 1.69	33.22 ± 2.99	18.96 ± 3.12	34.89 ± 6.26
Myrtenal	1.81 ± 0.17	2.12 ± 0.19	1.62 ± 0.12	2.21 ± 0.19
Verbenone	1.44 ± 0.15	1.64 ± 0.15	1.07 ± 0.08	1.87 ± 0.18
Cuminal	2.09 ± 0.16	2.36 ± 0.21	1.47 ± 0.14	2.72 ± 0.28
Ascaridole	0.43 ± 0.03	0.66 ± 0.11	0.47 ± 0.03	0.68 ± 0.09
***E*-nerolidol**	0.45 ± 0.03	0.91 ± 0.12	0.4 ± 0.08	0.36 ± 0.08
Widdrol	1.4 ± 0.15	1.46 ± 0.12	0.25 ± 0.05	0.12 ± 0.04
***epi*-Cubenol**	1.47 ± 0.13	1.25 ± 0.13	0.35 ± 0.03	0.18 ± 0.01[c]
γ-Eudesmol	0.44 ± 0.04	0.42 ± 0.03	0.29 ± 0.02	0.18 ± 0.01
***epi*-α-Muurolol**	4.09 ± 0.47	6.06 ± 0.43	3.93 ± 0.45	3.76 ± 0.28
β-Eudesmol	5.58 ± 0.6	8.7 ± 0.65	5.43 ± 0.41	6.41 ± 0.67
α-Cadinol	9.57 ± 0.91	16.31 ± 1.38	9.54 ± 1.17	15.26 ± 2.31
Oplopanone	0.55 ± 0.05	0.77 ± 0.06	0.52 ± 0.05	0.55 ± 0.04
Pinocarvone	0.26 ± 0.03	0.3 ± 0.03	0.22 ± 0.02	0.33 ± 0.03
Cryptone	15 ± 1.27	17 ± 1.52	10.78 ± 0.9	19.52 ± 1.43
Linalool	5.21 ± 0.37	2.96 ± 0.21[a]	5.29 ± 0.38	2.31 ± 0.17

The data are shown as the mean ± SEM of four plants per group. [a]$p < 0.05$ versus seeds of A. subulatum. [b]$p < 0.05$ versus seeds of A. xanthioides. [c]$p < 0.05$ versus fruits of A. subulatum.

Lower levels of α-thujene, phellandrenes, ascaridole, and pinocarvone were also observed in both species. Other components were also isolated and summarized in Table 1. Statistical analyses showed that *A. subulatum* contained significant correlation levels ($p < 0.05$) of sabinene component in fruits more than in seeds while having linalool in seeds more than in fruits ($p < 0.05$). For *A. xanthioides*, significantly higher levels ($p < 0.05$) of α-phellandrene were observed in fruits. In addition, cymene seed levels were higher compared to those of fruits in *A. xanthioides*. When compared to *A. xanthioides*, *A. subulatum* showed higher significance ($p < 0.05$) of carvacrol in seeds and lower significance ($p < 0.05$) of cubenol in fruits. α-Pinene from some *Salvia* spp. showed moderate antibacterial activity against multiple bacterial species including *Bacillus subtilis*, *S. aureus*, and *Staphylococcus epidermidis* with inhibition zones ranging from 13 to 15 mm and minimum inhibitory concentration (MIC) values of 3.7–7.5 mg/ml (Yousefzadi et al., 2007). Moreover, α-pinene had a strong antifungal potential against *Candida* species isolated from otomycosis patients, indicating its promising clinical utility as a natural drug (Nóbrega et al., 2021). In this regard, it was suggested that essential oil might bind to the bacterial cell surface and then penetrate the phospholipid bilayer of the cytoplasmic membrane. Consequentially, it results in the leakage of various vital intracellular constituents and leads to cell death (Lv et al., 2011). Thus, the presence of α-pinene and elemol components in considerably high amounts in *A. subulatum* and *A. xanthioides* makes the plant species extremely important sources of these oils due to their great benefits in fighting both plant and human pathogens.

3.3 Elemental analysis

In the seed and fruit of both species, nitrogen was the highest frequent element followed by potassium and then phosphorus (Figures 2A, B). *A. subulatum* contained lower levels of nitrogen (~45 mg/g dry weight) than *A. xanthioides* seed and fruit. Some elements were found in much lower amounts (less than 3 mg/g dry weight) including calcium, magnesium, sodium, and zinc. A significant increase in levels ($p < 0.05$) of magnesium was observed in the fruits of *A. subulatum* compared to seeds. However, zinc levels were significantly higher ($p < 0.05$) in fruits of *A. subulatum* compared to fruits of *A. xanthioides*. It was also reported that *Amomum longiligulare* is rich in microminerals (Chau et al., 2015). In this context, edible Zingiberaceae plants contained relatively high macroelement amounts including K, Ca, and Fe (Rachkeeree et al., 2018). Nitrogen is regarded as a critical macronutrient among all mineral nutrients including all living tissues within the plant, ranging from metabolism to resource allocation, growth, and development. Nitrogenous compounds play several roles in metabolism, including acting as protein building blocks and as precursors in the biosynthetic pathways of substances like lignin (Famiani et al., 2020; Yousaf et al., 2021). Phosphorus is a crucial nutrient in crop yields since many soils in their natural condition do not have enough accessible phosphorus to enhance crop yield (Sett and Soni, 2013). Potassium supplementation improved osmotic adjustment and water relations in a variety of crop species. K alleviates drought conditions in plant material primarily by managing stomatal closure and preserving stromal pH, thereby lowering photo-oxidative injury to chloroplasts (Santos et al., 2021). This means that these plants have excellent quality in improving the osmotic adjustment of water and preserving stromal pH.

3.4 Total phenolics and flavonoids

In determining the total phenolic and flavonoid components in both species *A. subulatum* and *A. xanthioides* in either seeds or fruits, we found that both species had no significant difference in the phenolic or flavonoid levels. However, there was a significant increase in the levels of caffeic acid and velutin ($p < 0.05$) in *A. xanthioides* fruits compared to seeds (Figures 3A, B). The phenolic–flavonoid profiling showed that only gallic acid, a phenolic compound, is produced in high amounts in the two species compared to the other screened compounds. Most of the phytochemical constituents in *Amomum* sp. are phenolics, flavonoids, and essential oils (Alkandahri et al., 2021). Smaller amounts of other phenolics including caffeic acid, ferulic acid, and catechin were found. Isolated flavonoids were quercetin, kaempferol, naringenin, luteolin, and apigenin. Similarly, quercetin was previously isolated from *A. longiligulare* (Chau et al., 2015).

Owing to its anticancer, antioxidant, antimicrobial, and anti-inflammatory capabilities, gallic acid has therapeutic applications in treating multiple medical conditions including GI infections, neurological diseases, and heart disorders (Kahkeshani et al., 2019; Shukla et al., 2022). Gallic acid was reported to induce apoptosis in leukemia HL60 cell lines owing to its antitumor and anti-proliferative effects observed in multiple studies (Maruszewska and Tarasiuk, 2019; Soto-Maldonado et al., 2019; Roba'ei et al., 2022). Gallic acid also showed potent anti-obesity properties by directly targeting adipose tissue and inhibiting lipogenesis (Dludla et al., 2018), which was reported upon administration of ethyl acetate extract of *A. xanthioides* to high-fat-diet mouse and rat models (Kim et al., 2015; Im et al., 2020). Regarding skin diseases, gallic acid was also reported to have an immunosuppressive role in the regulation of atopic dermatitis by inhibiting the production of certain pro-inflammatory cytokines and chemokines (Tsang et al., 2016). These studies further increase the medicinal importance of *A. subulatum* and *A. xanthioides* plants as sources of natural product-based drugs since they contain considerable amounts of gallic acid.

3.5 Vitamins

While screening for vitamins, small amounts of vitamins A, B, and K were found, which did not exceed 1 mg/g dry weight, while

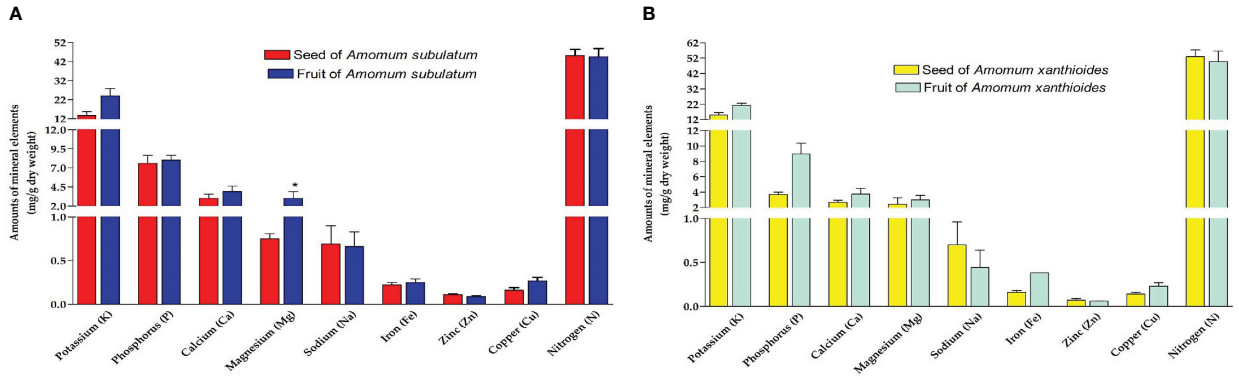


FIGURE 2
Amounts of mineral elements extracted (mg/g dry weight) in **(A)** seed and fruit of *A. subulatum*, **(B)** seed and fruit of *A. xanthioides*. The data are shown as the mean ± SEM of 4 plants per group. a $P < 0.05$ versus seeds of *A. subulatum*, c $P < 0.05$ versus fruits of *A. subulatum*.

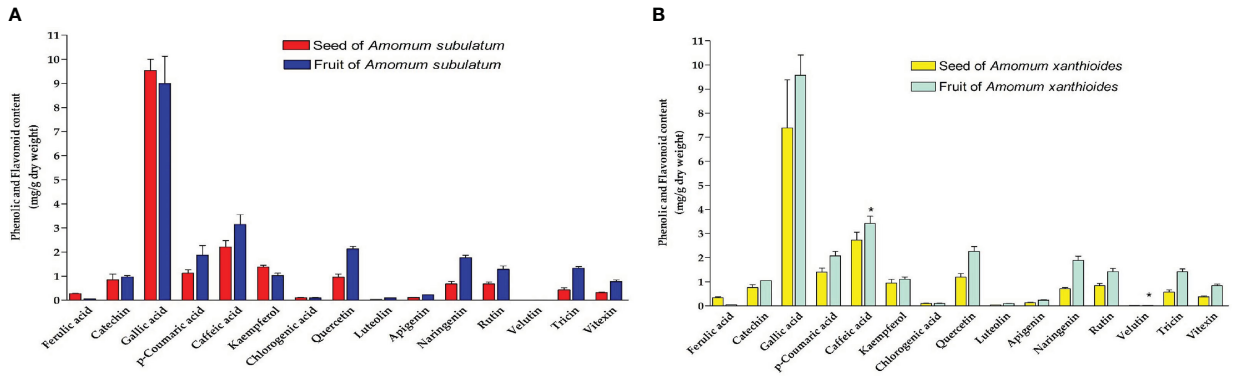


FIGURE 3
Phenolic and flavonoid content (mg/g dry weight) in **(A)** seed and fruit of *Amomum subulatum* and **(B)** seed and fruit of *Amomum xanthioides*. The data are shown as the mean ± SEM of four plants per group. [b] $p < 0.05$ versus seeds of *A. xanthioides*.

slightly higher amounts of vitamins C and E were found in seed and fruit of both species (Figures 4A, B). Despite the overall small amounts of vitamins, seeds of *A. xanthioides* had significantly ($p < 0.05$) higher levels of vitamin E and vitamin C in seeds when compared to *A. subulatum*. A significant increase ($p < 0.05$) of vitamin E and vitamin K levels was also observed in *A. xanthioides* fruits than in *A. subulatum* fruits. Vitamin E content in plants is thought to be linked to ripening and senescence. The non-volatile substances isolated from this *Amomum* including vitamins were also observed (Chau et al., 2015). Plants are subjected to oxidative stress throughout senescence, which leads to a rise in fatty acid-free radicals. Enhancing the content of vitamin E during senescence may be a stress-reduction technique for plants. In addition, by collaborating with other antioxidants, vitamin E can effectively eliminate the generation of reactive oxygen species (ROS). In plants, vitamin C is the most abundant and widely distributed water-soluble cellular antioxidant. Plants can avoid oxidative damage by directly scavenging ROS with vitamin C. Furthermore, vitamin C is essential for plant growth, development, and stress responses. It functions as a cofactor for many enzymes, regulates cell division, and influences cell expansion. It also regulates plant senescence (Paciolla et al., 2019; Zhang et al., 2020). These results revealed the antioxidant effects of these plants in the DPPH, FRAP, and anti-lipid peroxidation tests.

3.6 Amino acids

Amino acid profiling revealed that both species possess small amounts of all amino acids except for glutamine and glutamic acid, which were found in much higher amounts in both seed and fruit (Figures 5A, B). Generally, *A. subulatum* had higher amounts of the two amino acids than *A. xanthioides*. Leucine levels were significantly increased ($p < 0.05$) in seeds of *A. subulatum* compared with the fruits of the same species. A comprehensive chemical investigation found that *A. tsao-ko* contained a variety of chemical components including protein and amino acids (Xie et al., 2022). Among the amino acids, glutamine serves as a major amino donor for the synthesis of amino acids, nucleotides, and other nitrogen-containing compounds, in addition to its role in nutrition and metabolism (Kan et al., 2015).

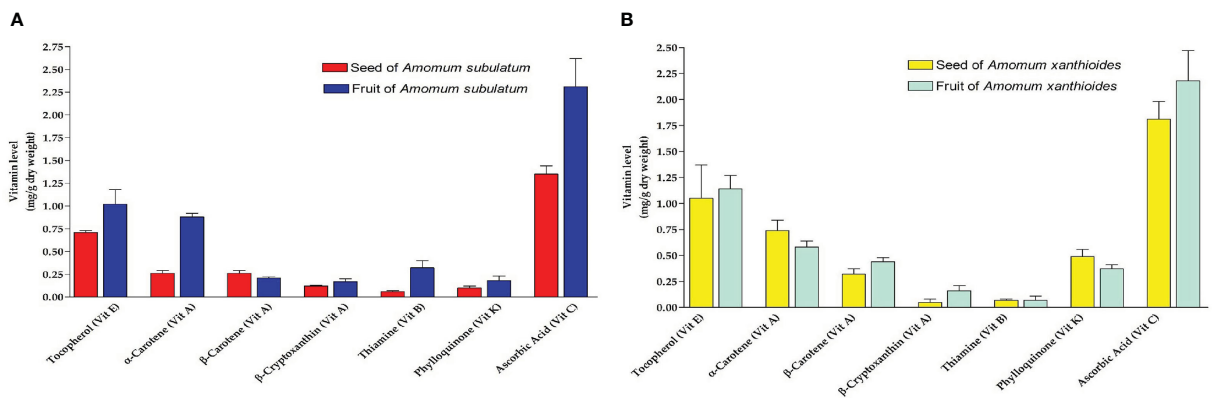


FIGURE 4
Vitamins levels (mg/g dry weight) isolated in **(A)** seed and fruit of *Amomum subulatum* and **(B)** seed and fruit of *Amomum xanthioides*. The data are shown as the mean ± SEM of four plants per group. [a] $p < 0.05$ versus seeds of *A. subulatum*, [c] $p < 0.05$ versus fruits of *A. subulatum*.

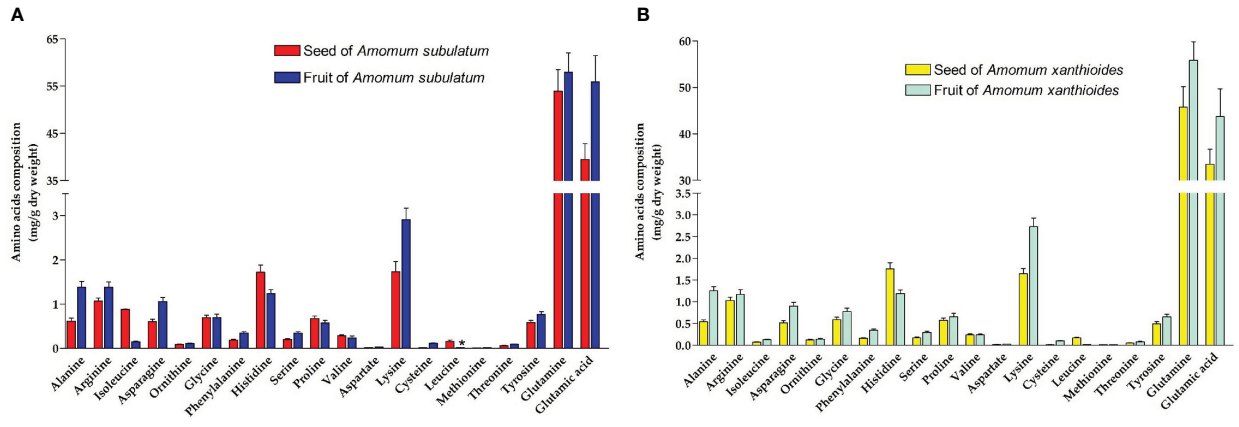


FIGURE 5
The amino acid compositions (mg/g dry weight) in **(A)** seed and fruit of *Amomum subulatum* and **(B)** seed and fruit of *Amomum xanthioides*. The data are shown as the mean ± SEM of four plants per group. [a] $p < 0.05$ versus seeds of *A. subulatum*.

3.7 Antioxidant biological activity

Multiple tests revealed the antioxidant activity of seed and fruit extracts of both species (Table 2). According to the DPPH assay, the seed extract of *A. subulatum* exhibited the highest antioxidant activity followed by the seed extract of *A. xanthioides*. Similarly, FRAP assay showed that the seed extract of the two species had the best antioxidant activity, i.e., 20.14 ± 1.11 and 21.18 ± 1.04 µmol Trolox g^{-1} DW for *A. subulatum* and *A. xanthioides*, respectively. In terms of anti-lipid peroxidation, higher values were obtained from the fruit extract of *A. subulatum* and the seed extract of *A. xanthioides*. No statistical significance was found between the antioxidant activity of seeds and fruits in the same plant or the two species. These results indicate the high antioxidant activity of these extracts that contributes to their usefulness in a variety of therapeutic applications. It was also observed that plants of the genus *Amomum* showed antioxidant and anti-inflammatory activities (Kwon et al., 2003). The presence of gallic acid and tannins may be the main contributor to the high antioxidant activity associated with *A. subulatum*, which was previously reported in various studies (Bhaswant et al., 2015; Gautam et al., 2016; Sharma et al., 2017). The intermediate antioxidant activity of essential oils isolated from two *A. subulatum* fruit samples collected from India and Saudi Arabia was formerly reported (Alam and Singh, 2021). The free radical scavenging activity at 1,000 µg/ml was approximately 85.27% in the Indian sample and 86.86% in the Saudi Arabian sample, and the IC_{50} values were 219.38 and 203.79 µg/ml for the Indian and Saudi Arabian samples, respectively. For *A. xanthioides*, antioxidant activity was previously evaluated for the water and ethanol extracts using DPPH assay, which showed that the water extract had higher antioxidant activity than the ethanolic one with IC_{50} values of 2.87 and 3.31 µg/ml for water and ethanol extracts, respectively (Myint et al., 2012).

3.8 Antimicrobial activity

Collectively, extracts from both species demonstrated strong antimicrobial effects against various pathogenic bacteria and fungi (Figures 6A, B). Ethanolic seed extracts of *A. subulatum* had the highest efficiency against four gram-negative bacterial species that cause serious human diseases, namely, *Enterobacter aerogenes*, *Proteus vulgaris*, *P. aeruginosa*, and *Salmonella typhimurium*. In addition, *P. aeruginosa* was also inhibited by the fruit extract of both *A. subulatum* and *A. xanthioides*. For the seed extract of *A. xanthioides*, large inhibition zones were formed against *P. vulgaris* and the fungus *C. albicans*. In this regard, *Amomum compactum* extracts damaged the membrane of fungal cells and inhibited the enzyme system of fungi to create a bland zone around the disc (Oh et al., 1951). The fruit extract of *A. xanthioides* caused high inhibition of *Staphylococcus saprophyticus* and *S. typhimurium* in addition to *P. aeruginosa*. While there were non-significant correlations between the antimicrobial activity of seeds and fruits between the two species, a statistically significant increase ($p < 0.05$)

TABLE 2 Antioxidant activity exhibited by *Amomum subulatum* and *Amomum xanthioides* extracts.

Assay	*Amomum subulatum*		*Amomum xanthioides*	
	Seed	Fruit	Seed	Fruit
DPPH (% inhibition)	78.26 ± 9.27	63.27 ± 3.33	68.21 ± 2.56	51.53 ± 1.72
FRAP (µmol Trolox g^{-1} DW)	20.14 ± 1.11	13.19 ± 1.15	21.18 ± 1.04	15.66 ± 1.61
Anti-lipid peroxidation	4.87 ± 0.34	6.08 ± 0.35	6.11 ± 0.55	5.12 ± 0.8

The data are shown as the mean ± SEM of four plants per group.
FRAP, ferric ion-reducing antioxidant power.

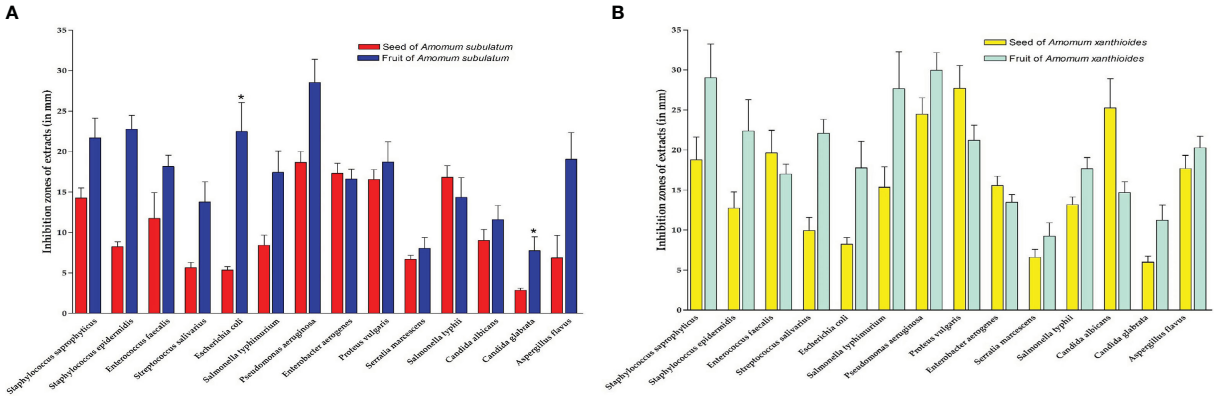


FIGURE 6
Inhibition zones (mm) of extracts in **(A)** seed and fruit of *Amomum subulatum* and **(B)** seed and fruit of *Amomum xanthioides*. The data are shown as the mean ± SEM of four plants per group. [a] $p < 0.05$ versus seeds of *A. subulatum*..

was observed in the antimicrobial activity of *A. subulatum* fruits with *E. coli* and *Candida glabrata* compared with seeds.

Indeed, the antimicrobial potential of both species was reported in various studies against the tested pathogens including *S. aureus* and *A. niger* (Satyal et al., 2012; Thinh et al., 2022). A concentration-dependent antibacterial activity was observed upon testing the essential oils of Indian and Saudi Arabian *A. subulatum* fruit samples against *Acinetobacter baumannii*, *E. coli*, and *P. aeruginosa*, with an inhibition zone range of 12–16 mm (Alam and Singh, 2021). In an older study, essential oils from Indian *A. subulatum* fruits exhibited antimicrobial activity similar to or higher than that of ciprofloxacin (as a standard) against *B. pumilus*, *S. aureus*, *S. epidermidis*, and *P. aeruginosa* (Agnihotri et al., 2012). For the seeds, methanolic seed extracts of *A. subulatum* exhibited reasonable antimicrobial activity at a dose of 800 mg/ml against *B. subtilis*, *Salmonella typhii*, and *P. aeruginosa*, when compared to tetracycline as a standard (Tijjani et al., 2012). Bacterial cell death might come from the effect of essential oil on membrane permeability, resulting in bacterial cell wall lysis (Bajpai et al., 2013). Generally, essential oils are considered the plant immune system due to their antimicrobial effect against a wide range of pathogens (Monica and Ioan, 2018). The pure form of essential oils from *A. subulatum* exhibited antimicrobial activity against various bacterial pathogens including *E. faecalis*, *S. aureus*, and *Shigella dysenteriae* forming inhibition zones of 14–23 mm (Tandukar et al., 2018). For the two species, the ratio of thyme oil was the highest among other oils in seed and fruit. In general, thyme oil and gallic acid could be the main factors promoting the antimicrobial properties of the two species due to their known antibacterial and antifungal activities and their high abundance among other oils (Borges et al., 2013; Sarjit et al., 2015; Lorenzo et al., 2019).

3.9 Molecular docking simulations

Sterol 14-alpha demethylase of the fungus *C. albicans* is a common target of azole drugs since it is essential in maintaining the fungus cell membrane integrity through catalyzing a vital step in ergosterol synthesis (Hargrove et al., 2017; Marek and Timmons, 2019; Zhang et al., 2019). The docking results of the proposed molecules are summarized in Table 3. The standard antifungal drug, posaconazole, formed the highest interaction with the receptor, which was indicated by the low binding affinity (−15.2 kcal/mol), followed by tocopherol (−9.6 kcal/mol) and the flavonoid kaempferol (−8.4 kcal/mol). The hydrophobicity of 3D surfaces shows that more hydrophobic interactions were formed between the ligands and the target molecules (Figure 7). Aromatic interactions visualize the most stable conformation of each ligand in the binding pocket with respect to the whole protein molecule (Figure S2). SAS structures show that a large surface area, shown in blue, is involved in the ligand binding, which increases the binding efficiency and further stabilizes the ligand in the protein binding site (Figure S3). Since the interactions mainly depended on the hydrophobic interactions, no formal charges or ionized bonds were formed, and a relatively neutral interface (white) is formed between the ligands and the receptor (Figure S4). Some molecules, including kaempferol, elemol, and glutamine, formed polar hydrogen bonds with the target receptor, which decreased the binding energy and contributed markedly to stabilizing the formed complex. The hydrogen bond donor/acceptor surface showed the polar areas (pink/green) between the ligand and the receptor due to the formation of hydrogen bonds (Figure S5). The 2D interaction diagram (Figure 8) showed the specific positions of amino acids involved in the interaction with each molecule as well as the types of chemical bonds formed between them in addition to their distances stated in Table 4.

The transcriptional regulator MvfR of *P. aeruginosa* is responsible for the regulation of multiple virulence genes that are responsible for the pathogenicity of *P. aeruginosa*; therefore, it is an important drug target (Déziel et al., 2005; Allegretta et al., 2017; Zender et al., 2020). Similar to the demethylase, the interaction of the proposed molecules with the MvfR was dependent on hydrophobic interactions rather than hydrophilic ones, except for glutamine. The docking results, in Table 3, showed that tocopherol and kaempferol formed stronger and more stable complexes with the receptor than the antibacterial agent 2-heptylquinolin-4(1*H*)-one (HQQ), which indicates the high potency of these compounds

10.3389/fpls.2023.1136961

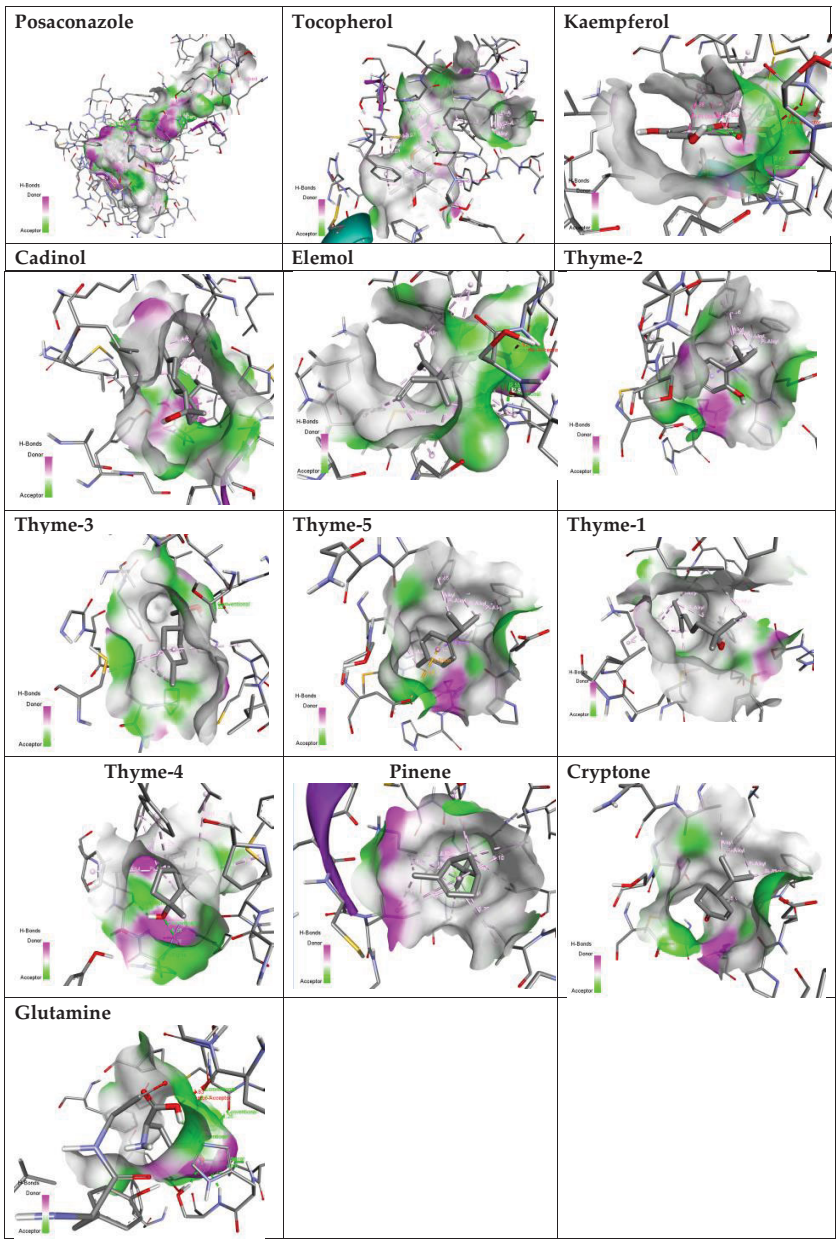


FIGURE 7
Hydrogen bond donor/acceptor surface view of the proposed molecules on sterol 14-alpha demethylase.

and their promising antibacterial capabilities. The binding affinity was −8.3 and −8.2 kcal/mol for tocopherol and kaempferol, respectively, while for HQQ, it was −8 kcal/mol. In addition, more residues were involved in the hydrophobic interactions between tocopherol and MvfR, while the kaempferol–MvfR complex involved the formation of one hydrogen bond in addition to the hydrophobic one, which further stabilizes the ligand through partial charge formation. A large hydrophobic surface was observed between the ligands and the receptor (Figure S6). Aromatic surfaces show the conformation of the docked molecules responsible for the calculated binding affinity (Figure S7), and the large solvent-accessible area further confirms the stable complex (Figure S8). As expected, very low charge bonds are formed between the ligands and the receptors since the reactions relied mostly on hydrophobic interactions leading to the formation of neutral interfaces (Figure S9). The hydrogen bond formation was observed in the cases of kaempferol, cryptone, thyme-3,4, and glutamine, which are shown in the hydrogen bond donor/acceptor surface view in Figure S10. The stable ligand conformation was increased due to the formation of multiple chemical bonds with many atoms of the ligand that hold the ligand steady and make it more stable in the binding site. This was highly observed in tocopherol, cadinol, and elemol (Figures 9, 10).

10.3389/fpls.2023.1136961

FIGURE 8
2D ligand interaction diagram of the proposed molecules with sterol 14-alpha demethylase.

TABLE 3 Docking results of the proposed molecules with transcriptional regulator MvfR (PDB: 6Q7U).

Name	Binding affinity	Hydrophobic residues	Distance	H-bond residues	Distance
HQQ	–8	LEU197- LEU208- ALA168- PRO238- ILE236- PRO129- ALA130	3.96076- 3.88127- 4.29327- 4.05885- 5.32035- 5.00021- 4.83028	–	–

(Continued)

10.3389/fpls.2023.1136961

TABLE 3 Continued

Name	Binding affinity	Hydrophobic residues	Distance	H-bond residues	Distance
Tocopherol	−8.3	LEU207- ALA102- ALA130- ALA168- ILE263- ILE149- PRO238- LEU197- LEU208- ILE236- PHE221- TYR258-	3.78142- 3.90905- 3.68213- 4.23219- 5.19753- 4.5559- 3.75019- 3.93976- 3.77058- 4.71286- 4.83515- 5.26525	–	–
Kaempferol	−8.2	ILE236- ALA168- LEU208- ALA102- ILE149- PRO238	3.55105- 5.47727- 5.29956- 4.45556- 4.3678- 4.76949	LEU207	2.58672
Cadinol	−7.7	ALA102- ALA168- LEU208- ILE236- ILE149- LEU197- PRO129- LEU208 PHE221	4.40483 4.41751 5.2447 4.02302 4.90871 4.30594 3.89275 3.92425 5.33667	–	–
Thyme-2	−6.8	LEU208 PHE221 ALA130 ALA168 PRO129 LEU197 ILE149 ILE236 PRO238 PRO129 LEU197	3.72919 5.10225 3.79191 4.29797 4.45527 3.69787 5.37132 4.76293 4.11078 4.9306 4.83712	–	–
Cryptone	−6.7	LEU197 LEU208 ILE236	4.605 4.22898 3.71001	ALA130	2.32825
Thyme-5	−6.7	LEU208 PHE221 ALA130 ALA168 LEU197 LEU208 ILE149 ILE236 PRO238 PRO129	3.65717 5.2505 3.81477 4.28436 3.82674 3.83104 5.27522 4.86006 4.17347 5.03709	–	–
Thyme-3	−6.6	ALA130 LEU197 LEU208 PRO129	3.77576 4.11458 4.22878 3.43155	GLN194 SER196 LEU208	2.95299 2.66477 2.66181
Elemol	−6.3	ALA102 ALA168 LEU208 ILE236 ILE149 PRO238	4.26534 4.49849 4.91296 4.30588 5.14165 4.3817	–	–

(Continued)

10.3389/fpls.2023.1136961

TABLE 3 Continued

Name	Binding affinity	Hydrophobic residues	Distance	H-bond residues	Distance
		PRO129 LEU207 PHE221	4.54934 4.0581 4.85289		
Thyme-4	−6.1	ALA130 LEU197 LEU208 PRO129 ILE149 PHE221	4.35225 5.41336 4.11982 3.64299 5.2382 5.4972	ILE236	2.42809
Pinene	−6	LEU208 ILE236 MET224 LEU197 PRO129 PHE221	5.01287 4.72332 5.08638 4.39848 4.14459 4.90766	-	-
Thyme-1	−5.9	ALA130 ALA168 ILE149 LEU197 LEU208 PRO129	3.71024 3.61391 4.78145 4.39862 4.07807 3.80864	-	-
Glutamine	−5.5	–	–	GLN194 LEU197:O LEU208:O HIS204:ND1 SER196	2.6024 2.85136 2.82554 2.73285 3.75025

TABLE 4 Docking results of the proposed molecules with sterol 14-alpha demethylase (PDB: *5FSA*).

Name	Binding affinity	Hydrophobic residues	Distance	H-bond residues	Distance
Posaconazole	−15.2	THR311 TYR64 PHE233 ALA61 PHE463 ARG469 ALA62 PRO230 ALA476 LEU370 PRO375 LEU376 ILE379 ILE471 ILE304 PHE58 TYR118 TYR132 HIS377 CYS470 ILE131 LEU121 MET508	3.90404 5.7893 5.48366 4.8006 4.63529 4.8041 4.46946 5.15083 4.37419 5.2972 3.72448 4.6252 4.56791 4.27882 4.67822 4.82018 4.67667 5.46517 4.45682 3.95546 5.2038 5.32755 5.47494	TYR132 ARG381 HIS468 TYR118 GLY307 MET508 HIS377	2.64558 2.24954 2.16623 3.06271 3.60646 3.74934 3.17384
Tocopherol	−9.6	LEU376 TYR118 ALA117 LEU121 LEU376	3.73737 3.82059 3.73786 4.95867 4.96483	–	–

(Continued)

10.3389/fpls.2023.1136961

TABLE 4 Continued

Name	Binding affinity	Hydrophobic residues	Distance	H-bond residues	Distance
		PRO230 MET508 VAL509 LEU88 LYS90 MET92 TYR118 TYR132 PHE228 PHE233 HIS377 PHE380 TYR401	4.70491 3.96657 4.80687 5.13365 4.65839 4.70436 5.24085 5.43382 4.89369 4.86514 4.49524 4.66532 4.53285		
Kaempferol	−8.4	PHE233 - PRO230 LEU376	4.76319 5.07271 5.07682	HIS377	2.41623
Cadinol	−7.5	LEU87 PRO230 PHE233 HIS377 PHE380	5.39462 5.02188 4.01347 4.60132 4.41745	-	-
Elemol	−7.5	PRO230 LEU87 TYR64 PHE233 HIS377 PHE380	5.18674 5.00227 4.19177 4.37583 4.96068 4.88153	HIS377	2.54609- 2.96656
Thyme-2	−6.4	ILE197 ALA218 PHE198 PHE213 TYR221	3.94221 4.3587 4.3256 4.57551 4.9718	–	–
Thyme-3	−6.4	PRO375 CYS470 PHE463	4.86757 5.10794 5.4915	THR315	2.31527
Thyme-5	−6.1	ILE197 ALA218 PHE198 PHE213 TYR221	3.94487 4.46078 4.307 4.57848 4.93474	-	-
Thyme-1	−5.9	LYS90 MET92 PHE233 PHE380 TYR401	4.62746 4.7503 4.40604 4.92622 4.26215	–	-
Thyme-4	−5.9	HIS377 LEU121 LEU376 MET508 LEU376 TYR118 PHE233 PHE380	3.86265 5.46847 5.06807 4.44248 4.01994 4.56292 4.82083 5.13759	HIS377 SER378	3.06152 2.23329
Pinene	−5.8	ILE131 LYS143 ALA146 ILE304 ILE471 LEU139 LEU300	5.10348 5.19261 4.32492 5.29182 5.25868 4.98531 5.15443	-	-

(Continued)

10.3389/fpls.2023.1136961

TABLE 4 Continued

Name	Binding affinity	Hydrophobic residues	Distance	H-bond residues	Distance
Cryptone	−5.7	ALA218 PHE198 ILE197 PHE213	3.95208 5.06528 4.62198 4.85444	-	-
Glutamine	−5.2	–	–	HIS377 SER378 SER507 MET508	3.03508 2.29239 2.61005 2.25831

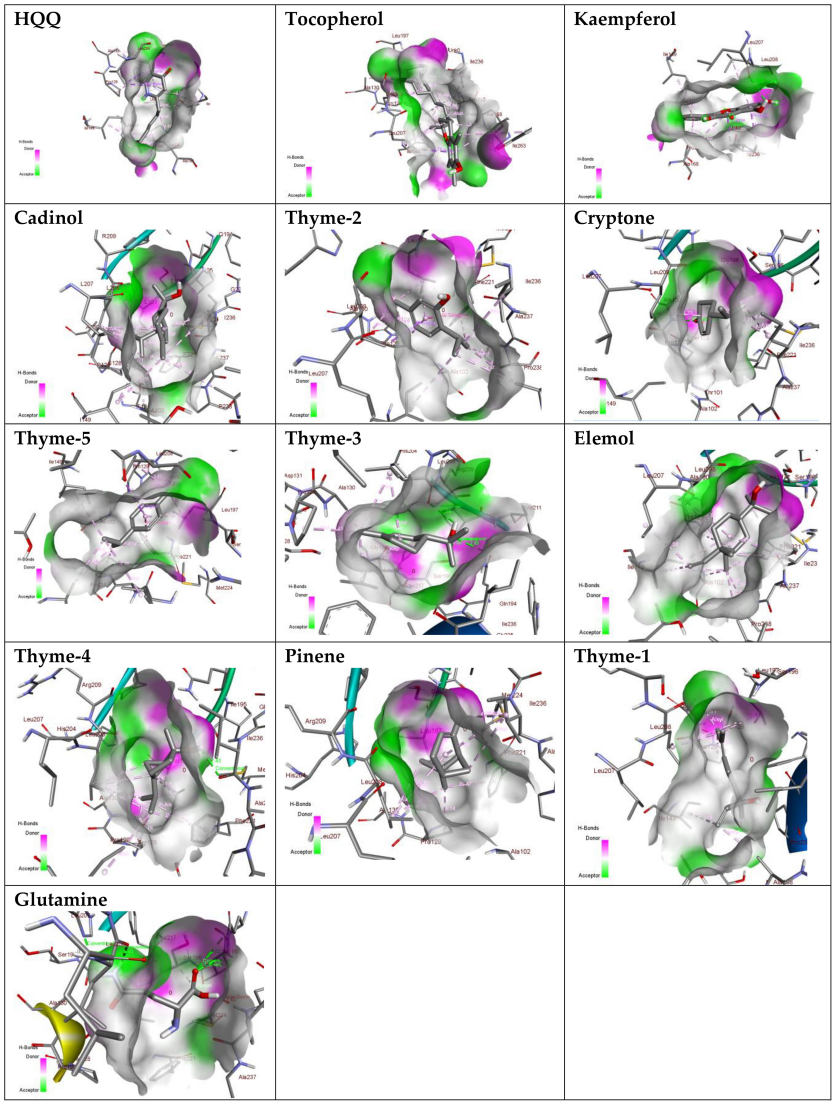


FIGURE 9
Hydrogen bond donor/acceptor surface view of the proposed molecules on MvfR.

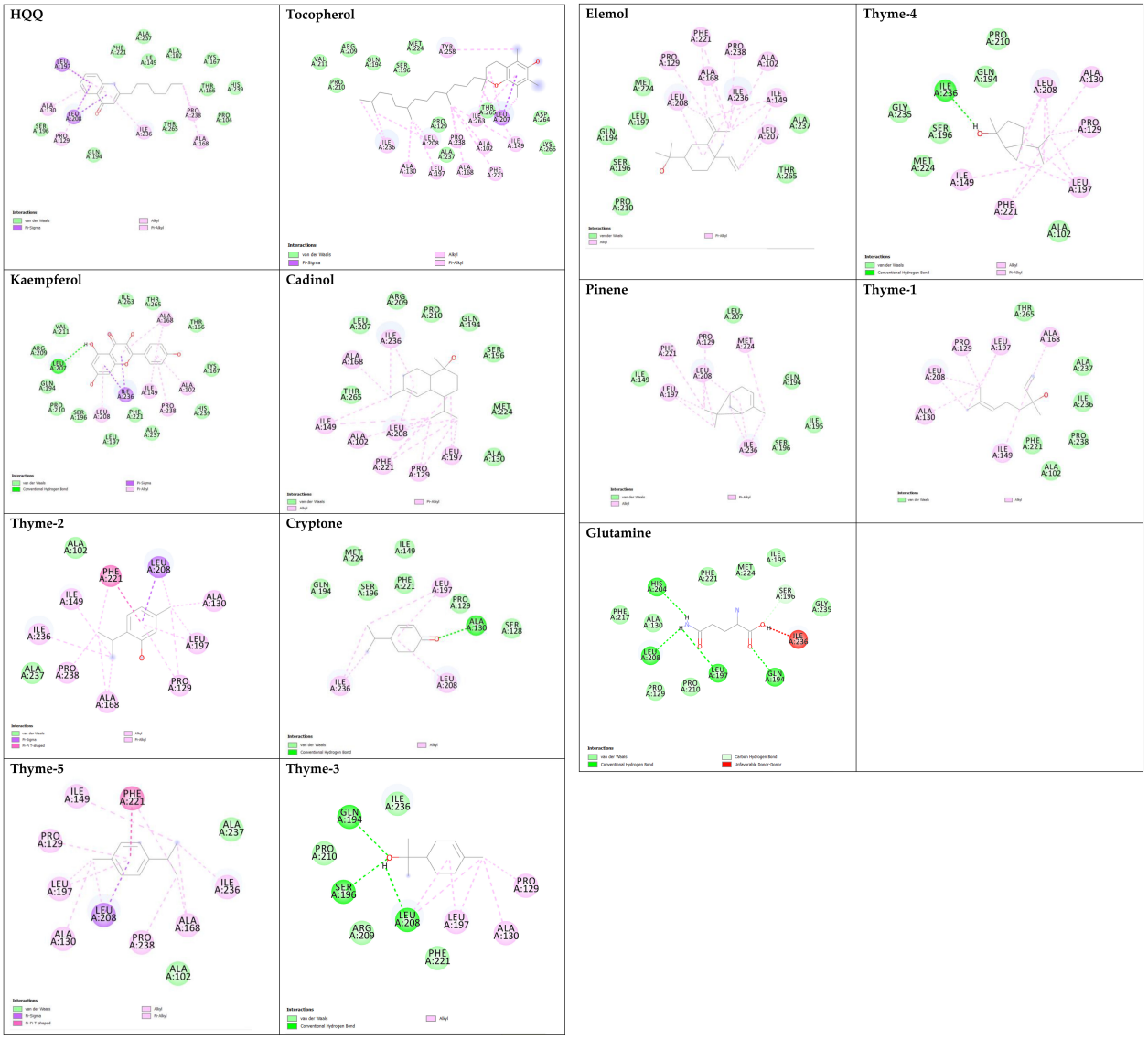


FIGURE 10
2D ligands interaction diagram of the proposed molecules with MvfR.

4 Conclusion

This study showed the composition and the major constituents of the medicinal plants *A. subulatum* and *A. xanthioides* as well as the therapeutic functions associated with them. Both species are considered good sources of antimicrobials including gallic acid, α-pinene, and thyme oil. The most effective antibacterial agents were ethanolic seed extracts of *A. subulatum* against *E. aerogenes*, *P. vulgaris*, *P. aeruginosa*, and *S. typhimurium*, while the antifungal effect was demonstrated by α-pinene oil against *Candida* species. The strong antioxidant activity was associated with tannins and gallic acid. Gallic acid had various medicinal benefits including apoptotic and anti-proliferative effects on cancer cells, potent anti-obesity activity on fatty mice and rats, and an immunosuppressive role in the regulation of skin diseases. The anti-inflammatory potential associated with essential oils, namely, thyme and α-pinene oils, can also be employed in immunosuppressive drugs. Finally, detailed molecular modeling studies indicated that the antimicrobial activities of these plants could be attributed to the high binding affinity of their bioactive compounds to bind to the active sites of the sterol 14-alpha demethylase and the transcriptional regulator MvfR. These findings further emphasize the significance of *A. subulatum* and *A. xanthioides* plants for the formulation and production of novel, less toxic, and more efficient pharmaceuticals in the future. However, clinical and *in vivo* studies are needed to validate the proposed therapeutic applications.

Data availability statement

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

Author contributions

All authors contributed to the article and approved the submitted version.

Funding

This research work was funded by Institutional Fund Projects under grant no. (IFPIP: 621-140-1443). The authors gratefully acknowledge the technical and financial support provided by the Ministry of Education and King AbdulAziz University, DSR, Jeddah, Saudi Arabia.

In memoriam

This article is dedicated to the memory of Prof. Fatimah O. Nassief, former Vice president of female campus at KAU and co-founder of the Prophetic Medicine Chair for her dedicated life in serving science and humanity.

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.2023.1136961/full#supplementary-material

References

AbdElgawad, H., Peshev, D., Zinta, G., Van den Ende, W., Janssens, I. A., and Asard, H. (2014). Climate extreme effects on the chemical composition of temperate grassland species under ambient and elevated CO2: a comparison of fructan and non-fructan accumulators. *PloS One* 9, e92044. doi: 10.1371/journal.pone.0092044

AbdElgawad, H., Saleh, A. M., Al Jaouni, S., Selim, S., Hassan, M. O., Wadaan, M. A. M., et al. (2019). Utilization of actinobacteria to enhance the production and quality of date palm (Phoenix dactylifera l.) fruits in a semi-arid environment. *Sci. Total Environ.* 665, 690–697. doi: 10.1016/j.scitotenv.2019.02.140

Abdel-Wahab BA, F., Abd El-Kareem, H., Alzamami, A., Fahmy, C. H., Elesawy, B., Mostafa Mahmoud, M., et al. (2022). Novel exopolysaccharide from marine bacillus subtilis with broad potential biological activities: Insights into antioxidant, anti-inflammatory, cytotoxicity, and anti-Alzheimer activity. *Metabolites* 12, 715. doi: 10.3390/metabo12080715

Agnihotri, S., and Wakode, S. (2010). Antimicrobial activity of essential oil and various extracts of fruits of greater cardamom. *Indian J. Pharm. Sci.* 72, 657–659. doi: 10.4103/0250-474X.78542

Agnihotri, S. A., Wakode, S. R., and Ali, M. (2012). Chemical composition, antimicrobial and topical anti-inflammatory activity of essential oil of amomum subulatum fruits. *Acta Poloniae Pharm.* 69, 1177–1181.

Alam, K., Pathak, D., and Ansari, S. (2011). Evaluation of anti-inflammatory activity of ammomum subulatum fruit extract. *Int. J. Pharm. Sci. Drug Res.* 3, 35–37.

Alam, A., and Singh, V. (2021). Composition and pharmacological activity of essential oils from two imported amomum subulatum fruit samples. *J. Taibah Univ. Med. Sci.* 16, 231–239. doi: 10.1016/j.jtumed.2020.10.007

Alkandahri, M. Y., Shafirany, M. Z., Rusdin, A., et al. (2021). Amomum compactum: A REVIEW OF PHARMACOLOGICAL STUDIES. *Plant Cell Biotechnol. AND Mol. Biol.* 4, 61–69. doi: 10.9734/bpi/caprd/v4/2775E

Allegretta, G., Maurer, C. K., Eberhard, J., Maura, D., Hartmann, R. W., Rahme, L., et al. (2017). In-depth profiling of MvfR-regulated small molecules in pseudomonas aeruginosa after quorum sensing inhibitor treatment. *Front. Microbiol.* 8, 924. doi: 10.3389/fmicb.2017.00924

Alzahrani, H. R., Kumakli, H., Ampiah, E., Mehari, T., Thornton, A. J., bbyak, C. M., et al. (2017). Determination of macro, essential trace elements, toxic heavy metal concentrations, crude oil extracts and ash composition from Saudi Arabian fruits and vegetables having medicinal values. *Arabian J. Chem.* 10, 906–913. doi: 10.1016/j.arabjc.2016.09.012

Ambreen, M., and Mirza, S. A. (2020). Evaluation of anti-inflammatory and wound healing potential of tannins isolated from leaf callus cultures of achyranthes aspera and ocimum basilicum. *Pakistan J. Pharm. Sci.* 33, 361–369.

Aslam, B., Wang, W., Arshad, M. I., Khurshid, M., Muzammil, S., Rasool, M. H., et al. (2018). Antibiotic resistance: a rundown of a global crisis. *Infection Drug Resist.* 11, 1645. doi: 10.2147/IDR.S173867

Bajpai, V. K., Sharma, A., and Baek, K.-H. (2013). Antibacterial mode of action of cudrania tricuspidata fruit essential oil, affecting membrane permeability and surface characteristics of food-borne pathogens. *Food Control* 32, 582–590. doi: 10.1016/j.foodcont.2013.01.032

Balunas, M. J., and Kinghorn, A. D. (2005). Drug discovery from medicinal plants. *Life Sci.* 78, 431–441. doi: 10.1016/j.lfs.2005.09.012

Bauer, A., and Brönstrup, M. (2014). Industrial natural product chemistry for drug discovery and development. *Natural product Rep.* 31, 35–60. doi: 10.1039/C3NP70058E

Bhaswant, M., Poudyal, H., Mathai, M. L., Ward, L. C., Mouatt, P., and Brown, L. (2015). Green and black cardamom in a diet-induced rat model of metabolic syndrome. *Nutrients* 7, 7691–7707. doi: 10.3390/nu7095360

Bhatia, S. P., Letizia, C. S., and Api, A. M. (2008). Fragrance material review on elemol. *Food Chem. Toxicol.* 46, S147–S148. doi: 10.1016/j.fct.2008.06.045

Borges, A., Ferreira, C., Saavedra, M. J., and Simões, M. (2013). Antibacterial activity and mode of action of ferulic and gallic acids against pathogenic bacteria. *Microb. Drug Resist.* 19, 256–265. doi: 10.1089/mdr.2012.0244

Chau, L. T. M., Thang, T. D., Huong, L. T., and Ogunwande, I. A. (2015). Constituents of essential oils from amomum longiligulare from Vietnam. *Chem. Nat. Compd.* 51, 1181–1183. doi: 10.1007/s10600-015-1525-z

Choi, Y., Choi, J. K., Jang, Y. H., Lee, S., Lee, S., Choi, J. H., et al. (2017). Anti –inflammatory effect of amomum xanthioides in a mouse atopic dermatitis model. *Mol. Med. Rep.* 16, 8964–8972. doi: 10.3892/mmr.2017.7695

Clark, J., and Switzer, R. (1977). Isolation and characterization of phosphatidyl choline from egg yolks. *Exp. Biochem.*, 187–190. doi: 10.1080/10826076.2014.962148

Corsalini, M., Inchingolo, F., Dipalma, G., Wegierska, A. E., Charitos, I. A., Potenza, M. A., et al. (2021). Botulinum neurotoxins (BoNTs) and their biological,

pharmacological, and toxicological issues: A scoping review. *Appl. Sci.* 11, 8849. doi: 10.3390/app11198849

Cowan, M. M. (1999). Plant products as antimicrobial agents. *Clin. Microbiol. Rev.* 12, 564–582. doi: 10.1128/CMR.12.4.564

Davies, J. (1994). Inactivation of antibiotics and the dissemination of resistance genes. *Science* 264, 375–382. doi: 10.1126/science.8153624

De, A. K., and De, M. (2019). "Functional and therapeutic applications of some important spices," in *The Role of Functional Food Security in Global Health*. 2019, 499–510. doi: 10.1016/B978-0-12-813148-0.00029-3

Delesa, D. A. (2018). Traditional medicinal plants for industrial application as natural food preservatives. *Int. J. Adv. Res. Biol. Sci.* 5, 85–94. doi: 10.26655/ijabbr.2018.1.1

Déziel, E., Gopalan, S., Tampakaki, A. P., et al. (2005). The contribution of MvfR to pseudomonas aeruginosa pathogenesis and quorum sensing circuitry regulation: multiple quorum sensing-regulated genes are modulated without affecting lasRI, rhlRI or the production of n-acyl-L-homoserine lactones. *Mol. Microbiol.* 55, 998–1014. doi: 10.1111/j.1365-2958.2004.04448.x

Dhakal, R., Dihingia, A., Ahmed, R. S., Gupta, D. D., Sahu, R. K., Dutta, P., et al. (2022). Prophylactic and therapeutic potential of active phytoconstituents from amomum subulatum roxb. *Food Front.* 4 (1), 60–84. doi: 10.1002/fft2.184

Diao, W.-R., Zhang, L.-L., Feng, S.-S., and Xu, J.-G. (2014). Chemical composition, antibacterial activity, and mechanism of action of the essential oil from amomum kravanh. *J. Food Prot.* 77, 1740–1746. doi: 10.4315/0362-028X.JFP-14-014

Dludla, P. V., Nkambule, B. B., Jack, B., Mkandla, Z., Mutize, T., Silvestri, S., et al. (2018). Inflammation and oxidative stress in an obese state and the protective effects of gallic acid. *Nutrients* 11, 23. doi: 10.3390/nu11010023

El Azab, I. H., Saied, E. M., Osman, A. A., Mehana, A. E., Saad, H. A., and Elkanzi, N. A. (2021). Novel n-bridged pyrazole-1-carbothioamides with potential antiproliferative activity: design, synthesis, *in vitro* and in silico studies. *Future Medicinal Chem.* 13, 1743–1766. doi: 10.4155/fmc-2021-0066

Famiani, F., Bonghi, C., Chen, Z.-H., Drincovich, M. F., Farinelli, D., Lara, M. V., et al. (2020). Stone fruits: Growth and nitrogen and organic acid metabolism in the fruits and seeds–a review. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.572601

Gaber, A., Alsanie, W. F., Kumar, D. N., Refat, M. S., and Saied, E. M. (2020). Novel papaverine metal complexes with potential anticancer activities. *Molecules* 25, 5447. doi: 10.3390/molecules25225447

Gaber, A., Refat, M. S., Belal, A. A. M., El-Deen, I. M., Hassan, N., Zakaria, R., et al. (2021). New mononuclear and binuclear Cu(II), Co(II), Ni(II), and Zn(II) thiosemicarbazone complexes with potential biological activity: Antimicrobial and molecular docking study. *Molecules* 26, 2288. doi: 10.3390/molecules26082288

Gautam, N., Bhattarai, R. R., Khanal, B. K. S., and Oli, P. (2016). Technology, chemistry and bioactive properties of large cardamom (Amomum subulatum roxb.): An overview. *Int. J. Appl. Sci. Biotechnol.* 4, 139–149. doi: 10.3126/ijasbt.v4i2.15104

Grover, J. K., Yadav, S., and Vats, V. (2002). Medicinal plants of India with anti-diabetic potential. *J. Ethnopharmacology* 81, 81–100. doi: 10.1016/S0378-8741(02)00059-4

Hargrove, T. Y., Friggeri, L., Wawrzak, Z., Qi, A., Hoekstra, W. J., Schotzinger, R. J., et al. (2017). Structural analyses of candida albicans sterol 14α-demethylase complexed with azole drugs address the molecular basis of azole-mediated inhibition of fungal sterol biosynthesis. *J. Biol. Chem.* 292, 6728–6743. doi: 10.1074/jbc.M117.778308

Hassan, M. O., Saleh, A. M., and AbdElgawad, H. (2018). Sonchus oleraceus residue improves nutritive and health-promoting value of common bean (Phaseolus vulgaris l.): A metabolic study. *J. Agric. Food Chem.* 66, 2092–2100. doi: 10.1021/acs.jafc.7b05821

Healey, R. D., Saied, E. M., Cong, X., Karsai, G., Gabellier, L., Saint-Paul, J., et al. (2022). Discovery and mechanism of action of small molecule inhibitors of ceramidases**. *Angewandte Chemie Int. Edition* 61, e202109967. doi: 10.1002/anie.202109967

Hussien, Z. N., and Ibrahim, M. M. (2020). Integrated management of onion pink root rot caused by setophoma terrestris. *Middle East J. Appl. Sci.* 10, 703–714. doi: 10.36632/mejas/2020.10.4.62

Im, H.-J., Hwang, S.-J., Lee, J.-S., Lee, S.-B., Kang, J.-Y., and Son, C.-G. (2020). Ethyl acetate fraction of amomum xanthioides ameliorates nonalcoholic fatty liver disease in a high-fat diet mouse model. *Nutrients* 12, 2433. doi: 10.3390/nu12082433

Itkin, M., Heinig, U., Tzfadia, O., Bhide, A. J., Shinde, B., Cardenas, P. D., et al. (2013). Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Science* 341, 175–179. doi: 10.1126/science.1240230

Jakob, E., and Elmadfa, I. (1996). Application of a simplified HPLC assay for the determination of phylloquinone (vitamin K1) in animal and plant food items. *Food Chem.* 56, 87–91. doi: 10.1016/0308-8146(95)00150-6

Kahkeshani, N., Farzaei, F., Fotouhi, M., Alavi, S. S., Bahramsoltani, R., Naseri, R., et al. (2019). Pharmacological effects of gallic acid in health and disease: A mechanistic review. *Iranian J. Basic Med. Sci.* 22, 225–237. doi: 10.22038/ijbms.2019.32806.7897

Kan, C.-C., Chung, T.-Y., Juo, Y.-A., and Hsieh, M.-H. (2015). Glutamine rapidly induces the expression of key transcription factor genes involved in nitrogen and stress responses in rice roots. *BMC Genomics* 16, 731. doi: 10.1186/s12864-015-1892-7

Khirallah, S. M., Ramadan, H. M. M., Aladl, H. A. A., Ayaz, N. O., Kurdi, L. A.F., Jaremko, M., et al. (2022a). Antidiabetic potential of novel 1,3,5-Trisubstituted-2-Thioxoimidazloidin-4-One analogues: Insights into α-glucosidase, α-amylase, and antioxidant activities. *Pharmaceuticals* 15, 1576. doi: 10.3390/ph15121576

Khirallah, S. M., Ramadan, H. M. M., Shawky, A., Qahl, S. H., Baty, R. S., Alqadri, N., et al. (2022b). Development of novel 1,3-Disubstituted-2-Thiohydantoin analogues with potent anti-inflammatory activity; *In vitro* and in silico assessments. *Molecules* 27, 6271. doi: 10.3390/molecules27196271

Kim, H.-G., Han, J.-M., Lee, J.-S., Suk Lee, J., and Son, C.-G. (2015). Ethyl acetate fraction of amomum xanthioides improves bile duct ligation-induced liver fibrosis of rat model *via* modulation of pro-fibrogenic cytokines. *Sci. Rep.* 5, 1–11. doi: 10.1038/srep14531

Konappa, N., Udayashankar, A. C., Krishnamurthy, S., Pradeep, C. K., Chowdappa, S., and Jogaiah, S., et al. (2020). GC–MS analysis of phytoconstituents from amomum nilgiricum and molecular docking interactions of bioactive serverogenin acetate with target proteins. *Sci. Rep.* 10, 16438. doi: 10.1038/s41598-020-73442-0

Kumar, A., Goyal, R., Kumar, S., Jain, S., Jain, N., and Kumar, P. (2015). Estrogenic and anti-alzheimer's studies of zingiber officinalis as well as amomum subulatum roxb.: the success story of dry techniques. *Med. Chem. Res.* 24, 1089–1097. doi: 10.1007/s00044-014-1166-y

Kwon, K.-B., Kim, J.-H., Lee, Y.-R., Lee, H.-K., Jeong, Y.-J., Rho, H.-W., et al. (2003). Amomum xanthoides extract prevents cytokine-induced cell death of RINm5F cells through the inhibition of nitric oxide formation. *Life Sci.* 73, 181–191. doi: 10.1016/s0024-3205(03)00267-4

Lai, L.-R., Hsieh, S.-C., Huang, H.-Y., and Chou, C.-C. (2013). Effect of lactic fermentation on the total phenolic, saponin and phytic acid contents as well as anti-colon cancer cell proliferation activity of soymilk. *J. bioscience bioengineering* 115, 552–556. doi: 10.1016/j.jbiosc.2012.11.022

Lamxay, V., and Newman, M. (2012). A revision of amomum (Zingiberaceae) in cambodia, laos and vietnam. *Edinburgh J. Bot.* 69, 99–206. doi: 10.1017/S0960428611000436

Li, G., Lu, Q., Wang, J., Hu, Q., Liu, P., Yang, Y., et al. (2021). Correlation analysis of compounds in essential oil of amomum tsaoko seed and fruit morphological characteristics, geographical conditions, locality of growth. *Agronomy* 11, 744. doi: 10.3390/agronomy11040744

Lorenzo, J. M., Mousavi Khaneghah, A., Gavahian, M., Marszałek, K., Eş, I., Munekata, P. E., et al. (2019). Understanding the potential benefits of thyme and its derived products for food industry and consumer health: From extraction of value-added compounds to the evaluation of bioaccessibility, bioavailability, anti-inflammatory, and antimicrobial activities. *Crit. Rev. Food Sci. Nutr.* 59, 2879–2895. doi: 10.1080/10408398.2018.1477730

Lv, F., Liang, H., Yuan, Q., and Li, C. (2011). *In vitro* antimicrobial effects and mechanism of action of selected plant essential oil combinations against four food-related microorganisms. *Food Res. Int.* 44, 3057–3064. doi: 10.1016/j.foodres.2011.07.030

Makhija, P., Handral, H. K., Mahadevan, G., Kathuria, H., Sethi, G., and Grobben, B. (2022). Black cardamom (Amomum subulatum roxb.) fruit extracts exhibit apoptotic activity against lung cancer cells. *J. Ethnopharmacology* 287, 114953. doi: 10.1016/j.jep.2021.114953

Marek, C. L., and Timmons, S. R. (2019). "9 - antimicrobials in pediatric dentistry," in *Pediatric dentistry*, *Sixth Edition*. Eds. A. J. Nowak, J. R. Christensen, T. R. Mabry, J. A. Townsend and M. H. Wells (Philadelphia: Elsevier), 128–141.e1. doi: 10.1016/B978-0-323-60826-8.00009-2

Maruszewska, A., and Tarasiuk, J. (2019). Antitumour effects of selected plant polyphenols, gallic acid and ellagic acid, on sensitive and multidrug-resistant leukaemia HL60 cells. *Phytotherapy Res.* 33, 1208–1221. doi: 10.1002/ptr.6317

McGregor, J. C., Quach, Y., Bearden, D. T., Smith, D. H., Sharp, S. E., and Guzman-Cottrill, J. A. (2014). Variation in antibiotic susceptibility of uropathogens by age among ambulatory pediatric patients. *J. Pediatr. Nurs.* 29, 152–157. doi: 10.1016/j.pedn.2013.09.001

Mohamed, D. I., Abou-Bakr, D. A., Ezzat, S. F., El-Kareem, H. F. A., Nahas, H. H. A., Saad, H. A., et al. (2021). Vitamin D3 prevents the deleterious effects of testicular torsion on testis by targeting miRNA-145 and ADAM17: In silico and *In vivo* study. *Pharmaceuticals* 14, 1222. doi: 10.3390/ph14121222

Mohamed, D. I., Alaa El-Din Aly El-Waseef, D., Nabih, E. S., El-Kharashi, O. A., Abd El-Kareem, H. F., Abo Nahas, H. H., et al. (2022a). Acetylsalicylic acid suppresses alcoholism-induced cognitive impairment associated with atorvastatin intake by targeting cerebral miRNA155 and NLRP3: *In vivo*, and in silico study. *Pharmaceutics* 14, 529. doi: 10.3390/pharmaceutics14030529

Mohamed, D. I., Ezzat, S. F., Elayat, W. M., El-Kharashi, O. A., El-Kareem, H. F. A., Nahas, H. H. A., et al. (2022b). Hepatoprotective role of carvedilol against ischemic hepatitis associated with acute heart failure *via* targeting miRNA-17 and mitochondrial dynamics-related proteins: An *In vivo* and in silico study. *Pharmaceuticals* 15, 832. doi: 10.3390/ph15070832

Monica, B., and Ioan, S. (2018). Essential oils from plants. *J. Biotechnol. Biomed. Sci.* 1, 35–43. doi: 10.14302/issn.2576-6694.jbbs-18-2489

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791. doi: 10.1002/jcc.21256

Mueller-Harvey, I. (2006). Unravelling the conundrum of tannins in animal nutrition and health. *J. Sci. Food Agric.* 86, 2010–2037. doi: 10.1002/jsfa.2577

Myint, P. P., Thu, K., Win, K. C., Win, N. N., and Ngwe, D. H. (2012). A study of antidiabetic and antioxidant activities of amomum xanthioides Wall.(Chinbaung hpalar) seed. *Universities Res. J.* 5(4), 19–32.

Nóbrega, J. R., Silva D de, F., de Andrade Júnior, F. P., de Sousa, P. M.S., Figueiredo, P. T.R., Cordeiro, L. V., et al. (2021). Antifungal action of α-pinene against candida spp. isolated from patients with otomycosis and effects of its association with boric acid. *Natural Product Res.* 35, 6190–6193. doi: 10.1080/14786419.2020.1837803

Oh, L., Nj, R., Al, F., and Rj, R. (1951). Protein measurement with the folin phenol reagent. *J. Biol. Chem.* 193(1), 265–275.

Ohkawa, H., Ohishi, N., and Yagi, K. (1979). Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction. *Analytical Biochem.* 95, 351–358. doi: 10.1016/0003-2697(79)90738-3

Paciolla, C., Fortunato, S., Dipierro, N., Paradiso, A., De Leonardis, S., Mastropasqua, L., et al. (2019). Vitamin c in plants: From functions to biofortification. *Antioxidants (Basel)* 8, 519. doi: 10.3390/antiox8110519

Pan, X., Xiao, H., Hu, X., and Liu, Z. L. (2022). Insecticidal activities of the essential oil of rhynchanthus beesianus rhizomes and its constituents against two species of grain storage insects. *Z. für Naturforschung C.* 78(1-2), 83–89.

Pandey, A. K., Kumar, P., Singh, P., Tripathi, N. N., and Bajpai, V. K. (2017). Essential oils: Sources of antimicrobials and food preservatives. *Front. Microbiol.* 7, 2161. doi: 10.3389/fmicb.2016.02161

Park, M., Cho, H., Jung, H., Lee, H., and Hwang, K. T. (2014). Antioxidant and anti-inflammatory activities of tannin fraction of the extract from black raspberry seeds compared to grape seeds. *J. Food Biochem.* 38, 259–270. doi: 10.1111/jfbc.12044

Pešić, M., and Stanković, S. (2015). Development of natural product drugs in a sustainable manner. *Brief United Nations Global Sustain. Dev. Rep.*

Prado-Audelo, D., Luisa, M., Cortés, H., Caballero-Florán, I. H., González-Torres, M., Escutia-Guadarrama, L., et al. (2021). Therapeutic applications of terpenes on inflammatory diseases. *Front. Pharmacol.* 2114. doi: 10.3389/fphar.2021.704197

Prasathkumar, M., Anisha, S., Dhrisya, C., Becky, R., and Sadhasivam, S. (2021). Therapeutic and pharmacological efficacy of selective Indian medicinal plants – a review. *Phytomedicine* Plus 1, 100029. doi: 10.1016/j.phyplu.2021.100029

Rachkeeree, A., Kantadoung, K., Suksathan, R., Puangpradab, R., Page, P. A., and Sommano, S. R. (2018). Nutritional compositions and phytochemical properties of the edible flowers from selected zingiberaceae found in Thailand. *Front. Nutr.* 5. doi: 10.3389/fnut.2018.00003

Rafieian-Kopaei, M. (2012). Medicinal plants and the human needs. *J. HerbMed Pharmacol.* 1(1), 1–2.

Roba'ei, D. R. M., Osman, W. N. W., Zainudin, H. Z. M., Aziz, M. Y. A., and Johari, S. A. T. T. (2022). Anti-proliferative and apoptosis inducing effects of Gallic acid on human acute myeloid leukaemia cell lines (HL-60). *Asian J. Med. Biomedicine* 6, 160–162. doi: 10.37231/ajmb.2022.6.S1.572

Saied, E. M., El-Maradny, Y. A., Osman, A. A., Darwish, A. M. G., Abo Nahas, H. H., Niedbała, G., et al. (2021). A comprehensive review about the molecular structure of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): Insights into natural products against COVID-19. *Pharmaceutics* 13, 1759. doi: 10.3390/pharmaceutics13111759

Saleem, M., Deters, B., de la Bastide, A., and Korzen, M. (2019). Antibiotics overuse and bacterial resistance. *Ann. Microbiol. Res.* 3, 93.

Saleh, A. M., and Madany, M. M. Y. (2015). Coumarin pretreatment alleviates salinity stress in wheat seedlings. *Plant Physiol. Biochem.* 88, 27–35. doi: 10.1016/j.plaphy.2015.01.005

Salehi, B., Mishra, A. P., Shukla, I., Sharifi-Rad, M., Contreras, M. M., Segura-Carretero, A., et al. (2018). Thymol, thyme, and other plant sources: Health and potential uses. *Phytotherapy Res.* 32, 1688–1706. doi: 10.1002/ptr.6109

Salehi, B., Upadhyay, S., Erdogan Orhan, I., Kumar Jugran, A., Jayaweera, L. D S., Dias, A. D., et al. (2019). Therapeutic potential of α-and β-pinene: A miracle gift of nature. *Biomolecules* 9, 738. doi: 10.3390/biom9110738

Samaha, D., Hamdo, H. H., Cong, X., Schumacher, F., Banhart, S., Aglar, Ö., et al. (2020). Liposomal FRET assay identifies potent drug-like inhibitors of the ceramide transport protein (CERT). *Chem. – A Eur. J.* 26, 16616–16621. doi: 10.1002/chem.202003283

Santos, E. F., Mateus, N. S., Rosário, M. O., Garcez, T. B., Mazzafera, P., and Lavres, J. (2021). Enhancing potassium content in leaves and stems improves drought tolerance of eucalyptus clones. *Physiologia Plantarum* 172, 552–563. doi: 10.1111/ppl.13228

Sarjit, A., Wang, Y., and Dykes, G. A. (2015). Antimicrobial activity of gallic acid against thermophilic campylobacter is strain specific and associated with a loss of calcium ions. *Food Microbiol.* 46, 227–233. doi: 10.1016/j.fm.2014.08.002

Sarungallo, Z. L., Hariyadi, P., Andarwulan, N., Purnomo, E. H., and Wada, M. (2015). Analysis of α-cryptoxanthin, β-cryptoxanthin, α -carotene, and β-carotene of pandanus conoideus oil by high-performance liquid chromatography (HPLC). *Proc. Food Sci.* 3, 231–243. doi: 10.1016/j.profoo.2015.01.026

Satyal, P., Dosoky, N. S., Kincer, B. L., and Setzer, W. N. (2012). Chemical compositions and biological activities of amomum subulatum essential oils from Nepal. *Nat. Prod. Commun.* 7, 1233–1236. doi: 10.1177/1934578X1200700935

Sett, R., and Soni, B. (2013). Foliar nitrogen, phosphorus and potassium content in trees in environmentally toxic plastic industry area. *J. Environ. Sci. Eng.* 55, 167–174.

Sharma, K., Kumar, V., Kaur, J., Tanwar, B., Goyal, A., Sharma, R., et al. (2021). Health effects, sources, utilization and safety of tannins: A critical review. *Toxin Rev.* 40, 432–444. doi: 10.1080/15569543.2019.1662813

Sharma, V., Lohia, N., Handa, V., and Baranwal, M. (2017). Amomum subulatum seed extract exhibit antioxidant, cytotoxic and immune-suppressive effect. *Indian J. Biochem. Biophys.* 54, 135–139.

Shukla, S., Singh, B., Singh, A., and Singh, C. (2022). Emerging and advanced drug delivery systems for improved biopharmaceutical attributes of gallic acid: A review. *Phytomedicine* Plus 2, 100369. doi: 10.1016/j.phyplu.2022.100369

Soto-Maldonado, C., Vergara-Castro, M., Jara-Quezada, J., Caballero-Valdés, E., Müller-Pavez, A., Zúñiga-Hansen, M. E., et al. (2019). Polyphenolic extracts of walnut (Juglans regia) green husk containing juglone inhibit the growth of HL-60 cells and induce apoptosis. *Electronic J. Biotechnol.* 39, 1–7. doi: 10.1016/j.ejbt.2019.02.001

Sreevidya, N., and Mehrotra, S. (2003). Spectrophotometric method for estimation of alkaloids precipitable with dragendorff's reagent in plant materials. *J. AOAC Int.* 86, 1124–1127. doi: 10.1093/jaoac/86.6.1124

Sudarsanan, D., Suresh Sulekha, D., and Chandrasekharan, G. (2021). Amomum subulatum induces apoptosis in tumor cells and reduces tumor burden in experimental animals *via* modulating pro-inflammatory cytokines. *Cancer Invest.* 39, 333–348. doi: 10.1080/07357907.2021.1878529

Tampe, J., Espinoza, J., Chacón-Fuentes, M., Quiroz, A., and Rubilar, M. (2020). Evaluation of drimys winteri (Canelo) essential oil as insecticide against acanthoscelides obtectus (Coleoptera: Bruchidae) and aegorhinus superciliosus (Coleoptera: Curculionidae). *Insects* 11, 335. doi: 10.3390/insects11060335

Tandukar, U., Lakhey, P. B., Shrestha, N., Dhungana, A., and Bhatt, T. D. (2018). Antimicrobial activity and chemical composition of essential oil of fruits of amomum subulatum roxb. *J. OF Plant Resour.* 16, 57.

Thinh, B. B., Chac, L. D., Hanh, D. H., Korneeva, A. A., Hung, N., and Igoli, J. O. (2022). Effect of extraction method on yield, chemical composition and antimicrobial activity of essential oil from the fruits of amomum villosum var. xanthioides. *J. Essential Oil Bearing Plants* 25, 28–37. doi: 10.1080/0972060X.2022.2049893

Thinh, B. B., Doudkin, R. V., and Thanh, V. Q. (2021). Chemical composition of essential oil of amomum xanthioides wall. ex baker from northern Vietnam. *Biointerface Res. Appl. Chem.* 11, 12275–12284. doi: 10.33263/BRIAC114.1227512284

Thompson, T. (2022). The staggering death toll of drug-resistant bacteria. *Nature.* doi: 10.1038/d41586-022-00228-x

Tijjani, M., Abdulrahman, F., and Sandabe, U. (2011). Chemical composition and anti-nociceptive effects of the leaf extract of vitexdoniana sweet. *J. Chem. Soc. Nigeria* 36, 213–219.

Tijjani, M. A., Dimari, G. A., Buba, S. W., and Khan, I. Z. (2012). In-vitro antibacterial properties and pre-liminary phytochemical analysis of amomum subulatum roxburg (Large cardamom). *J. Appl. Pharm. Sci.*, 69–73. doi: 10.7324/JAPS.2012.2511

Tong, Z., He, W., Fan, X., and Guo, A. (2021). Biological function of plant tannin and its application in animal health. *Front. Veterinary Sci.* 8.

Trott, O., and Olson, A. J. (2010). AutoDock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461. doi: 10.1002/jcc.21334

Tsang, M. S., Jiao, D., Chan, B. C., Hon, K.-L., Leung, P. C., Lau, C. B., et al. (2016). Anti-inflammatory activities of pentaherbs formula, berberine, gallic acid and chlorogenic acid in atopic dermatitis-like skin inflammation. *Molecules* 21, 519. doi: 10.3390/molecules21040519

Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *Pharm. Ther.* 40, 277.

Wangrawa, D. W., Kosgei, J., Machani, M., Opala, J., Agumba, S., Yaméogo, F., et al. (2022). Larvicidal activities and synergistic effects of essential oils against anopheles funestus and culex quinquefasciatus (Diptera: Culicidae) from kisumu, Kenya. *Psyche: A J. Entomology.* doi: 10.1155/2022/8302696

WHO. *Global Centre for Traditional Medicine in India (n.d.).* Avaialble at: https://www.who.int/news/item/25-03-2022-who-establishes-the-global-centre-for-traditional-medicine-in-india (Accessed November 6, 2022).

Xie, L., Yu, D., Li, Y., Ju, H., Chen, J., Hu, L., et al. (2022). Characterization, hypoglycemic activity, and antioxidant activity of methanol extracts from amomum tsao-ko: *in vitro* and *in vivo* studies. *Front. Nutr.* 9. doi: 10.3389/fnut.2022.869749

Yahaya, A. A. H., Salleh, W. M. N. H. W., Ab Ghani, N., Khamis, S., Rezali, N. S., and Juhari, M. A. A. (2022). Chemical composition and bioactivities of magnolia candollii h. keng essential oil. *Z. für Naturforschung C* 77, 519–523. doi: 10.1515/znc-2022-0100

Yousaf, M., Bashir, S., Raza, H., Shah, A. N., Iqbal, J., Arif, M., et al. (2021). Role of nitrogen and magnesium for growth, yield and nutritional quality of radish. *Saudi J. Biol. Sci.* 28, 3021–3030. doi: 10.1016/j.sjbs.2021.02.043

Yousefzadi, M., Sonboli, A., Karimi, F., Ebrahimi, S. N., Asghari, B., and Zeinali, A. (2007). Antimicrobial activity of some salvia species essential oils from Iran. *Z. für Naturforschung C* 62, 514–518. doi: 10.1515/znc-2007-7-809

Zarzuelo, A., and Crespo, E. (2002). The medicinal and non-medicinal uses of thyme. *In: Thyme. CRC Press* pp, 277–306.

Zender, M., Witzgall, F., Kiefer, A., Kirsch, B., Maurer, C. K., Kany, A. M., et al. (2020). Flexible fragment growing boosts potency of quorum-sensing inhibitors against pseudomonas aeruginosa virulence. *ChemMedChem* 15, 188–194. doi: 10.1002/cmdc.201900621

Zhang, J., Li, L., Lv, Q., Yan, L., Wang, Y., and Jiang, Y. (2019). The fungal CYP51s: Their functions, structures, related drug resistance, and inhibitors. *Front. Microbiol.* 10, 691. doi: 10.3389/fmicb.2019.00691

Zhang, Z.-W., Yang, X.-Y., Zheng, X.-J., Fu, Y.-F., Lan, T., Tang, X.-Y., et al. (2020). Vitamin e is superior to vitamin c in delaying seedling senescence and improving resistance in arabidopsis deficient in macro-elements. *Int. J. Mol. Sci.* 21, 7429. doi: 10.3390/ijms21197429

frontiers | Frontiers in Pharmacology

TYPE Review
PUBLISHED 17 July 2023
DOI 10.3389/fphar.2023.1214881

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Ahmad Faizal Abdull Razis,
University Putra Malaysia, Malaysia
Valentina Citi,
University of Pisa, Italy

*CORRESPONDENCE
Jaya Arora,
jaya890@gmail.com,
jayaarora@mlsu.ac.in
Vinoth Kumarasamy,
vinoth@ukm.edu.my

†PRESENT ADDRESS
Vetriselvan Subramaniyan,
Jeffrey Cheah School of Medicine and Health Sciences, Monash University, Petaling Jaya, Selangor, Malaysia

RECEIVED 14 May 2023
ACCEPTED 04 July 2023
PUBLISHED 17 July 2023

CITATION
Khan F, Joshi A, Devkota HP, Subramaniyan V, Kumarasamy V and Arora J (2023), Dietary glucosinolates derived isothiocyanates: chemical properties, metabolism and their potential in prevention of Alzheimer's disease.
Front. Pharmacol. 14:1214881.
doi: 10.3389/fphar.2023.1214881

Dietary glucosinolates derived isothiocyanates: chemical properties, metabolism and their potential in prevention of Alzheimer's disease

Farhana Khan[1], Abhishek Joshi[1], Hari Prasad Devkota[2], Vetriselvan Subramaniyan[3†], Vinoth Kumarasamy[4*] and Jaya Arora[1*]

[1]Laboratory of Bio-Molecular Technology, Department of Botany, Mohanlal Sukhadia University, Udaipur, Rajasthan, India, [2]Graduate School of Pharmaceutical Sciences, Kumamoto University, Kumamoto, Japan, [3]Department of Pharmacology, Center for Transdisciplinary Research, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India, [4]Department of Parasitology and Medical Entomology, Faculty of Medicine, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

Alzheimer's disease (AD) is the most prevalent form of dementia affecting millions of people worldwide. It is a progressive, irreversible, and incurable neurodegenerative disorder that disrupts the synaptic communication between millions of neurons, resulting in neuronal death and functional loss due to the abnormal accumulation of two naturally occurring proteins, amyloid β (Aβ) and tau. According to the 2018 World Alzheimer's Report, there is no single case of an Alzheimer's survivor; even 1 in 3 people die from Alzheimer's disease, and it is a growing epidemic across the globe fruits and vegetables rich in glucosinolates (GLCs), the precursors of isothiocyanates (ITCs), have long been known for their pharmacological properties and recently attracted increased interest for the possible prevention and treatment of neurodegenerative diseases. Epidemiological evidence from systematic research findings and clinical trials suggests that nutritional and functional dietary isothiocyanates interfere with the molecular cascades of Alzheimer's disease pathogenesis and prevent neurons from functional loss. The aim of this review is to explore the role of glucosinolates derived isothiocyanates in various molecular mechanisms involved in the progression of Alzheimer's disease and their potential in the prevention and treatment of Alzheimer's disease. It also covers the chemical diversity of isothiocyanates and their detailed mechanisms of action as reported by various *in vitro* and *in vivo* studies. Further clinical studies are necessary to evaluate their pharmacokinetic parameters and effectiveness in humans.

KEYWORDS

Alzheimer's disease, isothiocyanates, amyloid β, phosphorylated tau, glucosinolates

Abbreviations: AD, Alzheimer's disease; AITC, allylisothiocyanate; APP, amyloid precursor protein; Aβ, amyloid beta; BBB, blood brain barrier; ER, erucin; GLCs, glucosinolates; I3C, indole 3 carbinol; IL-1β, interleukin-1β; iNOS, inducible nitric oxide synthase; ITCs, isothiocyanates; MAPK, mitogen activated protein kinase; MO, *Moringa oleifera;* NF-kβ, nuclear factor kappa B; NFT, neurofibrillary tangles; PEITC, phenethylisothiocyanate; p-tau, phosphorylated tau; ROS, reactive oxygen species; SFN, sulforaphan; TNF-α, tumor necrosis factor- α.

1 Introduction

In the past few decades, owing to healthy habits and general improvements in lifestyle and medication, life expectancy has substantially increased; however, the prominent upward shift in age distribution has increased the prevalence of chronic diseases, including Alzheimer's disease (AD). AD slowly affects the brain and exhibits clear pathological changes in the hippocampus, the centre of memory and learning (Zhang et al., 2020). In AD, the propensity of neurotoxic proteins to form template or oligomers is higher and accelerates the conversion and aggregation of endogenous proteins, which eventually convert into fibrils (Schaffert and Carter, 2020). It can be sporadic or familial and AD cases are sporadic in most instances (Dorszewska et al., 2016). Disease modifying treatments primarily focused on reducing amyloid beta (senile plaques, Aβ) and tau (neurofibrillary tangles) load in the brain (Cammisuli et al., 2022). Despite many costly clinical trials ranging from pharmacological to hormonal treatments and immunotherapy, not even a single drug produced clinically significant results due to suboptimal dosing of drugs, unavailability of reliable biomarkers for early diagnosis and more specifically lack of detailed mechanistic approaches (Lashley et al., 2018; Loewenstein, 2022). The existing medication exert only moderate reduction of symptoms; therefore, AD remains symptomatic and can be controlled and prevented but uncured (Fernández and Ribeiro, 2018).

According to the World Alzheimer Report (2018), there are 50 million people living with dementia worldwide, of which 70–80 percent are AD patients, and by 2050 these numbers will be more than triple to 152 million (Patterson, 2018). From the data provided by the World Health Organization (WHO), it is an epidemic worldwide and has become a global burden (Cao et al., 2020). Death from AD has increased 123 percent between 2000–2005 and more than 60 percent cases are from low to middle income countries (Patterson, 2018). At the beginning of 21st century, AD remains a major biomedical challenge. Pharmaceutical companies and neurobiologists around the world are doing their efforts to develop novel FDA approved drugs such as acetyl cholinesterase (AChE) inhibitors (Donepezil, Rivastigmine and Galantamine) and NMDA (n-methyl D aspartate) receptor antagonist (Memantine) but they showed several side effects in phase II and III clinical trials. Common adverse effects of AChE inhibitors are diarrhea, nausea, vomiting, bradycardia, muscle twitching nightmares, etc., and memantine includes dizziness, headache, and lethargy (Ettcheto et al., 2018; Schneider, 2022).

The discovery of new natural pharmacologically active compounds is a widely growing field, as the synthesis of most the biomolecules is tough task (Ramawat and Arora, 2021). Consumption of antioxidant rich food and vegetables might improve brain function, minimize the possibilities of cognitive impairment, retard the process of aging, subsequent oxidation, and disease progression (Andrade et al., 2019). It is clinically proven that they enhance cellular metabolism and nourish brain cells; this safeguarding impact is more potent when isothiocyanates (ITCs) rich fruits and vegetables are specifically consumed (Esteve, 2020; Kamal et al., 2022). The propitious attributes of fruits and vegetables are related to their nutritional and functional components like minerals, vitamins, antioxidants and polyphenols. All of these molecules are found in cruciferous vegetables, however, the sulfurous compound GLCs that give them their distinctive pungent aroma and flavour set them apart. GLCs are stable chemically but biologically inactive and remain sequestered within plant compartment (Verkerk et al., 2009; Alexandre et al., 2020). Tissue damage and chewing are the main causes that lead to the formation of biologically active derivatives of GLCs such as ITCs by enzyme hydrolysis, which directly and indirectly regulate their activity and have been demonstrated to exert neuroprotective properties through multiple mechanisms (Tian et al., 2018).

Generally, there are three major hypothesis, i.e., AChE, amyloid, and tau, which are primarily implicated in Alzheimer's disease management and prevention. Beside them, neuroinflammation is another important response target involving biochemical events activating resident cells of the central nervous system (CNS), which may induce the entire process of AD. It is initiated by aberrant astrocytes and microglial activation, which leads to the release of different inflammatory mediators such as nitric oxide (NO), prostaglandin E2 (PGE-2), reactive oxygen species (ROS), cytokines and chemokines (Kraft and Harry, 2011). Furthermore, it elevates the level of proinflammatory cytokines such as tumor necrosis factor (TNF-α), interleukin-1β (IL-1β) and interleukin-6 (IL-6), which are responsible for neuronal death (Xia et al., 2015). Controlling microglia and astrocytes activation can therefore be a therapeutic approach in the prevention and management of AD. Recently, it has been shown that ITCs possess neuroprotective effects through the modulation of different signalling pathways (Latronico et al., 2021). In oxidative stress and inflammation control, nuclear factor-kβ (NF-kβ) and nuclear erythroid related factor 2 (Nrf2) are two main regulators (Fão et al., 2019). They may primarily be attributed to their peculiar ability to activate the Nrf2/ARE pathway (Giacoppo et al., 2015). ITCs significantly decrease NF-kβ translocation with the inhibition of proinflammatory cytokines (Latronico et al., 2021). Hydrogen sulphide (H_2S) is another important signal molecule in CNS; it could represent an intriguing strategy for the treatment of neurodegenerative diseases (Tabassum and Jeong, 2019; Sharif et al., 2023). Beside this, it also play a key role in many aspects of human health like in antiproliferation, cardioprotection, chemoprevention, etc. (Martelli et al., 2020). It also interacts with redox system regulating cellular oxidative stress and ROS (Kabil and Banerjee, 2010). There is a strong relationship between H_2S and aging, as consistent significant decline of H_2S levels has been observed in AD patients (Disbrow et al., 2021). H_2S is a relevant player accounting for different biophysiological effects of Brassicaceae plants, for example, Allyl isothiocyanate (AITC) from black mustard (*B. nigra*), benzyl-ITC from garden cress (*Lepidium sativum*), erucin form *Eruca* sp., *B. oleirecia*, etc. and 4-hydorxybenzyl-ITC from white mustard (*B. alba*) are some important naturally occurring ITCs. Among these selected ITCs, benzyl ITC is the most effective H_2S donor, exhibiting remarkable H_2S release followed by AITC (Citi et al., 2014). Recently, available literature clearly demonstrated that the role of natural ITCs as H_2S donor (Martelli et al., 2020). It is a pleiotropic mediator that affects different element in inflammatory cascade specially NF-kβ and Nrf2 signalling (Zhao et al., 2023).

Another important effect of ITCs is apoptotic suppression as they can intervene and arrest the mitochondrial apoptotic pathway (Dinkova-Kostova and Kostov, 2012). Deposition of Aβ and hyperphosphorylated tau proteins is a crucial event in AD as pathology several studies demonstrated the pharmacological potencies of ITCs against these two hallmarks and their toxicity by intervene in its cascade such as APP cleavage, BACE1 expression,

oligomerization of seeded proteins, phosphorylation and dephosphorylation assembly, etc. (Morroni et al., 2018; Asif et al., 2022). ITCs could therefore be considered as a promising source of medicine and for the treatment and management of AD. This review focuses on the knowledge regarding the direct and indirect mechanisms by which GLCs-derived ITCs intervene in inhibition of AChE, neurotoxic proteins (Aβ and tau) and neuroinflammation cascade.

2 Glucosinolates (GLCs) and isothiocyanates (ITCs)

2.1 Sources from foods

Glucosinolates (GLCs), a group of sulphur containing glycosides and their hydrolysis products, i.e., isothiocyanates (ITCs) are abundantly found in the family Brassicaceae which encompasses our daily vegetables including cabbage, broccoli, mustard, white radish, radish, kale, turnip, oilseed rape, collard greens, daikon, kohlrabi, wasabi, cauliflower, Brussels, etc. (Cancer et al., 2004; Shree et al., 2022). These metabolites distinguish them from other plant families and are responsible for pungent smell and bitter taste (Verkerk et al., 1998; Barba et al., 2016). Besides this, they are also found in *Moringa oleifera* (drumstick tree), a plant from the family Moringaceae; in contrast with other Brassicaceae plants, only aromatic GLCs have been identified in *M. oleifera* (Lopez-Rodriguez et al., 2020). More than 200 GLCs have already been characterized so far, although a small number of these compounds are present in closely related taxonomic groups and not all are present in plants that people consume (Fahey et al., 2001; Agerbirk and Olsen, 2012). Its content varies between different cultivars and plant species even in plant parts such as seeds, stems, roots, and leaves, while the highest amount is present in young tissues (Blažević and Mastelić, 2009). These variations arise from several factors (genetic, nutrient and environmental) and growth conditions (temperature, nutrient availability and water content).

2.2 Chemical properties

GLCs are structurally thiohydroximates containing S-linked β-glucopyranosyl and O-linked sulfate residues with different side chains derived from amino acids (Agerbirk and Olsen, 2012). They are synthesized by different amino acid precursors such as phenylalanine, tryptophan, and methionine, which give rise to molecules with side chain R (Table 1; Figure 1). All known GLCs display structural homogeneity with different R groups in producing their corresponding ITCs responsible for various biological activities (Agerbirk and Olsen, 2012). On the basis of their side chain they are characteristically subdivided into three groups (Ali et al., 2018; Huke et al., 2021) as shown in Table 1: i) long chain length aliphatic; ii) short to medium chain length aliphatic (only C3 and C3 or C4 with C5) and iii) simple aryl aliphatic such as benzyl, phenyl, hydroxybenzyl GLCs; highly substituted aryl aliphatic such as dihydroxy, dimethoxy and trimethoxy benzyl GLCs. C3-C5 aliphatic GLCs are commonly found in *Brassica* species (Bennett et al., 2004).

ITCs are a specific type of compound derived from the hydrolysis of GLCs along with nitriles and thiocyanates. The entire conversion is catalyzed by endogenous myrosinase (thioglucoside glucohydrolase) enzyme released after chopping and chewing of raw vegetables or physical damage such as insect attack (Oliviero et al., 2018). Myrosinase reacts

TABLE 1 Trivial name, side chain structure and dietary plant source of Glucosinolates and Isothiocyanates.

GLCs trivial name	ITCs trivial name	Side chain name and structure of R group	Main dietary source
Aliphatic group			
Sinigrin (Glucobrassicin)	Allyl ITC	CH_2 = CH-CH_2-2-Propenyl	Cabbage, horseradish, wasabi, mustard Cartea and Velasco. (2008)
Glucoerucin	Erucin	CH_3-S-CH_2-CH_2-CH_2-CH_2-4-Methyl thiobutyl	Turnip, kohlrabi, arugula, broccoli seeds Avato and Argentieri. (2015)
Glucoraphanin	Sulforaphane	CH_3-SO-CH_2-CH_2-CH_2-CH_2-4-Methylsulphinylbutyl	Broccoli, cauliflower, kale, brussels sprout, cabbage Fahey et al. (2001)
Glucoraphenin	Sulforaphane	CH_3-SO-CH = CH-CH_2-CH_2-4-Methylsulfinyl-3-butenyl	Radish, brussels sprout Fahey et al. (2001); Avato and Argentieri. (2015)
Glucoraphasatin	*Raphasatin*	CH_3-S-CH = CH-CH_2-CH_2-4- methylsulfanyl 3-butenyl	Japanese Daikon Jaafaru et al. (2019b)
Glucoiberin	Iberin	CH_3-SO-CH_2-CH_2-CH_2-3-Methylsulfinylpropyl	Broccoli, cabbage Fahey et al. (2001); Cancer et al. (2004)
Aromatic group			
Glucotropaeolin	Benzyl ITC	C_6H_5-CH_2-Benzyl	Wasabi and mustard Mithen et al. (2000); Verkerk et al. (2009)
Gluconasturtiin	Phenylethyl ITC	C_6H_5-$(CH_2)_2$-2-Phenylethyl	Watercress, radish, turnips, broccoli, kale Cartea and Velasco. (2008)
Glucomoringin	Moringin	$C_{13}H_{15}O_5^-$	Drumstick tree Lopez-Rodriguez et al. (2020)
Indolyl ITC			
Indol-3-yl-methylglucosinolate	Indole 3-carbinol	C_8H_6N-CH_2OH 1H-Indol-3-yl-methanol	All vegetables Amarakoon et al. (2023)

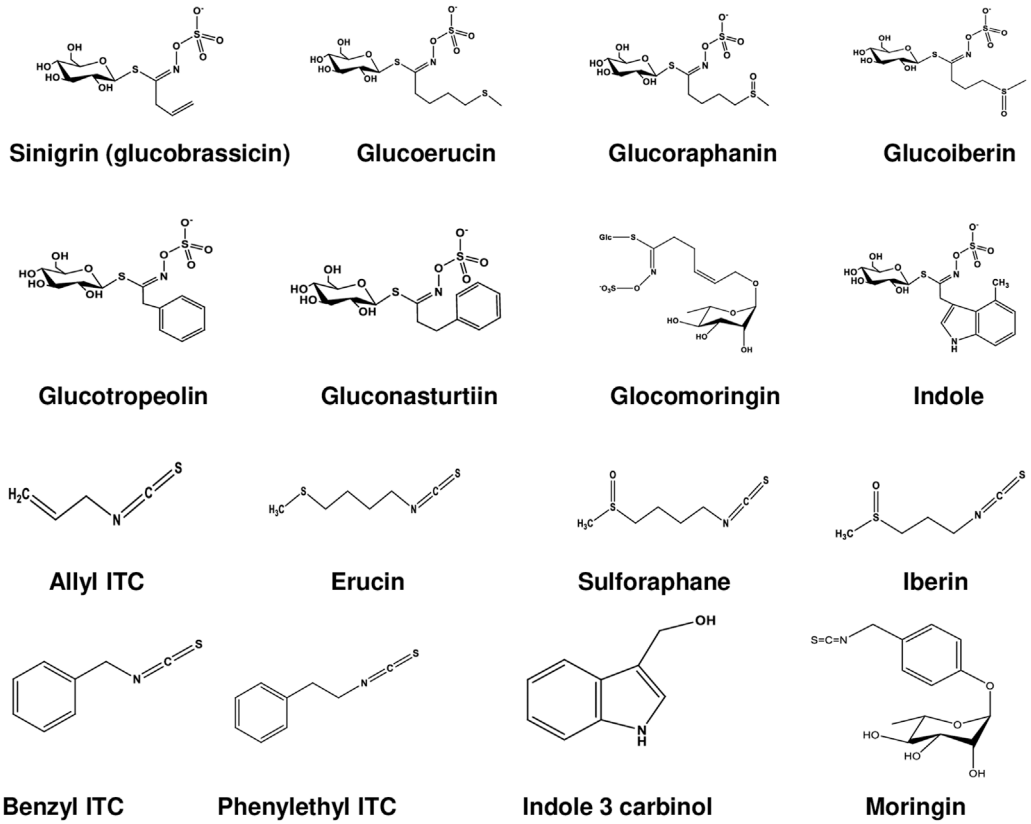


FIGURE 1
Chemical structures of glucosinolates and isothiocyanates.

with GLCs by hydrolytically cleaving thio-linked glucose and forms active ITCs by an unstable intermediate thiohydroximate-O-sulfonate after immediate rearrangement depending on the corresponding substrate (GLCs), pH, temperature, epithiospecifier proteins (ESP), ferrous ions and thiocyanate forming proteins (TFP) (Sikorska-Zimny and Beneduce, 2021) as shown in Figure 2. Extraction and isolation of GLCs and their hydrolysis product ITCs are still challenging due to their sensitive nature. In recent years, different methods have been developed for the detection and quantification of GLCs and ITCs, mainly UHPLC-DAD-ESI-MS and HPLC-DAD-ESI-MS for GLCs (Devkota, 2020) and HPLC-DAD and UHPLC-HRMS/MS for ITCs (Karanikolopoulou et al., 2021). If myrosinase is denatured during ingestion, GLCs metabolism can also be triggered by gut microbiota (Luang-In et al., 2014). In such conditions, GLCs are absorbed in the stomach and then transit to the small intestine and colon where they hydrolyzed by microbiota (Barba et al., 2016). Long cooking time and high cooking temperature (>80°C) triggered myrosinase denaturation and significant GLCs and ITCs loss (more than 90%), but after ingestion, gut bacteria promote the conversion of GLCs into ITCs, which are then absorbed; therefore, a preferable method is steaming over boiling the raw food to minimize metabolite loss (Barba et al., 2016; Shakour et al., 2022).

2.3 Bioavailability of GLCs and ITCs

Bioavailability is an essential parameter that determines the action of metabolites. It represents absorption, distribution, metabolism, and excretion unlike drugs, where the oral concentration is predetermined. It depends on the number of food products, which is highly variable (Gupta and Robinson, 2017). It is evidently proved that ITCs are absorbed in higher amounts by passive diffusion from the gastrointestinal tract after ingestion into blood capillaries where they bind with free plasma proteins (thiocarbomylation) and pass into tissues cells where they affect their biophysiological mechanism (Kołodziejski et al., 2019). In a recent investigation, it was observed that broccoli converts gut microbiota to healthier profile, which coincides with myrosinase activity (Sikorska-Zimny and Beneduce, 2021). Most studies conducted among humans revealed that mercapturic acid pathway is involved in ITCs metabolism. One study using human urine explained that the ITCs can be absorbed indirectly through cylcocondensation determined by measuring plasma ITCs level after oral dose through high performance liquid chromatography with tandem mass spectrometry (HPLC-MS/MS) (Zhang and Zhang, 2017). Another study conducted on a rat model using radiolabel ITCs (14C) as an oral dose revealed the rapid absorption of ITCs, but the structure of individuals affects liposolubility (Chang et al., 2012). Both investigations observed that ITCs entered into enterocytes and glutathione S-transferase (GST)

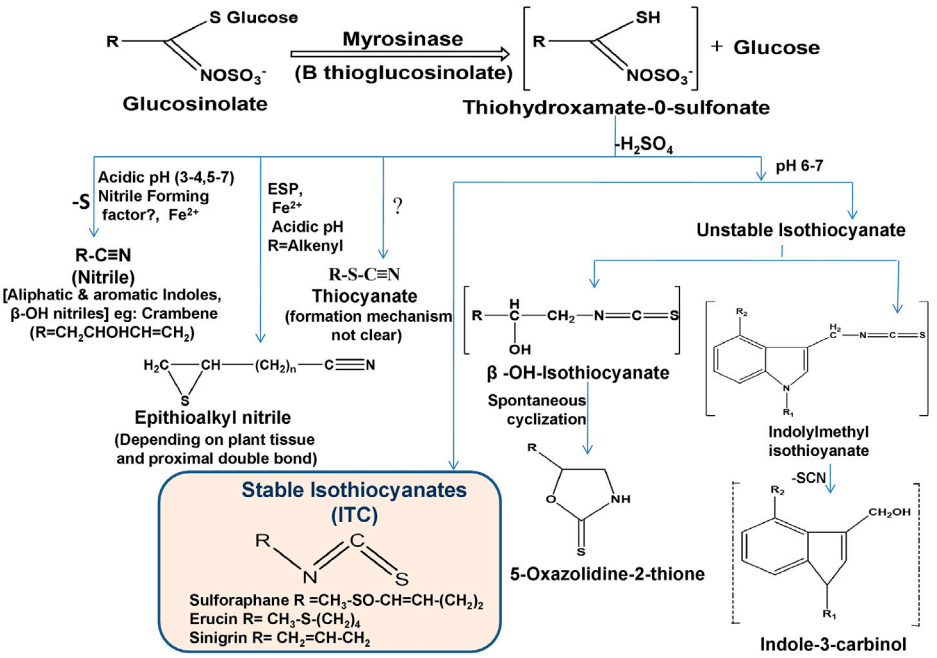


FIGURE 2
Enzymatic hydrolysis reaction of GLCs and their corresponding breakdown products (ESP; epithiospecifier protein).

conjugated with glutathione favors internal accumulation and concentration gradient. Kidney and liver are involved in entire conversion because the liver contains high levels of glutathione and highest GST activity and plays a crucial role in xenobiotic detoxification by supporting accumulation of conjugated ITCs (Esteve, 2020). These conjugated ITCs are converted to mercapturic acid derivatives, which are implicated by the kidney due to the presence of γ-glutamyltranspeptidase (γ-GT), N-acetyltransferase (AT), and cysteinylglycinase (CGase), after they are excreted in urine (Shakour et al., 2022).

3 Role in neuroprotection, AChE inhibition, and neuroinflammatory mechanism

Neurons are the building blocks of the CNS, incapable of reproducing and replacing themselves. Several pathological disorders including AD are caused by the accumulation of reactive oxygen species (ROS) in cells (Deshmukh et al., 2017). The ability of a compound to possess anti-inflammatory, antioxidative, and/or antiapoptotic properties is currently used to establish neuroprotective and neuroinflammatory functions (Dinkova-Kostova and Kostov, 2012). ITCs were reported to play a protective effect in acute and chronic AD (Kamal et al., 2022). A variety of ITCs have been experimentally proven (Table 2) to reduce oxidative stress, inflammation, excitotoxicity, misfolded proteins, and mitochondrial dysfunction, and prevent programmed cell death (Connolly et al., 2021). Through the activation of ARE (antioxidant response element) driven genes, ITCs are strong Nrf-2 (nuclear factor erythroid factor 2) activators. They strongly suppress inflammation via NF-kβ (nuclear factor kappa light chain enhancer of activated β cells) pathway (Sita et al., 2016).

A deficient and non-equilibrium cholinergic neurotransmission is responsible for the pathophysiology of learning and memory resulting behavioral disturbance, progressive loss of cognition and daily routine function (Hoyer, 2004; Craig et al., 2011). In context with the cholinergic hypothesis, decreasing the amount of acetylcholine in the hippocampus and cerebral cortex leads to the dysregulation of ChAT and premature loss of basal forbidden cholinergic neurons (Burčul et al., 2018; Hampel et al., 2019). One of the most significant properties of ITCs is AChE inhibition implicated in acetylcholine neurotransmission (Figure 3). In one study, 11 different ITCs were evaluated for their AChE inhibitory and anti - inflammation properties; the most promising inhibitory activity among 11 ITCs was demonstrated by phenyl isothiocyanate and its derivatives. The most potent AChE inhibitory activity was shown by 2-methoxyphenyl ITC with IC_{50} value of 0.57 mM. Human COX-2 enzyme was also used to evaluate the anti-inflammatory activity, ranking phenyl ITC and 2-methoxy phenylITC as the most potent with 99% inhibition at 50 μM (Burčul et al., 2018). Moringine-specific benzyl ITC from *Moringa Oleifera* modulated the Nrf2/AER pathway, proinflammatory biomarkers, and apoptotic pathway in different mouse and rat models (Galuppo et al., 2014, (Galuppo et al., 2015). In another mouse model (LPS induced), it was found that ITCs effectively decreased TNF-α, IL-1β, IL-6 and inhibited NF-kβ (Sailaja et al., 2022). It also downregulated senescence as it promoted neuronal repair in *in vitro* Aβ induce SH5Y5Y cells (Silvestro et al., 2021).

TABLE 2 The beneficial effects and mechanism of action of ITCs on various models of Alzheimer's disease.

Compound or extract	Experimental model	Pharmacological effects	Mechanism of action	References
6-(Methylsulfinyl) hexyl ITC (6-MSITC)	*in vitro,* cell line	Slow down the progression of inflammation	Slow down pro inflammatory cytokines expression and increased Nrf2	Chen et al. (2010)
	in vitro, LPS activated murine macrophage RAW 264 cell line	Reduced neuroinflammation	Strongly suppressed COX-2, iNOS and cytokines and attenuated the expression of these factors	Uto et al. (2005)
	in vivo, murin AD model	Decreased apoptosis and neuroinflammation	Inhibited phosphorylation of ERK, GSK3, decreased inflammatory cytokines and activate of caspase	Morroni et al. (2018)
	in vitro, IMR-32 neuronal cell lines	Exerted neuroprotective effect by reducing oxidative stress	Targeted Nrf-2 mediated oxidative stress through changes in gene expression (DNA microarray)	Trio et al. (2016)
Phenethyl ITC(PEITC)	*in vitro,* cell line	Decreased inflammation	Initiated Nrf2, modulate Nrf2/AER signalling pathway	Qin et al. (2015)
	in vivo, transgenic mice model	Reduced inflammation, activated cytoprotective pathway	Restored Nrf2 expression	Boyanapalli et al. (2014), Dayalan Naidu et al. (2018)
	in vitro LPS-activated rat astrocytes	Anti-inflammatory	Downregulated MAPK/ERK signalling	Dayalan Naidu et al. (2018); Latronico et al. (2021)
Moringin	*in vivo,* rat model	Enhanced cognition	Modulated Nrf2/AER pathway and pro inflammatory biomarkers	Galuppo et al. (2015)
	in vivo, mouse model	Abolished inflammation	Modulated apoptotic pathway and downregulate pro inflammatory cytokines	Galuppo et al. (2014)
	in vitro, Aβinduced- SHSY5Y cells	Promoted neuronal repair and slowdown Alzheimer's disease progression	Downregulated senescence, autophagy and mitophagy pathway	Silvestro et al. (2021)
	in vivo, lipopolysaccharide induced C57BL/6 mice model	Immunomodulatory and anti-inflammatory	Decreased pro inflammatory biomarkers (TNF-α, IL-1β, IL-6) in C2C12 myoblast, inhibited NF-kβ	Sailaja et al. (2022)
Erucin	*in vitro,* cell line	Stopped inflammation	Counteracted pro inflammatory markers expression, inhibited NF-kβ signalling pathway	Yehuda et al. (2012); Qin et al. (2015)
	in vitro, cell lines and *in vivo,* animal model	Decreased inflammation	Balanced Erk1/2, P38 and JNK signalling by Nrf2 pathway	Wagner et al. (2015)
	in vitro, LPS induced microglial cell line	Decreased inflammation	Decreased NO production, increased H_2S levels	Sestito et al. (2019)
Moringa oleifera extract	*in vivo*, colchicine and ethyl Choline induced rat model	Reduced neuronal cell death, ameliorated memory impairment and improved spatial memory	Upregulated phase II antioxidant enzymes, SOD and catalase	Ganguly and Guha. (2008); Sutalangka et al. (2013)
	in vivo, cadmium and alcoholic beverage induced Wistar rats	Neuroprotection	Reduced the activated astrocytes in frontal cortex	Omotoso et al. (2019)
	in vitro primary hippocampal neurons culture	Promoted neurite outgrowth and promoted neuronal survival	Increased NSE, decreased GFAP	Hannan et al. (2014)
	in vivo, NDD/Al induced temporo-cortical degenerated mice model	Reduced neurodegeneration	AChE inhibitory activity	Ekong et al. (2017)
	in vivo, NDD/hippocampal neuro-degenerated rat model	Enhanced memory and cognition	Maintained neuron integrity and cholinergic transmission	Adebayo et al. (2021)

(Continued on following page)

TABLE 2 (*Continued*) The beneficial effects and mechanism of action of ITCs on various models of Alzheimer's disease.

Compound or extract	Experimental model	Pharmacological effects	Mechanism of action	References
	in vivo, scopolamine induced mice model with spatial memory deficit	Improved spatial memory function	Altered the endogenous antioxidants, pro inflammatory mediators, elevatedAChE activity and promoted chromatolysis of cortical hippocample neurons	Onasanwo et al. (2021)
	in vivo lead acetate induced Wistar rat model	Ameliorated oxidative stress, inflammation and apoptosis	Protected neuronal cells via attenuation of NF-kβ signalling	Alqahtani and Albasher (2021)
	in vivo, CCl4 induced mice model	Modulated neuroinflammation and oxidative stress	Modulated TLR4/2MyD88/NF-kβ signalling	Mahmoud et al. (2022)
Sulforaphane	*in vitro*, human neuroblastoma cell line (SH-SY5Y)	Inhibited apoptosis	Modulated Bax/Bcl2	Lee et al. (2013)
	in vitro, murine neuroblastoma cell line (Neuro 2A and N1E-115)	Increased proteasome activity	Enhanced Nrf2 pathway	Park et al. (2009)
	in vivo, $AlCl_3$ and D-galactose induced mice model	Ameliorated cognitive impairment	Modulated Nrf2/ARE pathway	Zhang et al. (2014)
	in vivo mice model	Reduced inflammatory markers in glial and hippocampal cells, protected neurons	ITH12674 (melatonin sulforaphan hybrid) induced Nrf2 and scavenged free radicals	Michalska et al. (2020)
	in vivo, scopolamine induced mice model (C57BL/6) and *in vitro* scopolamine treated primary cortical neurons	Improved memory, cognition and cholinergic neurotransmission	Inhibited Acetyl cholinesterase (AChE)	Lee et al. (2014)
	in vitro, Swedish mutant mouse model (N2a/APPswe cells)	Inhibited Aβ generated neuroinflammation and oxidation	Epigenetic modification of Nrf2	Zhao et al. (2018)
	in vitro, human THP-1 macrophages (induced by $A\beta_{1-42}$)	Suppressed neuroinflammation	Downregulated NF-kβ pathway and preserved MERTK	Jhang et al. (2018)
	in vitro, amyloid induced microglial cells	Induced neuroinflammation	Increased microglial phagocytic activity	Chilakala et al. (2020)
	in vitro, dopaminergic SH-SY5Y human cells and LPS stimulated microglial BV2 cells	Prevented mitochondrial impairment and suppress neuroinflammation	InhibitedHO-1 enzyme	Brasil et al. (2023)
	in vivo, LPS induced rat model	Reduced inflammation	Suppressed LPS induced NF-kβ pathway, modulated TRAF6 and RIPI ubiquitination by cezanne	Wang et al. (2020)
Allyl isothiocyanate (AITC)	*in vitro*, neuroinflammatory model (NDD/LPS induced N2a neuroblastoma, BV2 murine microglia and C6 glioma cells)	Improved outgrowth of neurite and dysregulated apoptotic pathway	Suppressed NF-kβ/TNF-α/JNK signalling	Subedi et al. (2017)
	in vitro, cultured Schwann cells	Reduced neurogenic inflammation	Activated ROS dependent TRPA1	De Logu et al. (2022a)
	in vitro, murine RAW264.7 macrophages cell line, *in vivo* C57BL/6 mice	Suppressed inflammation	Decreased NF-kβ, downregulated pro inflammatory cytokine (IL-1β) and nitric oxide synthase, increased Nrf-2 and heme-oxygenase-1	Wagner et al. (2012)
	in vivo, cryogenic injury mice model	Increased plasticity markers level, regulate antioxidant genes	Decreased NF-kβ, GFAP, IL1β, IL-6, BBB permeability, increasing GAP43 and neural cell adhesion molecule	Caglayan et al. (2019)
Indole-3-carbinol (I3C)	*in vitro*, NDD/LPS induced BV-2 microglia (hyper activated)	Anti-apoptotic and anti-n -euroinflammatory activity, reduced microglial activation in hippocampus	Inhibited NF-kβ	Lee et al. (2014)

(Continued on following page)

TABLE 2 (*Continued*) The beneficial effects and mechanism of action of ITCs on various models of Alzheimer's disease.

Compound or extract	Experimental model	Pharmacological effects	Mechanism of action	References
	in vitro, PC12 neuronal cells (NDD/glutamate excitotoxicity)	Inhibited apoptotic pathway	Inhibited caspase 8 and 3, scavenged ROS	Jeong et al. (2015)
	in vivo, mice model	Suppressed neuroinflammation and oxido-nitrosoactive stress in brain	Decreased BDNF, GSH, increased levels of nitrites, malondialdihyde IL-1β, TNF-α	Huang et al. (2022)

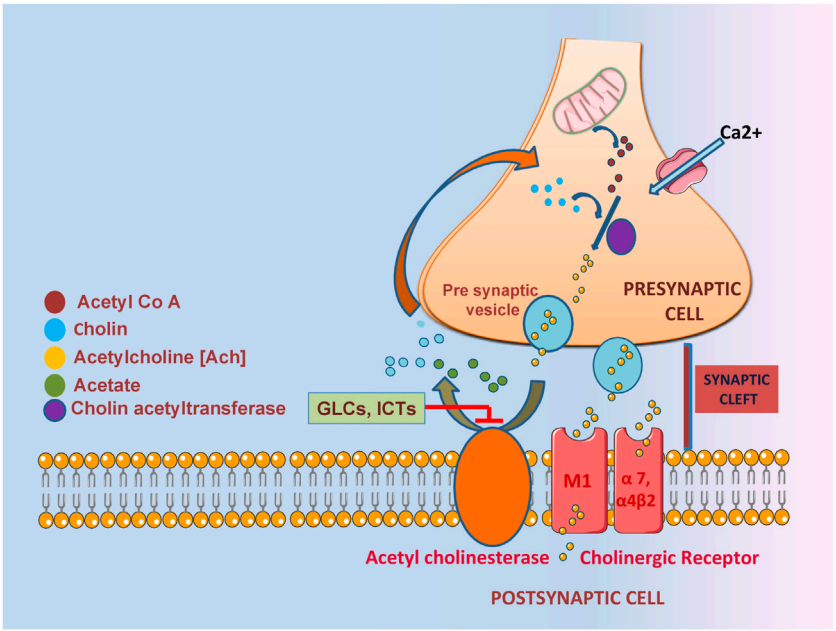


FIGURE 3
The role of GLCs derived ITCs in AChE inhibition characterized by impaired acetylcholine neurotransmission.

Through different mechanisms (explained in Table 2), SFN prevented cognitive impairment, reduced the Aβ and tau biomarkers, oxidative stress, inflammation and neurodegeneration in experimental models (Kim, 2021). SFN was able to improve spatial and contextual memory through the Y-maze test and counteract the Aβ aggregate induced memory deficits induced by intracerebroventricular (ICV) injection in a mouse model (Kim, 2021). In the hippocampus and frontal cortex, SFN increased cholinacetyltransferase (ChAT) expression, decreased acetylcholine esterase (AChE) activity, and raised the level of acetylcholine (AChE) (Lee et al., 2014). In another study on a transgenic AD mouse model, it was observed that SFN not only reduced the production and deposition of Aβ plaques in the hippocampus and cerebral cortex but also it is associated with neurobehavioral deficit (Zhang et al., 2015; 2017). The neuroinflammatory inhibition is through the activation of Nrf2/HO-1 pathway and inhibition of JNK/AP-1/NF-Kβ by SFN. SFN significantly increased proteasome activity and enhance Nrf-2 pathway in murine neuroblastoma cell lines (Park et al., 2009). It also modulated the Nrf2/ARE pathway in an $AlCl_3$-and D-galactose induced mice (Zhang et al., 2014).

Neurogenesis has been shown to be enhanced by AITC and PEITC. AChE inhibitory activity in AD revealed that PEITC inhibited the enzyme more effectively than benzyl ITC and AITC (Burčul et al., 2018). In another study, PEITC inhibited Akt activation, suppressed NO production through INF induction, and had an anti-inflammatory effect (Okubo et al., 2010). PEITC showed a protective effect by modulating the MAPK pathway (Ma et al., 2017). Experimental findings revealed that in LPS-induced inflammation model, AITC showed a neuroprotective effect mediated through downregulation of JNK/NF-kβ/TNF-α signaling (Subedi et al., 2017). It also activated ROS-dependent TRPA1 signaling, resulting in neurogenic inflammation reduction in cultured Schwann cells *in vitro* (De Logu et al., 2022a; De Logu et al., 2022b). PEITC decreased inflammation and activated the cytoprotective pathway in transgenic mice model by modulating Nrf2/AER pathway and restoring Nrf-2 expression (Boyanapalli et al., 2014; Dayalan Naidu et al., 2018). In another study using LPS-activated rat astrocyte culture, PEITC significantly

downregulated MAPK/ERK signaling and influenced the inflammatory pathway (Latronico et al., 2021). Increasing evidences suggests that cytochrome p450 is fundamental for brain homeostasis and function while phase II enzyme such as glutathione *S*-transferase play a key role in redox homeostasis. Modulation of these enzymes can be achieved by ITCs, in the recent studies glucuronosyltranseferase expression increase by sulforaphane in HepG2 cells, in another study erucin and phenethyl ITC elevated glucuronosyltranseferase activity in rat liver slices (Abdull Razis and Mohd Noor, 2013).

Moringa oleifera extract (MOE) decreased the neuritis resulting from naturally occurring cellular injury, with the development of multipolar primary process (Hannan et al., 2014). It also suppressed oxidative stress, MDA, nitrite and TNF-α, increased SOD and inflammation and improved spatial memory and cholinergic neurotransmission by reducing AChE activity and loss of cortico-hippocampus neurons in rat model fed with *M. oleifera* seeds in dose dependent manner (Onasanwo et al., 2021). *Moringa oleifera* extract also scavenged free radicals produced by NO, iNOS and nitrotyrosine increase Nrf2 in LPS-activated macrophages and downregulated antioxidative genes; HO-1, GST-P1 and NQO- (Jaja-Chimedza et al., 2017). In another study, it significantly inhibited AChE and reduced neurodegeneration in an NDD/Al - induced temporocortical degenerated mice model (Ekong et al., 2017). *Moringa oleifera* - supplemented male Wistar rats showed improved memory when evaluated by the Morris water Maze test and significantly reduced AChE levels in brain tissues in a dose-dependent manner (Adebayo et al., 2021). In another observation, GMC-ITC treated neuronal cells (SH-SY5Y) significantly alleviate oxidative stress condition by reducing ROS level ((Jaafaru et al., 2019a). Glucomoringin ITC (GMC-ITC) isolated from *M. oleifera* seeds abrogated oxidative stress and showed neuroprotective activity against cytotoxic neuroblastoma cells (SH-SY5Y) induced by H_2O_2, gene expression study of detoxifying markers (phase II) by GMC-ITC revealed that all involved genes significantly express themselves. It also decreased the expression of NF-kβ and increased the expression of Ikβ, Nrf2, SOD-1, NQO1 and Nf-kβ respectively (Jaafaru et al., 2019b). *Eruca sativa* extract (ESE) with a high amount of erucin (ER) prevented cell death and degeneration induced by LPS in NSC-34 motor neurons exposed to LPS-stimulated macrophage cell culture medium by inhibiting FasL (tumor necrosis factor ligand superfamily number 6 expression) and suppressing pro-inflammatory mediators (attenuates TLR4, COX-2 expression of TNF-α level) (Gugliandolo et al., 2018). Erucin decreased inflammation in different cell line models, counteracted proinflammatory marker expression, and balanced Erk1/2, P38, and JNK signaling (Yehuda et al., 2012; Wagner et al., 2015). Indol 3 carbinol (I3C) is another promising candidate found in vegetables; it reduces the free radical production in neuronal cells (Mammana et al., 2019). It also showed the potent radical scavenging activity by chelating already produced free radical species (Giacoppo et al., 2015). In another study, it suppressed the expression of NO, COX-2, and iNOS in the brain, which prevented apoptosis and inflammation by inhibiting NF-kβ and IB phosphorylation (Kim et al., 2014). Furthermore, it decreased BDNF, GHS and increased TNF-α, IL1-β in mice model, it also helped in suppression of neurodegeneration (Huang et al., 2022). In another experiment, researchers explored the antioxidant and anti-inflammatory activity of SFN and ERN as H_2S donor through the combination with rivastigmine in microglia and neuronal cell line (SH-SY5Y). Result revealed that both derivatives show significant antioxidant and anti inflammatory activities in microglial cell line, expression of antioxidant defense protein (GSH) was also induced in neuronal cell line. It significantly decreased the ROS production and NO release in microglial BV-2 cells. Further Erucin exerted a time dependent Nrf2 activation in SH-SY5Y cells (Sestito et al., 2019). When anti-inflammatory effect of erucin was evaluated in LPS-challenged umbilical vein endothelial cells (HUVECs), it significantly prevented the increase of ROS, TNF-α levels and decreased COX-2. It also induced NF-kβ (Ciccone et al., 2022).

4 Potential role of GLCs and ITCs against pathological hallmarks and their neurotoxicity

The brain of people suffering from Alzheimer's disease shows remarkable accumulations of two neurotoxic proteins Aβ and tau (Cao et al., 2020). So far, several Alzheimer's plaque and tau inhibitors from different sources are available they can target different mechanistic steps of fibril formation. One of the inhibitors that are widely used to stop protein aggregation is GLCs derivatives ITCs as they are consumed as a part of our daily diet (Lopez-Rodriguez et al., 2020). In Table 3, we have discussed some of the GLCs derived ITCs, proposed as the potential inhibitor of misfolded Aβ and tau aggregation and their induced toxicity by different mechanisms and modulation of multiple pathways (Figures 4, 5) as described earlier (Grande et al., 2020). Recent investigations suggested that they may directly interact with misfolded proteins during very early stages of the aggregation cascade by binding and stabilizing unfolded proteins and redirecting the aggregation pathways to form amorphous nontoxic fibrils, blocking seeding and further conformational changes that result in neurotoxicity and cell death.

6-(Methylsulfinyl) hexyl isothiocyanate (6-MSITC) from *Wasabia japonica* was evaluated against amyloidosis in a murine mice model in which 6-MSITC was induced by intra cerebroventricular injection of $Aβ_{1-42}$ oligomers. Behavioral analysis revealed that it reduced Aβ1-42 induced memory impairment in hippocampus tissues, increased ROS, and decreased glutathione levels following $Aβ_{1-42}$ injection (Morroni et al., 2018). In another study, the authors observed that $Aβ_{25-35}$ induced mitochondrial dependent cell death was blocked by SFN through Nrf2-associated manner (Brasil et al., 2023). Clinically, it inhibited Aβ, reduced its burden, and increased the expression of p75NTR in an intransgenic mouse model (Zhang et al., 2015). In another investigation, SFN was found to suppress Aβ deposition, improve cognition, and locomotor function in aluminum and D-galactose-induced mouse model (Zhang et al., 2017). It modulated the Aβ expression related markers followed CDK5 overexpression inhibition in primary neurons, further it reduced $Aβ_{1-42}$ induced neurotoxicity and its deposition in TgCRND8-transgenic mice brains. It also suppressed tau phosphorylation at specific sites (Yang et al., 2023). It reduced and altered hyperphosphorylated tau proteins in embryonic hippocampal

TABLE 3 Beneficial effects of ITCs against pathological hallmarks and their neurotoxicity.

Plant/Compound	Mechanism of action	Pharmacological effectiveness	Test scale	References
Against amyloid beta oligomerization and toxicity				
Wasabia japonica (6-methylsulfinyl hexyl ITC)	Increased glutathione levels and ROS in hippocampus by $A\beta_{1-42}$ injection were reduced	Neuroprotection against $A\beta_{1-42}$ and ameliorates $A\beta_{1-42}$ induced memory impairments	*in-vivo,* murine model, induced by intra cerebrovascular injection of $A\beta_{1-42}$	Morroni et al. (2018)
Indole-3 carbinol (I3C)	High affinity molecular recognition and reduced Aβ fragments by heteromeric interaction	Reduced amyloid production	*in-vitro,* biochemical method	Cohen et al. (2006)
Moringa oleifera extract	Downregulated BACE1	Decreased Aβ production, rescued cognitive impairment and enhanced the reduced synaptic proteins synapsin, synapsophysin, PSD93 and PSD95	*in-vivo,* hyperhomocysteinemia (HHcY) induced AD model	Mahaman et al. (2018)
	Deactivated calpain by ↓ intracellular Ca 2+, reduced ca2+ signaling and prevent cell death	Decreased cytosolic cysteine protease caplain activity	*in-vivo,* hyperhomocysteinemia (HHcy) induced rat model (AD like pathology)	Mahaman et al. (2018)
	Increased Aβ immunoexpression was significantly abolished, sustained the brain-Zn content	Decreased the aggregation and accumulation of Aβ	*in-vivo,* ACR induced forty male Sprague Dawley rat treated with MO-ZnONP	Dahran et al. (2023)
Sulforaphane	Increased levels of HSP-70 co-chaperons and CHIP (Aβ metabolism influencers)	Reduced monomeric and polymeric forms of Aβ, but do not affect m-RNA expression, ameliorated memory deficits	*in-vivo*, triple transgenic mouse model (3×Tg-AD)	Li et al. (2018)
	Decreased oxidative stress and neuroinflammation (generator of Aβ)	Significantly inhibited Aβ aggregation, ameliorated neurobehavioral deficits peroxidation in brain	*in-vivo*, 6-month-old PS1V97L transgenic (Tg) mice	Zhang et al. (2015)
	Modulated the amyloid expression related markers, inhibited the overexpression of CDK5 in primary neurons	Reduced the $A\beta_{1-42}$ deposition and related neurotoxicity	*in-vivo,* TgCRND8 transgenic mice model	Yang et al. (2023)
	Inhibited cathepsin-B and caspase-1 dependent NLRP3 inflammasome activation induced by Aβ monomers (1–42)	Reduced Aβ induced neurotoxicity	*in-vitro,* human THP-1 macrophages like cells	An et al. (2016)
	Alleviated several downstream pathological changes including oxidative stress and neuroinflammation	Significantly inhibited the generation of Aβ aggregates promotes spatial learning and memory	*in-vivo,* PS1V97L transgenic mice model	Hou et al. (2018)
Against tau hyperphosphorylation and toxicity				
Moringa oleifera extract	Not known	Decreased hyperphosphorylated tau at different sites (S-199, S-404, S-396 and T-231)	*in-vivo,* hyperhomocysteinemia (HHcy) induced rat model (AD like pathology)	Mahaman et al. (2018)
	Reduced sensory dysfunction and motor deficits, abolished immunoexpression of phosphorylated tau proteins	Reduced ACR induced neurotoxicity and tau proteins	*in-vivo,* ACR induced forty male Sprague Dawley rat treated with MO-ZnONP	Dahran et al. (2023)
Sulforaphane	Increased levels of HSP-70 co-chaperons and CHIP (Aβ metabolism influencers)	Reduced protein levels of tau and hyperphosphorylated tau, ameliorated memory deficits	*in-vivo,* triple transgenic mouse model (3×Tg-AD)	Lee et al. (2014)
	Suppressed phosphorylation of tau at specific sites, markedly suppressed the CDK5/p25	Reduced tau protein hyperphosphorylation in the brain and improved synaptic plasticity	*in-vivo,* TgCRND8 transgenic mice model	Yang et al. (2023)
	Altered phosphorylated tau at threonine 181 and serine991/ 202 distribution within astrocytes	Reduced hyperphosphorylated tau proteins in astrocytes under hypoglycaemic condition	*in vitro,* embryonic hippocampal rat astrocytes	Komiskey et al. (2022)

(Continued on following page)

TABLE 3 (*Continued*) Beneficial effects of ITCs against pathological hallmarks and their neurotoxicity.

Plant/Compound	Mechanism of action	Pharmacological effectiveness	Test scale	References
	Significantly inhibited hyperphosphorylated tau proteins at Ser396, Ser404 and Thr 205 site, enhanced the ration of p-GSK-3β(Ser9)/GSK-3β and p-Akt (Ser473)/Akt in hippocampus	Reduced the accumulation of phosphorylated tau in hippocampus and related toxicity	*in-vivo*,streptozotocin induced rat model	Yang et al. (2020)
	Significantly expressed the NDP52 induced by Nrf2 and facilitated clearance of p-tau proteins	Reduced the phosphorylated tau proteins	*in-vivo*, C57BL/6J mice model	Jo et al. (2014)

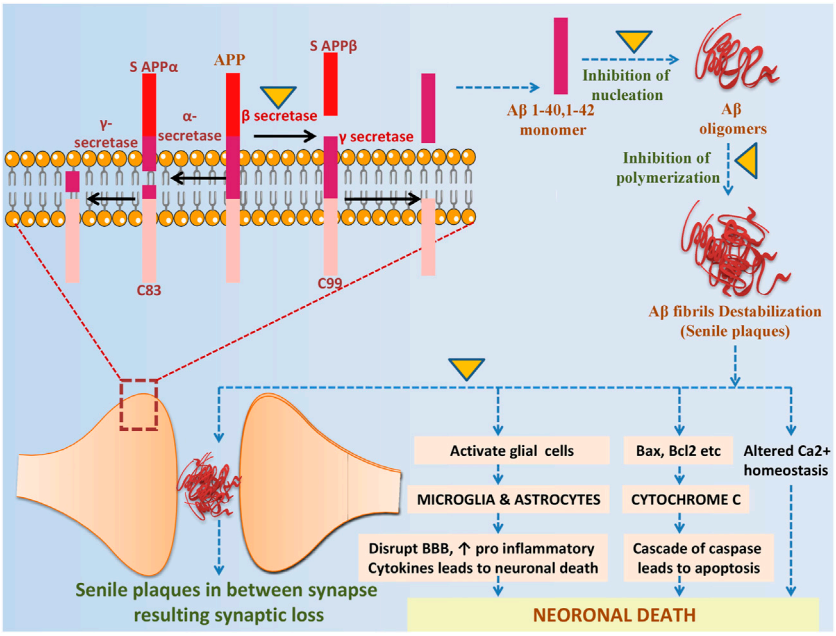


FIGURE 4
The potential role of ITCs in Aβ metabolism and related toxicity: sAPPα and C83 (membrane associated fragment) are formed by nonamyloidogenic pathway in which APP is cleaved by α-secretase, while in amyloidogenic pathway APP is cleaved by β-secretase producing S APPβ and C99 fragment, γ-secretase then processed the C99 and release Aβ. ITCs prevent from amyloidogenic cleavage by inhibiting β-secretase, further it inhibits nucleation, polymerization and plaques formation. It directly intervenes in Aβ induced neurotoxicity by altering Ca2+ homeostasis, downregulating cascade of caspase and in reducing inflammation.

rat astrocytes under hypoglycaemic condition at Th 181 and Sr 991/202 within astrocytes (Komiskey et al., 2022). It induced NDP52 by Nrf2 and cleared the phosphorylated tauproteins in mice model (Jo et al., 2014).Through high affinity molecular recognition by heteromeric interaction of Aβ plaques, I3C were found to strongly reduce Aβ fibril formation as observed in microscopic examination by TEM analysis (Cohen et al., 2006).

M.oleifera is profoundly used against chronic diseases including AD. Mitochondrial apoptotic genes profile through GMC-ITC pre-treated SH-SY5Y neuronal cells revealed that it protect the cells against oxidative stress via apoptotic pathway, it significantly downregulate the expression of Bax, CASP3, CASP8, CASP9, Apaf-1, cyt-c, p-53 genes and upregulate Bcl2 gene in mitochondrial apoptotic signalling pathway (Jaafaru et al., 2019a). In another study GMC-ITC from the seeds of *M. oleifera* significantly decreased the expression of BACE1, APP and increased the expression of MAPT tau genes in H_2O_2 induced cytotoxic neuroblastoma cell (SH-SY5Y) (Jaafaru et al., 2019b). It decreases Aβ production and enhance the synaptic proteins in HHcY induced AD model bydown regulating BACE1. It also played crucial role in Ca^{2+} homeostasis, as it deactivated calpain by decreasing intracellular Ca^{2+} resulting cytosolic protease calpain activity reduction in HHcY induced rat model (Mahaman et al., 2018). In another study conducted on MO-ZnONP treated Sprague Dawley rat model it reduced the Aβ accumulation and helped in sustained brain-Zn content (Dahran et al., 2023).

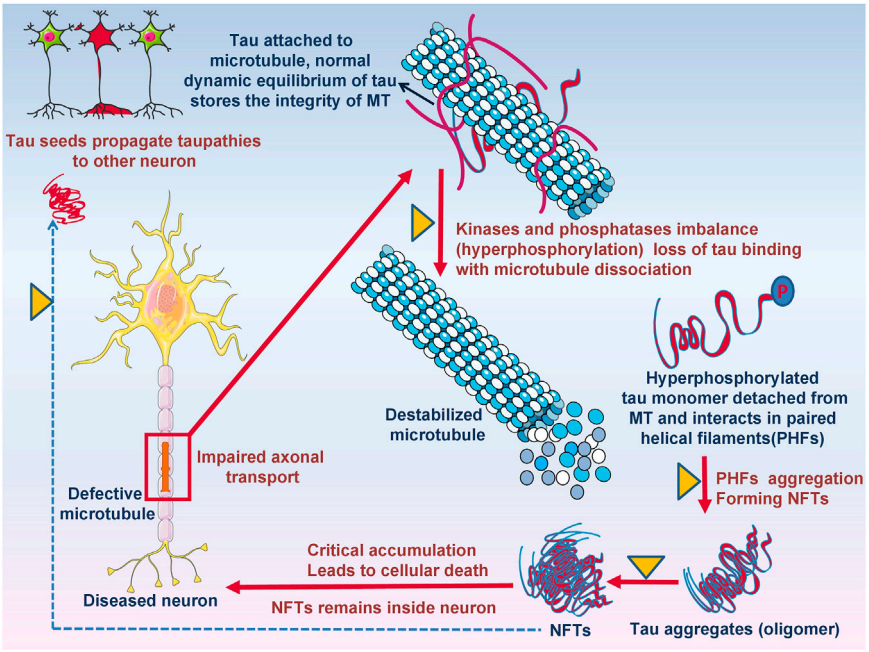


FIGURE 5
The potential role of ITCs in disease modification, targeting tau protein and its aggregation. Defective microtubules resulting in impaired axonal transport due to kinases and phosphatase imbalance resulting destabilized microtubule formation. Detached hyperphosphorylated tau monomers oligomerized and form NFTs leads to cellular death.

5 Conclusion

GLCs derived ICTs are important bioactive natural products that are found in many Brassicaceae plants and few plants from other families. *In vitro* and animal studies have reported their beneficial effects in neuroprotection and they are reported to enhance cellular metabolism, nourish brain cells, and reduce risk factors associated with neurodegeneration. ITCs inhibit inflammatory mediators, oxidative stress, cellular stress signaling, and improve behavioural measures. They also easily cross the blood brain barrier to interact with particular targets implicated in AD pathogenesis. However, there is no sufficient clinical evidence to prove these effects in humans. Future studies should focus to evaluate their pharmacokinetic parameters and effectiveness in humans.

Author contributions

JA contributed to the study conception and design. The first draft of the manuscript was written by FK. All tables and figures have been prepared by FK. AJ has performed literature survey. JA, HPD, AJ, VS, and VK revised and updated the manuscript. All authors read and approved the final 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

References

Abdull Razis, A. F., De Nicola, G. R., Pagnotta, E., Iori, R., and Ioannides, C. (2013). A glucosinolate-rich extract of Japanese Daikon perturbs carcinogen-metabolizing enzyme systems in rat, being a potent inducer of hepatic glutathione S-transferase. *Eur. J. Nutr.* 52, 1279–1285. doi:10.1007/s00394-012-0397-2

Abdull Razis, A. F., and Mohd Noor, N. (2013). Cruciferous vegetables: Dietary phytochemicals for cancer prevention. *Asian Pac. J. Cancer Prev.* 14, 1565–1570. doi:10.7314/APJCP.2013.14.3.1565

Adebayo, O. G., Wopara, I., Aduema, W., Ebo, O. T., and Umoren, E. B. (2021). Long-term consumption of Moringa oleifera-supplemented diet enhanced neurocognition, suppressed oxidative stress, acetylcholinesterase activity and neuronal degeneration in rat's hippocampus. *Drug Metab. Pers. Ther.* 36, 223–231. doi:10.1515/dmpt-2020-0189

Agerbirk, N., and Olsen, C. E. (2012). Glucosinolate structures in evolution. *Phytochemistry* 77, 16–45. doi:10.1016/j.phytochem.2012.02.005

Alexandre, E. M. C., Moreira, S. A., Pinto, C. A., Pintado, M., and Saraiva, J. A. (2020). "Analysis of glucosinolates content in food products," in *Glucosinolates: Properties, recovery, and applications* (Elsevier), 213–250. doi:10.1016/B978-0-12-816493-8.00007-X

Ali, S. S., Ahmad, N., Jamal Gilani, S., and Ali Khan, N. (2018). Isothiocyanates: A review. *Res. J. Pharmacogn.* 5, 71–89. doi:10.22127/RJP.2018.58511

Alqahtani, W. S., and Albasher, G. (2021). Moringa oleifera Lam. extract rescues lead-induced oxidative stress, inflammation, and apoptosis in the rat cerebral cortex. *J. Food Biochem.* 45, e13579. doi:10.1111/jfbc.13579

Amarakoon, D., Lee, W., Tamia, G., and Lee, S. (2023). Indole-3-Carbinol: Occurrence, health-beneficial properties, and cellular/molecular mechanisms. *Annu. Rev. Food Sci. Technol.* 14, 347–366. doi:10.1146/annurev-food-060721-025531

An, Y. W., Jhang, K. A., Woo, S.-Y., Kang, J. L., and Chong, Y. H. (2016). Sulforaphane exerts its anti-inflammatory effect against amyloid-β peptide via STAT-1 dephosphorylation and activation of Nrf2/HO-1 cascade in human THP-1 macrophages. *Neurobiol. Aging* 38, 1–10. doi:10.1016/j.neurobiolaging.2015.10.016

Andrade, S., Ramalho, M. J., Loureiro, J. A., and Pereira, M. do C. (2019). Natural compounds for Alzheimer's disease therapy: A systematic review of preclinical and clinical studies. *Int. J. Mol. Sci.* 20, 2313. doi:10.3390/ijms20092313

Asif, M., Kala, C., Gilani, S. J., Imam, S. S., Mohamad, T., Naaz, F., et al. (2022). Protective effects of isothiocyanates against Alzheimer's disease. *Curr. Tradit. Med.* 8, 1–10. doi:10.2174/2215083807666211109121345

Avato, P., and Argentieri, M. P. (2015). Brassicaceae: A rich source of health improving phytochemicals. *Phytochem. Rev.* 14, 1019–1033. doi:10.1007/s11101-015-9414-4

Barba, F. J., Nikmaram, N., Roohinejad, S., Khelfa, A., Zhu, Z., and Koubaa, M. (2016). Bioavailability of glucosinolates and their breakdown products: Impact of processing. *Front. Nutr.* 3, 24. doi:10.3389/fnut.2016.00024

Bennett, R. N., Mellon, F. A., and Kroon, P. A. (2004). Screening crucifer seeds as sources of specific intact glucosinolates using ion-pair high-performance liquid chromatography negative ion electrospray mass spectrometry. *J. Agric. Food Chem.* 52, 428–438. doi:10.1021/jf030530p

Blažević, I., and Mastelić, J. (2009). Glucosinolate degradation products and other bound and free volatiles in the leaves and roots of radish (Raphanus sativus L). *Food Chem.* 113, 96–102. doi:10.1016/j.foodchem.2008.07.029

Boyanapalli, S. S. S., Paredes-Gonzalez, X., Fuentes, F., Zhang, C., Guo, Y., Pung, D., et al. (2014). Nrf2 knockout attenuates the anti-inflammatory effects of phenethyl isothiocyanate and curcumin. *Chem. Res. Toxicol.* 27, 2036–2043. doi:10.1021/tx500234h

Brasil, F. B., de Almeida, F. J. S., Luckachaki, M. D., Dall'Oglio, E. L., and de Oliveira, M. R. (2023). The isothiocyanate sulforaphane prevents mitochondrial impairment and neuroinflammation in the human dopaminergic SH-SY5Y and in the mouse microglial BV2 cells: Role for heme oxygenase-1. *Metab. Brain Dis.* 38, 419–435. doi:10.1007/s11011-022-00990-x

Burčul, F., Generalić Mekinić, I., Radan, M., Rollin, P., and Blažević, I. (2018). Isothiocyanates: Cholinesterase inhibiting, antioxidant, and anti-inflammatory activity. *J. Enzyme Inhib. Med. Chem.* 33, 577–582. doi:10.1080/14756366.2018.1442832

Caglayan, B., Kilic, E., Dalay, A., Altunay, S., Tuzcu, M., Erten, F., et al. (2019). Allyl isothiocyanate attenuates oxidative stress and inflammation by modulating Nrf2/HO-1 and NF-κB pathways in traumatic brain injury in mice. *Mol. Biol. Rep.* 46, 241–250. doi:10.1007/s11033-018-4465-4

Cammisuli, D. M., Cipriani, G., and Castelnuovo, G. (2022). Technological solutions for diagnosis, management and treatment of Alzheimer's disease-related symptoms: A structured review of the recent scientific literature. *Int. J. Environ. Res. Public Health* 19, 3122. doi:10.3390/ijerph19053122

Cancer, I. A. for R. on, Strategies, I. W. G., theof, E. C.-P., and Organization, W. H. (2004). *Cruciferous vegetables, isothiocyanates and indoles*. IARC.

Cao, Q., Tan, C.-C., Xu, W., Hu, H., Cao, X.-P., Dong, Q., et al. (2020). The prevalence of dementia: A systematic review and meta-analysis. *J. Alzheimer's Dis.* 73, 1157–1166. doi:10.3233/JAD-191092

Cartea, M. E., and Velasco, P. (2008). Glucosinolates in Brassica foods: Bioavailability in food and significance for human health. *Phytochem. Rev.* 7, 213–229. doi:10.1007/s11101-007-9072-2

Chang, Q., Wang, G.-N., Li, Y., Zhang, L., You, C., and Zheng, Y. (2012). Oral absorption and excretion of icaritin, an aglycone and also active metabolite of prenylflavonoids from the Chinese medicine Herba Epimedii in rats. *Phytomedicine* 19, 1024–1028. doi:10.1016/j.phymed.2012.05.017

Chen, J., Uto, T., Tanigawa, S., Yamada-Kato, T., Fujii, M., and Hou, D.-X. (2010). Microarray-based determination of anti-inflammatory genes targeted by 6-(methylsulfinyl) hexyl isothiocyanate in macrophages. *Exp. Ther. Med.* 1, 33–40. doi:10.3892/etm_00000006

Chilakala, R. R., Manchikalapudi, A. L., Kumar, A., and Sunkaria, A. (2020). Sulforaphane attenuates Aβ oligomers mediated decrease in phagocytic activity of microglial cells. *Neuroscience* 429, 225–234. doi:10.1016/j.neuroscience.2020.01.002

Ciccone, V., Piragine, E., Gorica, E., Citi, V., Testai, L., Pagnotta, E., et al. (2022). Anti-inflammatory effect of the natural H2S-donor erucin in vascular endothelium. *Int. J. Mol. Sci.* 23, 15593. doi:10.3390/ijms232415593

Citi, V., Martelli, A., Testai, L., Marino, A., Breschi, M. C., and Calderone, V. (2014). Hydrogen sulfide releasing capacity of natural isothiocyanates: Is it a reliable explanation for the multiple biological effects of brassicaceae? *Planta Med.* 80, 610–613. doi:10.1055/s-0034-1368591

Cohen, T., Frydman-Marom, A., Rechter, M., and Gazit, E. (2006). Inhibition of amyloid fibril formation and cytotoxicity by hydroxyindole derivatives. *Biochemistry* 45, 4727–4735. doi:10.1021/bi051525c

Connolly, E. L., Sim, M., Travica, N., Marx, W., Beasy, G., Lynch, G. S., et al. (2021). Glucosinolates from cruciferous vegetables and their potential role in chronic disease: Investigating the preclinical and clinical evidence. *Front. Pharmacol.* 12, 767975. doi:10.3389/fphar.2021.767975

Craig, L. A., Hong, N. S., and McDonald, R. J. (2011). Revisiting the cholinergic hypothesis in the development of Alzheimer's disease. *Neurosci. Biobehav. Rev.* 35, 1397–1409. doi:10.1016/j.neubiorev.2011.03.001

Dahran, N., Abd-Elhakim, Y. M., Mohamed, A. A.-R., Abd-Elsalam, M. M., Said, E. N., Metwally, M. M. M., et al. (2023). Palliative effect of Moringa olifera-mediated zinc oxide nanoparticles against acrylamide-induced neurotoxicity in rats. *Food Chem. Toxicol.* 171, 113537. doi:10.1016/j.fct.2022.113537

Dayalan Naidu, S., Suzuki, T., Yamamoto, M., Fahey, J. W., and Dinkova-Kostova, A. T. (2018). Phenethyl isothiocyanate, a dual activator of transcription factors NRF2 and HSF1. *Mol. Nutr. Food Res.* 62, 1700908. doi:10.1002/mnfr.201700908

De Logu, F., De Siena, G., Landini, L., Marini, M., Souza Monteiro de Araujo, D., Albanese, V., et al. (2022a). Non-neuronal TRPA1 encodes mechanical allodynia associated with neurogenic inflammation and partial nerve injury in rats. *Br. J. Pharmacol.* 180, 1232–1246. doi:10.1111/bph.16005

De Logu, F., Nassini, R., Hegron, A., Landini, L., Jensen, D. D., Latorre, R., et al. (2022b). Schwann cell endosome CGRP signals elicit periorbital mechanical allodynia in mice. *Nat. Commun.* 13, 646–719. doi:10.1038/s41467-022-28204-z

Deshmukh, P., Unni, S., Krishnappa, G., and Padmanabhan, B. (2017). The keap1–nrf2 pathway: Promising therapeutic target to counteract ROS-mediated damage in cancers and neurodegenerative diseases. *Biophys. Rev.* 9, 41–56. doi:10.1007/s12551-016-0244-4

Devkota, H. P. (2020). "Analysis of glucosinolates," in *Recent advances in natural products analysis* (Elsevier), 651–661. doi:10.1016/B978-0-12-816455-6.00020-2

Dinkova-Kostova, A. T., and Kostov, R. V. (2012). Glucosinolates and isothiocyanates in health and disease. *Trends Mol. Med.* 18, 337–347. doi:10.1016/j.molmed.2012.04.003

Disbrow, E., Stokes, K. Y., Ledbetter, C., Patterson, J., Kelley, R., Pardue, S., et al. (2021). Plasma hydrogen sulfide: A biomarker of Alzheimer's disease and related dementias. *Alzheimer's Dement.* 17, 1391–1402. doi:10.1002/alz.12305

Dorszewska, J., Prendecki, M., Oczkowska, A., Dezor, M., and Kozubski, W. (2016). Molecular basis of familial and sporadic Alzheimer's disease. *Curr. Alzheimer Res.* 13, 952–963. doi:10.2174/1567205013666160314150501

Ekong, M. B., Ekpo, M. M., Akpanyung, E. O., and Nwaokonko, D. U. (2017). Neuroprotective effect of Moringa oleifera leaf extract on aluminium-induced temporal cortical degeneration. *Metab. Brain Dis.* 32, 1437–1447. doi:10.1007/s11011-017-0011-7

Esteve, M. (2020). Mechanisms underlying biological effects of cruciferous glucosinolate-derived isothiocyanates/indoles: A focus on metabolic syndrome. *Front. Nutr.* 7, 111. doi:10.3389/fnut.2020.00111

Ettcheto, M., Sánchez-López, E., Gómez-Mínguez, Y., Cabrera, H., Busquets, O., Beas-Zarate, C., et al. (2018). Peripheral and central effects of memantine in a mixed preclinical mice model of obesity and familial Alzheimer's disease. *Mol. Neurobiol.* 55, 7327–7339. doi:10.1007/s12035-018-0868-4

Fahey, J. W., Zalcmann, A. T., and Talalay, P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56, 5–51. doi:10.1016/S0031-9422(00)00316-2

Fão, L., Mota, S. I., and Rego, A. C. (2019). Shaping the Nrf2-ARE-related pathways in Alzheimer's and Parkinson's diseases. *Ageing Res. Rev.* 54, 100942. doi:10.1016/j.arr.2019.100942

Fernández, S. S. M., and Ribeiro, S. M. L. (2018). Nutrition and alzheimer disease. *Clin. Geriatr. Med.* 34, 677–697. doi:10.1016/j.cger.2018.06.012

Galuppo, M., Giacoppo, S., De Nicola, G. R., Iori, R., Navarra, M., Lombardo, G. E., et al. (2014). Antiinflammatory activity of glucomoringin isothiocyanate in a mouse model of experimental autoimmune encephalomyelitis. *Fitoterapia* 95, 160–174. doi:10.1016/j.fitote.2014.03.018

Galuppo, M., Giacoppo, S., Iori, R., De Nicola, G. R., Bramanti, P., and Mazzon, E. (2015). Administration of 4-(α-L-rhamnosyloxy)-benzyl isothiocyanate delays disease phenotype in SOD1G93A rats: A transgenic model of amyotrophic lateral sclerosis. *Biomed. Res. Int.* 2015, 259417–259512. doi:10.1155/2015/259417

Ganguly, R., and Guha, D. (2008). Alteration of brain monoamines and EEG wave pattern in rat model of Alzheimer's disease and protection by Moringa oleifera. *Indian J. Med. Res.* 128, 744–751.

Giacoppo, S., Galuppo, M., Montaut, S., Iori, R., Rollin, P., Bramanti, P., et al. (2015). An overview on neuroprotective effects of isothiocyanates for the treatment of neurodegenerative diseases. *Fitoterapia* 106, 12–21. doi:10.1016/j.fitote.2015.08.001

Grande, G., Qiu, C., and Fratiglioni, L. (2020). Prevention of dementia in an ageing world: Evidence and biological rationale. *Ageing Res. Rev.* 64, 101045. doi:10.1016/j.arr.2020.101045

Gugliandolo, A., Giacoppo, S., Ficicchia, M., Aliquò, A., Bramanti, P., and Mazzon, E. (2018). Eruca sativa seed extract: A novel natural product able to counteract neuroinflammation. *Mol. Med. Rep.* 17, 6235–6244. doi:10.3892/mmr.2018.8695

Gupta, P. K., and Robinson, J. R. (2017). *). Oral controlled-release delivery. Treatise Control.* drug Deliv., 255–313.

Hampel, H., Mesulam, M.-M., Cuello, A. C., Khachaturian, A. S., Vergallo, A., Farlow, M. R., et al. (2019). Revisiting the cholinergic hypothesis in Alzheimer's disease: Emerging evidence from translational and clinical research. *J. Prev. Alzheimer's Dis.* 6, 2–15. doi:10.14283/jpad.2018.43

Hannan, M. A., Kang, J.-Y., Mohibbullah, M., Hong, Y.-K., Lee, H., Choi, J.-S., et al. (2014). Moringa oleifera with promising neuronal survival and neurite outgrowth promoting potentials. *J. Ethnopharmacol.* 152, 142–150. doi:10.1016/j.jep.2013.12.036

Hou, T.-T., Yang, H.-Y., Wang, W., Wu, Q.-Q., Tian, Y.-R., and Jia, J.-P. (2018). Sulforaphane inhibits the generation of amyloid-β oligomer and promotes spatial learning and memory in Alzheimer's disease (PS1V97L) transgenic mice. *J. Alzheimer's Dis.* 62, 1803–1813. doi:10.3233/JAD-171110

Hoyer, S. (2004). Glucose metabolism and insulin receptor signal transduction in Alzheimer disease. *Eur. J. Pharmacol.* 490, 115–125. doi:10.1016/j.ejphar.2004.02.049

Huang, C., Pan, S., Yang, R., Ma, Y.-Y., Lu, X., You, Q., et al. (2022). Indole-3-carbinol selectively prevents chronic stress-induced depression-but not anxiety-like behaviors via suppressing pro-inflammatory cytokine production and oxido-nitrosative stress in the brain. *Front. Pharmacol.* 331, 829966. doi:10.3389/fphar.2022.829966

Huke, C. D., Taylor, L. J., Argent, S. P., and Kays, D. L. (2021). Catalyst-free hydrophosphinylation of isocyanates and isothiocyanates under low-added-solvent conditions. *ACS Sustain. Chem. Eng.* 9, 10704–10709. doi:10.1021/acssuschemeng.1c02907

Jaafaru, M. S., Nordin, N., Rosli, R., Shaari, K., Bako, H. Y., Noor, N. M., et al. (2019a). Prospective role of mitochondrial apoptotic pathway in mediating GMG-ITC to reduce cytotoxicity in H2O2-induced oxidative stress in differentiated SH-SY5Y cells. *Biomed. Pharmacother.* 119, 109445. doi:10.1016/j.biopha.2019.109445

Jaafaru, M. S., Nordin, N., Rosli, R., Shaari, K., Bako, H. Y., Saad, N., et al. (2019b). Neuroprotective effects of glucomoringin-isothiocyanate against H2O2-Induced cytotoxicity in neuroblastoma (SH-SY5Y) cells. *Neurotoxicology* 75, 89–104. doi:10.1016/j.neuro.2019.09.008

Jaja-Chimedza, A., Graf, B. L., Simmler, C., Kim, Y., Kuhn, P., Pauli, G. F., et al. (2017). Biochemical characterization and anti-inflammatory properties of an isothiocyanate-enriched moringa (Moringa oleifera) seed extract. *PLoS One* 12, e0182658. doi:10.1371/journal.pone.0182658

Jeong, J. H., Kim, J.-J., Bak, D. H., Yu, K. S., Lee, J. H., Lee, N. S., et al. (2015). Protective effects of indole-3-carbinol-loaded poly (lactic-co-glycolic acid) nanoparticles against glutamate-induced neurotoxicity. *J. Nanosci. Nanotechnol.* 15, 7922–7928. doi:10.1166/jnn.2015.11219

Jhang, K. A., Park, J.-S., Kim, H.-S., and Chong, Y. H. (2018). Sulforaphane rescues amyloid-β peptide-mediated decrease in MerTK expression through its anti-inflammatory effect in human THP-1 macrophages. *J. Neuroinflammation* 15, 75–12. doi:10.1186/s12974-018-1112-x

Jo, C., Gundemir, S., Pritchard, S., Jin, Y. N., Rahman, I., and Johnson, G. V. W. (2014). Nrf2 reduces levels of phosphorylated tau protein by inducing autophagy adaptor protein NDP52. *Nat. Commun.* 5, 3496–3513. doi:10.1038/ncomms4496

Kabil, O., and Banerjee, R. (2010). Redox biochemistry of hydrogen sulfide. *J. Biol. Chem.* 285, 21903–21907. doi:10.1074/jbc.R110.128363

Kamal, R. M., Abdull Razis, A. F., Mohd Sukri, N. S., Perimal, E. K., Ahmad, H., Patrick, R., et al. (2022). Beneficial health effects of glucosinolates-derived isothiocyanates on cardiovascular and neurodegenerative diseases. *Molecules* 27, 624. doi:10.3390/molecules27030624

Karanikolopoulou, S., Revelou, P.-K., Xagoraris, M., Kokotou, M. G., and Constantinou-Kokotou, V. (2021). Current methods for the extraction and analysis of isothiocyanates and indoles in cruciferous vegetables. *Analytica* 2, 93–120. doi:10.3390/analytica2040011

Kim, H. W., Kim, J., Kim, J., Lee, S., Choi, B.-R., Han, J.-S., et al. (2014). 3, 3′-Diindolylmethane inhibits lipopolysaccharide-induced microglial hyperactivation and attenuates brain inflammation. *Toxicol. Sci.* 137, 158–167. doi:10.1093/toxsci/kft240

Kim, J. (2021). Pre-clinical neuroprotective evidences and plausible mechanisms of sulforaphane in Alzheimer's Disease. *Int. J. Mol. Sci.* 22, 2929. doi:10.3390/ijms22062929

Kołodziejski, D., Piekarska, A., Hanschen, F. S., Pilipczuk, T., Tietz, F., Kusznierewicz, B., et al. (2019). Relationship between conversion rate of glucosinolates to isothiocyanates/indoles and genotoxicity of individual parts of Brassica vegetables. *Eur. Food Res. Technol.* 245, 383–400. doi:10.1007/s00217-018-3170-9

Komiskey, H., Peffley, D., Dang, T., McCain, J., McGallagher, D., Plascencia, M., et al. (2022). *Effect of sulforaphane on hyperphosphorylation of tau*.

Kraft, A. D., and Harry, G. J. (2011). Features of microglia and neuroinflammation relevant to environmental exposure and neurotoxicity. *Int. J. Environ. Res. Public Health* 8, 2980–3018. doi:10.3390/ijerph8072980

Lashley, T., Schott, J. M., Weston, P., Murray, C. E., Wellington, H., Keshavan, A., et al. (2018). Molecular biomarkers of Alzheimer's disease: Progress and prospects. *Dis. Model. Mech.* 11, dmm031781. doi:10.1242/dmm.031781

Latronico, T., Larocca, M., Milella, S., Fasano, A., Rossano, R., and Liuzzi, G. M. (2021). Neuroprotective potential of isothiocyanates in an *in vitro* model of neuroinflammation. *Inflammopharmacology* 29, 561–571. doi:10.1007/s10787-020-00772-w

Lee, C., Park, G. H., Lee, S.-R., and Jang, J.-H. (2013). Attenuation of-amyloid-induced oxidative cell death by sulforaphane via activation of NF-E2-related factor 2. *Oxid. Med. Cell. Longev.* 2013, 313510. doi:10.1155/2013/313510

Lee, S., Kim, J., Seo, S. G., Choi, B.-R., Han, J.-S., Lee, K. W., et al. (2014). Sulforaphane alleviates scopolamine-induced memory impairment in mice. *Pharmacol. Res.* 85, 23–32. doi:10.1016/j.phrs.2014.05.003

Li, S., Jin, M., Liu, L., Dang, Y., Ostaszewski, B. L., and Selkoe, D. J. (2018). Decoding the synaptic dysfunction of bioactive human AD brain soluble Aβ to inspire novel therapeutic avenues for Alzheimer's disease. *Acta Neuropathol. Commun.* 6, 121–216. doi:10.1186/s40478-018-0626-x

Loewenstein, R. J. (2022). Dissociation debates: Everything you know is wrong. *Dialogues Clin. Neurosci.* 20, 229–242. doi:10.31887/DCNS.2018.20.3/rloewenstein

Lopez-Rodriguez, N. A., Gaytán-Martínez, M., de la Luz Reyes-Vega, M., and Loarca-Piña, G. (2020). Glucosinolates and isothiocyanates from Moringa oleifera: Chemical and biological approaches. *Plant Foods Hum. Nutr.* 75, 447–457. doi:10.1007/s11130-020-00851-x

Luang-In, V., Narbad, A., Nueno-Palop, C., Mithen, R., Bennett, M., and Rossiter, J. T. (2014). The metabolism of methylsulfinylalkyl-and methylthioalkyl-glucosinolates by a selection of human gut bacteria. *Mol. Nutr. Food Res.* 58, 875–883. doi:10.1002/mnfr.201300377

Ma, Y.-S., Hsiao, Y.-T., Lin, J.-J., Liao, C.-L., Lin, C.-C., and Chung, J.-G. (2017). Phenethyl isothiocyanate (PEITC) and benzyl isothiocyanate (BITC) inhibit human melanoma A375. S2 cell migration and invasion by affecting MAPK signaling pathway *in vitro*. *Anticancer Res.* 37, 6223–6234. doi:10.21873/anticanres.12073

Mahaman, Y. A. R., Huang, F., Wu, M., Wang, Y., Wei, Z., Bao, J., et al. (2018). Moringa oleifera alleviates homocysteine-induced Alzheimer's disease-like pathology and cognitive impairments. *J. Alzheimer's Dis.* 63, 1141–1159. doi:10.3233/JAD-180091

Mahmoud, M. S., El-Kott, A. F., AlGwaiz, H. I. M., and Fathy, S. M. (2022). Protective effect of Moringa oleifera Lam. leaf extract against oxidative stress, inflammation, depression, and apoptosis in a mouse model of hepatic encephalopathy. *Environ. Sci. Pollut. Res.* 29, 83783–83796. doi:10.1007/s11356-022-21453-x

Mammana, S., Gugliandolo, A., Cavalli, E., Diomede, F., Iori, R., Zappacosta, R., et al. (2019). Human gingival mesenchymal stem cells pretreated with vesicular moringin nanostructures as a new therapeutic approach in a mouse model of spinal cord injury. *J. Tissue Eng. Regen. Med.* 13, 1109–1121. doi:10.1002/term.2857

Martelli, A., Citi, V., Testai, L., Brogi, S., and Calderone, V. (2020). Organic isothiocyanates as hydrogen sulfide donors. *Antioxidants Redox Signal* 32, 110–144. doi:10.1089/ars.2019.7888

Michalska, P., Buendia, I., Duarte, P., FernandezMendivil, C., Negredo, P., Cuadrado, A., et al. (2020). Melatonin-sulforaphane hybrid ITH12674 attenuates glial response *in vivo* by blocking LPS binding to MD2 and receptor oligomerization. *Pharmacol. Res.* 152, 104597. doi:10.1016/j.phrs.2019.104597

Mithen, R. F., Dekker, M., Verkerk, R., Rabot, S., and Johnson, I. T. (2000). The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. *J. Sci. Food Agric.* 80, 967–984. doi:10.1002/(SICI)1097-0010(20000515)80:7<967::AID-JSFA597>3.0.CO;2-V

Morroni, F., Sita, G., Graziosi, A., Turrini, E., Fimognari, C., Tarozzi, A., et al. (2018). Protective effects of 6-(methylsulfinyl) hexyl isothiocyanate on Aβ1-42-induced cognitive deficit, oxidative stress, inflammation, and apoptosis in mice. *Int. J. Mol. Sci.* 19, 2083. doi:10.3390/ijms19072083

Okubo, T., Washida, K., and Murakami, A. (2010). Phenethyl isothiocyanate suppresses nitric oxide production via inhibition of phosphoinositide 3-kinase/Akt-induced IFN-γ secretion in LPS-activated peritoneal macrophages. *Mol. Nutr. Food Res.* 54, 1351–1360. doi:10.1002/mnfr.200900318

Oliviero, T., Verkerk, R., and Dekker, M. (2018). Isothiocyanates from Brassica vegetables—Effects of processing, cooking, mastication, and digestion. *Mol. Nutr. Food Res.* 62, 1701069. doi:10.1002/mnfr.201701069

Omotoso, O. D., Olumorin, O. I., Sunday, A., Aderemi, A. S., and Ogbonna, E. (2019). Neuroprotective properties of Moringa Oleifera in cadmium and herbal alcoholic beverage induced frontal cortex damage in Wistar Rats. *Ip. Indian J. Neurosci.* 5, 206–213. doi:10.18231/j.ijn.2019.034

Onasanwo, S. A., Adamaigbo, V. O., Adebayo, O. G., and Eleazer, S. E. (2021). Moringa oleifera-supplemented diet protect against cortico-hippocampal neuronal degeneration in scopolamine-induced spatial memory deficit in mice: Role of oxido-inflammatory and cholinergic neurotransmission pathway. *Metab. Brain Dis.* 36, 2445–2460. doi:10.1007/s11011-021-00855-9

Park, H.-M., Kim, J.-A., and Kwak, M.-K. (2009). Protection against amyloid beta cytotoxicity by sulforaphane: Role of the proteasome. *Arch. Pharm. Res.* 32, 109–115. doi:10.1007/s12272-009-1124-2

Patterson, C. (2018). *World alzheimer report 2018.*

Qin, C.-Z., Zhang, X., Wu, L.-X., Wen, C.-J., Hu, L., Lv, Q.-L., et al. (2015). Advances in molecular signaling mechanisms of β-phenethyl isothiocyanate antitumor effects. *J. Agric. Food Chem.* 63, 3311–3322. doi:10.1021/jf504627e

Ramawat, K. G., and Arora, J. (2021). "Medicinal plants domestication, cultivation, improvement, and alternative technologies for the production of high value therapeutics: An overview," in *Medicinal plants, sustainable development and biodiversity*. Editor H. M. Ekiert (Springer Nature Switzerland AG), 28, 1–29.

Sailaja, B. S., Hassan, S., Cohen, E., Tmenova, I., Farias-Pereira, R., Verzi, M. P., et al. (2022). Moringa isothiocyanate-1 inhibits LPS-induced inflammation in mouse myoblasts and skeletal muscle. *PLoS One* 17, e0279370. doi:10.1371/journal.pone.0279370

Schaffert, L.-N., and Carter, W. G. (2020). Do post-translational modifications influence protein aggregation in neurodegenerative diseases: A systematic review. *Brain Sci.* 10, 232. doi:10.3390/brainsci10040232

Schneider, L. S. (2022). A critical review of cholinesterase inhibitors as a treatment modality in Alzheimer's disease. *Dialogues Clin. Neurosci.* 2, 111–128. doi:10.31887/DCNS.2000.2.2/lschneider

Sestito, S., Pruccoli, L., Runfola, M., Citi, V., Martelli, A., Saccomanni, G., et al. (2019). Design and synthesis of H2S-donor hybrids: A new treatment for Alzheimer's disease? *Eur. J. Med. Chem.* 184, 111745. doi:10.1016/j.ejmech.2019.111745

Shakour, Z. T., Shehab, N. G., Gomaa, A. S., Wessjohann, L. A., and Farag, M. A. (2022). Metabolic and biotransformation effects on dietary glucosinolates, their bioavailability, catabolism and biological effects in different organisms. *Biotechnol. Adv.* 54, 107784. doi:10.1016/j.biotechadv.2021.107784

Sharif, A. H., Iqbal, M., Manhoosh, B., Gholampoor, N., Ma, D., Marwah, M., et al. (2023). Hydrogen sulphide-based therapeutics for neurological conditions: Perspectives and challenges. *Neurochem. Res.* 48, 1981–1996. doi:10.1007/s11064-023-03887-y

Shree, B., Kumar, S., Sharma, S., and Katoch, V. (2022). Functional significance of underutilized high value cruciferous vegetables-an exotic gleam in the gloomy guise of their functional importance. *South Afr. J. Bot.* 145, 420–437. doi:10.1016/j.sajb.2022.02.028

Sikorska-Zimny, K., and Beneduce, L. (2021). The metabolism of glucosinolates by gut microbiota. *Nutrients* 13, 2750. doi:10.3390/nu13082750

Silvestro, S., Chiricosta, L., Gugliandolo, A., Iori, R., Rollin, P., Perenzoni, D., et al. (2021). The moringin/α-CD pretreatment induces neuroprotection in an *in vitro* model of Alzheimer's disease: A transcriptomic study. *Curr. Issues Mol. Biol.* 43, 197–214. doi:10.3390/cimb43010017

Sita, G., Hrelia, P., Tarozzi, A., and Morroni, F. (2016). Isothiocyanates are promising compounds against oxidative stress, neuroinflammation and cell death that may benefit neurodegeneration in Parkinson's disease. *Int. J. Mol. Sci.* 17, 1454. doi:10.3390/ijms17091454

Subedi, L., Venkatesan, R., and Kim, S. Y. (2017). Neuroprotective and anti-inflammatory activities of allyl isothiocyanate through attenuation of JNK/NF-κB/TNF-α signaling. *Int. J. Mol. Sci.* 18, 1423. doi:10.3390/ijms18071423

Sutalangka, C., Wattanathorn, J., Muchimapura, S., and Thukham-mee, W. (2013). Moringa oleifera mitigates memory impairment and neurodegeneration in animal model of age-related dementia. *Oxid. Med. Cell. Longev.* 9, 695936, doi:10.1155/2013/695936

Tabassum, R., and Jeong, N. Y. (2019). Potential for therapeutic use of hydrogen sulfide in oxidative stress-induced neurodegenerative diseases. *Int. J. Med. Sci.* 16, 1386–1396. doi:10.7150/ijms.36516

Tian, S., Liu, X., Lei, P., Zhang, X., and Shan, Y. (2018). Microbiota: A mediator to transform glucosinolate precursors in cruciferous vegetables to the active isothiocyanates. *J. Sci. Food Agric.* 98, 1255–1260. doi:10.1002/jsfa.8654

Trio, P. Z., Fujisaki, S., Tanigawa, S., Hisanaga, A., Sakao, K., and Hou, D.-X. (2016). DNA microarray highlights Nrf2-mediated neuron protection targeted by Wasabi-derived isothiocyanates in IMR-32 cells. *Gene Regul. Syst. Bio* 10, 73–83. GRSB-S39440. doi:10.4137/GRSB.S39440

Uto, T., Fujii, M., and Hou, D.-X. (2005). Inhibition of lipopolysaccharide-induced cyclooxygenase-2 transcription by 6-(methylsulfinyl) hexyl isothiocyanate, a chemopreventive compound from Wasabia japonica (Miq) Matsumura, in mouse macrophages. *Biochem. Pharmacol.* 70, 1772–1784. doi:10.1016/j.bcp.2005.09.023

Verkerk, R., Dekker, M., and Jongen, W. M. F. (1998). *Glucosinolates: Natural toxicants in food*, 29–53. ed by Watson D.

Verkerk, R., Schreiner, M., Krumbein, A., Ciska, E., Holst, B., Rowland, I., et al. (2009). Glucosinolates in Brassica vegetables: The influence of the food supply chain on intake, bioavailability and human health. *Mol. Nutr. Food Res.* 53, S219. doi:10.1002/mnfr.200800065

Wagner, A. E., Boesch-Saadatmandi, C., Dose, J., Schultheiss, G., and Rimbach, G. (2012). Anti-inflammatory potential of allyl-isothiocyanate–role of Nrf2, NF-κB and microRNA-155. *J. Cell. Mol. Med.* 16, 836–843. doi:10.1111/j.1582-4934.2011.01367.x

Wagner, A. E., Sturm, C., Piegholdt, S., Wolf, I. M. A., Esatbeyoglu, T., De Nicola, G. R., et al. (2015). Myrosinase-treated glucoerucin is a potent inducer of the Nrf2 target gene heme oxygenase 1—studies in cultured HT-29 cells and mice. *J. Nutr. Biochem.* 26, 661–666. doi:10.1016/j.jnutbio.2015.01.004

Wang, Z., Chen, Q., Wang, J., Yu, L., and Chen, L. (2020). Sulforaphane mitigates LPS-induced neuroinflammation through modulation of Cezanne/NF-κB signalling. *Life Sci.* 262, 118519. doi:10.1016/j.lfs.2020.118519

Xia, Y., Kong, L., Yao, Y., Jiao, Y., Song, J., Tao, Z., et al. (2015). Osthole confers neuroprotection against cortical stab wound injury and attenuates secondary brain injury. *J. Neuroinflammation* 12, 155–211. doi:10.1186/s12974-015-0373-x

Yang, W., Liu, Y., Xu, Q.-Q., Xian, Y.-F., and Lin, Z.-X. (2020). Sulforaphene ameliorates neuroinflammation and hyperphosphorylated tau protein via regulating the PI3K/Akt/GSK-3β pathway in experimental models of Alzheimer's disease. *Oxid. Med. Cell. Longev.* 2020, 4754195. doi:10.1155/2020/4754195

Yang, W., Xu, Q.-Q., Yuan, Q., Xian, Y.-F., and Lin, Z.-X. (2023). Sulforaphene, a CDK5 Inhibitor, attenuates cognitive deficits in a transgenic mouse model of Alzheimer's disease via reducing Aβ Deposition, tau hyperphosphorylation and synaptic dysfunction. *Int. Immunopharmacol.* 114, 109504. doi:10.1016/j.intimp.2022.109504

Yehuda, H., Soroka, Y., Zlotkin-Frušić, M., Gilhar, A., Milner, Y., and Tamir, S. (2012). Isothiocyanates inhibit psoriasis-related proinflammatory factors in human skin. *Inflamm. Res.* 61, 735–742. doi:10.1007/s00011-012-0465-3

Zhang, J., Zhang, R., Zhan, Z., Li, X., Zhou, F., Xing, A., et al. (2017). Beneficial effects of sulforaphane treatment in Alzheimer's disease may be mediated through reduced HDAC1/3 and increased P75NTR expression. *Front. Aging Neurosci.* 9, 121. doi:10.3389/fnagi.2017.00121

Zhang, L., Chen, C., Mak, M. S. H., Lu, J., Wu, Z., Chen, Q., et al. (2020). Advance of sporadic Alzheimer's disease animal models. *Med. Res. Rev.* 40, 431–458. doi:10.1002/med.21624

Zhang, R., Miao, Q.-W., Zhu, C.-X., Zhao, Y., Liu, L., Yang, J., et al. (2015). Sulforaphane ameliorates neurobehavioral deficits and protects the brain from amyloid β deposits and peroxidation in mice with Alzheimer-like lesions. *Am. J. Alzheimer's Dis. Other Dementias®* 30, 183–191. doi:10.1177/1533317514542645

Zhang, R., Zhang, J., Fang, L., Li, X., Zhao, Y., Shi, W., et al. (2014). Neuroprotective effects of sulforaphane on cholinergic neurons in mice with Alzheimer's disease-like lesions. *Int. J. Mol. Sci.* 15, 14396–14410. doi:10.3390/ijms150814396

Zhang, S., and Zhang, S. (2017). Oral absorption, distribution, metabolism, and excretion of icaritin in rats by Q-TOF and UHPLC–MS/MS. *Drug Test. Anal.* 9, 1604–1610. doi:10.1002/dta.2188

Zhao, F., Zhang, J., and Chang, N. (2018). Epigenetic modification of Nrf2 by sulforaphane increases the antioxidative and anti-inflammatory capacity in a cellular model of Alzheimer's disease. *Eur. J. Pharmacol.* 824, 1–10. doi:10.1016/j.ejphar.2018.01.046

Zhao, Z., Guo, W., Xu, C., Wang, Q., Mao, C., and Wan, M. (2023). Physiological functions and donor design of hydrogen sulfide and its application in central nervous system diseases. *Chem. Eng. J.* 452, 139089. doi:10.1016/j.cej.2022.139089

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 09 August 2023
DOI 10.3389/fpls.2023.1214602

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Dae-Kyun Ro,
University of Calgary, Canada
Gajendra Singh Jeena,
Seoul National University,
Republic of Korea

*CORRESPONDENCE
Toshiya Muranaka
muranaka@bio.eng.osaka-u.ac.jp

RECEIVED 30 April 2023
ACCEPTED 20 July 2023
PUBLISHED 09 August 2023

CITATION
Istiandari P, Yasumoto S, Seki H, Fukushima EO and Muranaka T (2023) Class I and II NADPH-cytochrome P450 reductases exhibit different roles in triterpenoid biosynthesis in *Lotus japonicus*. *Front. Plant Sci.* 14:1214602. doi: 10.3389/fpls.2023.1214602

Class I and II NADPH-cytochrome P450 reductases exhibit different roles in triterpenoid biosynthesis in *Lotus japonicus*

Pramesti Istiandari[1], Shuhei Yasumoto[1,2], Hikaru Seki[1,2], Ery Odette Fukushima[1,3] and Toshiya Muranaka[1,2]*

[1]Department of Biotechnology, Graduate School of Engineering, Osaka University, Suita, Japan, [2]Industrial Biotechnology Initiative Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Suita, Japan, [3]Plant Translational Research Group, Universidad Regional Amazónica IKIAM, Tena, Ecuador

Cytochrome P450 monooxygenases (CYPs) are enzymes that play critical roles in the structural diversification of triterpenoids. To perform site-specific oxidations of the triterpene scaffold, CYPs require electrons transferred by NADPH-cytochrome P450 reductase (CPR), which is classified into two main classes, class I and class II, based on their structural difference. *Lotus japonicus* is a triterpenoids-producing model legume with one CPR class I gene (*LjCPR1*) and a minimum of two CPR class II genes (*LjCPR2-1* and *LjCPR2-2*). CPR classes I and II from different plants have been reported to be involved in different metabolic pathways. By performing gene expression analyses of *L. japonicus* hairy root culture treated with methyl jasmonate (MeJA), this study revealed that *LjCPR1, CYP716A51*, and *LUS* were down-regulated which resulted in no change in betulinic acid and lupeol content. In contrast, *LjCPR2s, bAS, CYP93E1*, and *CYP72A61* were significantly upregulated by MeJA treatment, followed by a significant increase of the precursors for soyasaponins, i.e. β-amyrin, 24-OH β-amyrin, and sophoradiol content. Triterpenoids profile analysis of *LORE1* insertion and hairy root mutants showed that the loss of the *Ljcpr2-1* gene significantly reduced soyasaponins precursors but not in *Ljcpr1* mutants. However, *Ljcpr1* and *Ljcpr2-1* mutants showed a significant reduction in lupeol and oleanolic, ursolic, and betulinic acid contents. Furthermore, *LjCPR1*, but not *LjCPR2*, was crucial for seed development, supporting the previous notion that CPR class I might support plant basal metabolism. This study suggests that CPR classes I and II play different roles in *L. japonicus* triterpenoid biosynthesis.

KEYWORDS

CRISPR/Cas9, cytochrome P450 monooxygenases (CYP), LORE1, *Lotus japonicus*, NADPH-cytochrome P450 reductases (CPR), triterpenoid biosynthesis

1 Introduction

As one of the model legumes, *Lotus japonicus* is known to accumulate diverse phytochemicals, especially triterpenoid saponins (Suzuki et al., 2019). Triterpenoid saponins are beneficial as anti-cancer and anti-inflammatory agents for humans and are also known to play an essential role as plant defensive compounds against pathogenic bacteria and herbivores. Oxidosqualene cyclases (OSCs) catalyze the first step in triterpenoid biosynthesis by performing the cyclization of 2,3-oxidosqualene to various triterpene scaffolds. Cytochrome P450 monooxygenases (CYPs) are the enzyme most responsible for the structural diversity of triterpenoids due to their capability to perform site-specific oxidations (e.g., the introduction of hydroxyl, ketone, aldehyde, carboxyl, or epoxy groups) on various triterpene skeletons to produce triterpenoid sapogenins (aglycones). The triterpenoid sapogenins are modified with different sugar moieties by glycosyltransferases to produce triterpenoid glycosides, referred to as triterpenoid saponins.

The availability of an *L. japonicus* genome database, mutant library, and the establishment of its hairy root transformation make this plant an excellent platform for studying triterpenoid biosynthesis and its regulatory mechanisms (Urbański et al., 2013). Many triterpenoid biosynthetic genes in *L. japonicus* have been characterized. At least five *L. japonicus* OSCs have been identified, among which are *LjbAS*, *LjAaS*, and *LjLUS*. Also, the functions of three *LjCYP*s involved in triterpenoid biosynthesis, that is, *LjCYP716A51*, *LjCYP72A61*, and *LjCYP93E1*, have been described before (Suzuki et al., 2019). Compared to other tissues, the roots accumulate the highest amount of total triterpenoids in *L. japonicus* plants (Suzuki et al., 2019). Thus, the roots of *L. japonicus* are the best choice for studying triterpenoid biosynthesis.

To perform site-specific oxidation of the triterpene scaffold, CYPs require electrons transferred by its redox partner, NADPH-cytochrome P450 reductase (CPR). Plants have multiple CPR genes, depending on the species, unlike mammals and fungi, which have one CPR gene (Mizutani and Ohta, 1998; Jensen and Møller, 2010). Plant CPRs are branched into two classes, CPR classes I and II (Qu et al., 2015). CPR class I generally has a shorter N-terminal membrane sequence than CPR class II (Parage et al., 2016). CPR class I is reported to be constitutively expressed and plays a role in primary or basal constitutive metabolism, while CPR class II is inducible by environmental stimuli and is involved in defense mechanisms through plant secondary metabolism (Parage et al., 2016). Different tissue expression profiles of CPR classes I and II were reported in *Withania somnifera* (Rana et al., 2013), *Panax ginseng* (Zou et al., 2021), *Camellia sinensis* (Huang et al., 2021), and *Catharanthus roseus* (Parage et al., 2016). The differences in the protein sequences and expression levels of CPR classes I and II suggest that they have different roles in plants.

CPR and CYP are both membrane-bound proteins that have been reported to be present in microsomes in a ratio of 1:15 (Shephard et al., 1983). The significantly low ratio of CPR : CYP implied a competition between a vast number of different CYPs over a small number of CPR, in which systematic regulation is required to perform their functions properly. Knocking down *CPR2* gene in *C. roseus* significantly reduced the total monoterpene indole alkaloid content, while knocking down *CPR1* did not show any change (Parage et al., 2016). ATR2 mutation decreased the electron transfer to three CYPs involved in lignin-related phenolic metabolites, C4H, C3H1, and F5H1, but had small effects on other CYPs involved in glucosinolate and flavonol glycoside biosynthesis of *Arabidopsis thaliana* (Sundin et al., 2014).

Transcript and metabolite profiling of stress/elicitor-treated plants or cell cultures can be used to determine gene function in secondary metabolism (Misra et al., 2014). CPR class II genes from *W. somnifera*, *P. ginseng*, *C. sinensis*, and *C. roseus* were highly induced by methyl jasmonate (MeJA) treatment, whereas their CPR class I genes showed less or no induction (Rana et al., 2013; Parage et al., 2016; Huang et al., 2021; Zou et al., 2021). Interestingly, this different effect of phytohormone treatment was also observed in *OSC*s and *CYP*s. RT-PCR analysis of MeJA-treated *Ocimum basilicum* showed that *ObbAS1* and *ObCYP2* were significantly and continually induced until 12 h of treatment, while the phytohormone effect on *ObbAS2*, *ObCYP1*, and *ObCYP3* was not that apparent (Misra et al., 2014). These results implied that there could also be a differential regulation of plant CPR class in triterpenoid biosynthesis of *L. japonicus* upon MeJA elicitation.

Mutants are a powerful tool for investigating gene functions, including the impact of different CPR classes on *L. japonicus*. Given its status as an extensively studied model legume, several methods have been employed to generate mutants in *L. japonicus*. The Lotus Retrotransposon 1 (*LORE1*) mutant library has been used to study the genome-wide mutagenesis of *L. japonicus*. *LORE1* is a long-terminal repeat of transposable elements induced during tissue culture regeneration in the germ line of *L. japonicus* (Urbański et al., 2013). By tagging *LORE1* elements, a useful mutant library was created. Targeted mutagenesis using the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (CRISPR/Cas9) in *L. japonicus* hairy roots has facilitated the investigation of triterpenoid biosynthetic gene function (Suzuki et al., 2019).

Therefore, this study aims to elucidate the role of different classes of LjCPRs in triterpenoid biosynthesis *in planta*. We first analyzed the effect of MeJA treatment on triterpenoid biosynthetic genes and metabolite profiles of *L. japonicus* hairy roots. This study showed that *L. japonicus* NADPH-cytochrome P450 reductase class I (*LjCPR1*) and II (*LjCPR2*) have different regulations upon MeJA addition and revealed different sets of triterpenoid biosynthetic genes co-regulated with either *LjCPR* class I or II. To confirm the involvement of LjCPR classes on these triterpenoids *in planta*, we analyzed the triterpenoid profile of *Ljcpr1* and *Ljcpr2-1 LORE1* insertion mutant plants. We observed different effects on triterpenoid profiles in *Ljcpr1* and *Ljcpr2-1* mutants. To further confirm this result, we generated *Δljcpr1* knocked-out hairy root mutants by CRISPR/Cas9, which showed a similar triterpenoid profile with *Ljcpr1 LORE1* insertion mutants. We also observed physiological changes of the *Ljcpr LORE1* insertion mutants, which revealed another possible role of LjCPR classes I and II in other metabolic pathways. This study demonstrated for the first time that CPR classes I and II have different roles in triterpenoid biosynthesis *in planta*.

2 Materials and methods

2.1 Phylogenetic and gene co-expression analysis of LjCPRs

LjCPR gene sequences were obtained from Miyakojima MG-20 (Li et al., 2020) and the Gifu v1.2 genome database (https://lotus.au.dk/expat/) by using BLAST with *A. thaliana* CPR1 and CPR2 as the query. Phylogenetic analysis on LjCPRs was conducted using 29 other CPR sequences from different plant species, which were obtained from the web-based resource for *Arabidopsis* P450, cytochrome *b*, NADPH-cytochrome P450 reductases, and family 1 glycosyltransferases (www.P450.kvl.dk), NCBI (https://www.ncbi.nlm.nih.gov/), and various plant genome databases. The classification of CPR classes I and II was based on *A. thaliana* CPR1 and CPR2, as previously reported (Mizutani and Ohta, 1998). Accession numbers of CPR genes and genome databases used in this study are listed in Supplementary Table S1. Multiple amino acid sequences were aligned using ClustalW and were used for tree construction using the neighbor-joining method with MEGA7. Gene co-expression analysis of *L. japonicus* was performed using an online transcriptomic database from the gene expression atlas web server https://lotus.au.dk/expat/ using the Gifu v1.2 genome database version. Based on the sequence identity between CPR genes of Miyakojima MG-20 and Gifu ecotype, gene IDs LotjaGi1g1v0345200.1, LotjaGi4g1v0301400.3, and LotjaGi4g1v0301300.1 from Gifu genome database were used for *LjCPR1*, *LjCPR2-1*, and *LjCPR2-2* genes in this study (Supplementary Table S2).

2.2 Plant materials and germination treatment

L. japonicus Gifu B-129 wild-type (WT) and *LORE1* insertion lines (Fukai et al., 2012; Urbański et al., 2012) were provided by Miyazaki University, Japan, and Aarhus University, Denmark, through the National BioResource Project (NBRP). Seeds of *L. japonicus* were surface-sterilized using 2% (v/v) sodium hypochlorite and 0.02% (v/v) Tween 20 for 15 min in a seesaw shaker, rinsed three times with ultrapure water obtained from a Milli-Q Synthesis system (Millipore, Burlington, MA, USA), and placed onto a 0.8% agar plate. The seeds were allowed to germinate at 23°C for 4 days in the dark and 2 days under a 16-h light:8-h dark photoperiod.

2.3 Chemicals

β-Amyrin, α-amyrin, lupeol, erythrodiol, uvaol, oleanolic acid, ursolic acid, and asiatic acid were purchased from Extrasynthese (Genay, France). Betulin and MeJA were purchased from Sigma-Aldrich (St. Louis, MO, USA). Betulinic acid was purchased from Tokyo Chemical Industry (Tokyo, Japan). Soyasapogenol B and soyasapogenol A were purchased from Tokiwa Phytochemical (Chiba, Japan). Sophoradiol, 24-hydroxy-β-amyrin, and soyasapogenol E were kindly provided by Dr. Kiyoshi Ohyama (Tokyo Institute of Technology, Japan).

2.4 Hairy root induction

Induction of hairy roots was performed as reported previously (Suzuki et al., 2019), with slight modifications. *Agrobacterium rhizogenes* ATCC15834 strains were cultured on YEB plates for 2 days and suspended in sterilized water. The roots of 7-day-old WT seedlings were cut off, and *A. rhizogenes* were infected into the cross-sections of hypocotyls. After co-cultivation for 4 days, the infected seedlings were cultured on cefotaxime-containing hairy root elongation (HRE) solid medium for 2 weeks under a 16-h light:8-h dark photoperiod. After dissection, hairy roots were cultured under dark conditions. The root tip of a randomly chosen healthy WT hairy root clone was subcultured in 5 mL of cefotaxime-containing HRE liquid medium for 2 weeks and then transferred to 5 mL of HRE liquid medium without antibiotics with shaking at 90 rpm for another 2 weeks. Isolated hairy roots were cultured for 2 months at room temperature with subculturing every 3–4 weeks. Finally, hairy roots were cultured in 100 mL of HRE liquid medium at 25°C with shaking at 90 rpm for 4 weeks.

2.5 Methyl jasmonate preparation and addition

MeJA elicitation was conducted to test its effect on triterpenoid biosynthesis in *L. japonicus* hairy roots. MeJA preparation was performed as reported previously (Akhgari et al., 2019). A 20-mM MeJA stock solution was made by dissolving it in 40% (v/v) ethanol and then was filter sterilized (0.22 μm). The 4-week culture of WT hairy root was cut into similar portions (roughly 200 mg fresh weight each), cultured into 100 mL of HRE liquid medium without antibiotics, and incubated at 25°C with shaking at 90 rpm for another 4 weeks. The 4-week-old hairy root cultures were supplemented with final concentrations of 100 μM of MeJA. For control cultures, equal volumes (500 μL) of 40% ethanol were added to 100 mL of culture medium. The hairy roots were incubated under the same conditions as mentioned above and collected after 0, 3, 6, and 12 h for gene expression analysis and 0, 12, 24, and 24 h for metabolite analysis. The samples were flash-frozen in liquid nitrogen and stored at −80°C until use.

2.6 Quantitative real-time PCR

Total RNA was extracted from 100 mg of 4-week-old frozen *L. japonicus* MeJA-treated and control hairy roots using RNeasy Plant Mini Kit (Qiagen, Germantown, MD, USA). The RNA obtained was purified using the After Tri-Reagent RNA Clean-Up Kit (Favorgen Biotech Corp., Ping Tung, Taiwan) after digesting contaminated genomic DNA with recombinant DNase I (RNase-free) (TaKaRa Bio, Shiga, Japan). First-strand cDNA was synthesized from purified total RNA by PrimeScript RT Master Mix (Perfect Real

Time) (TaKaRa Bio). qPCR analysis was performed using The LightCycler® 96 (Roche, Basel, Switzerland) and FastStart Essential DNA Green Master (Roche). The primers used for qPCR analysis are listed in Supplementary Table S3. The expression of ubiquitin (UBQ) gene was analyzed as a reference gene.

2.7 *Ljcpr1* and *Ljcpr2* loss-of-function mutant lines

The loss-of-function mutant lines used in this study were the *LORE1* mutant collection with Gifu B-129 genetic background annotated into the Miyakojima MG-20 v3.0 genome assembly. *LjCPR1* and *LjCPR2-1* were mapped into gene IDs Lj1g3v1548790.1 and Lj4g3v2107220.1 in the Miyakojima MG-20 v3.0 genome version, respectively. Several *LORE1* lines mapped to each gene ID were selected from *LORE1* mutant library (lotus.au.dk) and screened to obtain homozygous mutants. Two independent lines of *Ljcpr1* (30003941 and 30059903) and *Ljcpr2-1* (30037476 and 30065390) homozygous insertion mutants were obtained. Each mutant line contains other exonic or intronic *LORE1* insertions other than *LjCPR* genes (Supplementary Table S4). These chosen mutant lines were cultivated in soil, and their progenies were cultivated in a hydroponic system.

For soil-cultured plants, the 7-day-old WT and mutant seedlings were moved to pots with a mixture of soil and vermiculite and cultivated for 3 months. The produced seeds were then collected for hydroponic cultivation. The seed pods from three independent lines for each *Ljcpr1* and *Ljcpr2-1* homozygous insertion mutant were counted. Then, the pod length of 26 randomly selected pods from each *Ljcpr1* and *Ljcpr2-1* homozygous insertion mutant line was measured using a digital vernier caliper. The 7-day-old WT and mutant seedlings of the soil-cultured plant progenies were first cultured in 5-mL tubes containing basal nutrient solution. After 2 weeks, the hydroponic plants were scaled up into 50-mL tubes containing basal nutrient solution and cultivated until the flowering stage. The hydroponic medium was renewed weekly.

Genomic DNA was extracted and purified from the leaves using FavorPrep™ Plant Genomic DNA Extraction Mini Kit (Favorgen Biotech Corp.) to screen for homozygous *Ljcpr1* and *Ljcpr2-1* mutants using PCR with CPR-specific primers and a *LORE1*-specific primer (Supplementary Table S5; Supplementary Figure S1). PCR was performed using KOD FX Neo following the manufacturer's instructions (Toyobo, Osaka, Japan) using the same DNA concentration for all samples. Triterpenoids were extracted from the soil- and hydroponic-cultured roots of homozygous and heterozygous mutants and were analyzed as described below.

2.8 Generation of *Ljcpr1* knockout mutant hairy root lines

The multiplex guide RNA (gRNA)-expressing CRISPR-Cas9 vector, pMgP237-2A-GFP (Hashimoto et al., 2018), was used for genome editing of *L. japonicus* hairy root. The target sequences of the gRNAs (Supplementary Figure S2) were selected from *LjCPR1* gene using the web-based tool CRISPRdirect (https://crispr.dbcls.jp/) (Naito et al., 2015). Two gRNA target sequences were simultaneously transferred into the pMgP237-2A-GFP vector as described previously (Nakayasu et al., 2018), generating the T1/T2-pMgP237 vector. Three sets of different gRNA designs (Nos. 2, 4, and 5) targeting *LjCPR1* gene were constructed using the primers listed in Supplementary Table S6. *A. rhizogenes* ATCC15834 was transformed with the pMgP237 empty vector or the T1/T2-pMgP237 vector.

The induction of hairy roots was described above. Crude genomic DNA extraction and PCR were performed using KOD FX Neo following the manufacturer's instructions (Toyobo). Mutagenesis was confirmed using PCR with specific primers for each gRNA design (Supplementary Table S6; Supplementary Figure S3) and analyzed using a heteroduplex mobility assay (HMA) and the MCE-202 MultiNA microchip electrophoresis system (Shimadzu, Kyoto, Japan) following the manufacturer's instructions. The target sequences amplified from putative mutants were cloned into the pJET1.2/blunt vector (CloneJET PCR Cloning Kit; Thermo Fisher Scientific, Waltham, MA, USA). Insertion and deletion mutations were confirmed by sequencing several randomly selected clones.

Multiple sequences of nucleotides and amino acids were aligned using MEGA11 software (Tamura et al., 2021). The protein structure of *LjCPR1* from wild-type and L1-4.2 mutant lines was modeled using the SWISS-MODEL (Arnold et al., 2006). The LjCPR1 models were constructed using two crystal structure templates, A0A0R4J338.1.A (NADPH-cytochrome P450 reductase of *Glycine max*) and 5GXU.1.A (NADPH-cytochrome P450 reductase 2 of *A. thaliana*), with 87.7% and 72.8% identity, respectively. These two templates were used due to the different properties exhibited by the crystal structures. The crystal structure of A0A0R4J338.1.A presents the complete protein structure of *G. max* CPR, including the transmembrane region. Meanwhile, 5GXU.1.A is a truncated *A. thaliana* CPR crystal structure without the transmembrane region and includes the position of the covalently bonded flavin mononucleotide (FMN) ligand on the CPR protein. The protein models were then visualized using PyMOL (DeLano, 2020) to compare the conformations of the wild-type and mutant proteins.

2.9 Triterpenoid extraction from *L. japonicus* plants and hairy roots

Triterpenoid extraction was performed as reported previously (Suzuki et al., 2019), with slight modifications. Plants at the flowering stage and hairy roots were lyophilized and powdered using a multi-bead shocker (Yasui Kikai, Osaka, Japan). Powdered tissues (20.00 ± 0.3 mg) were extracted three times with 1 mL of methanol using a sonication-assisted method. Completely dry extracts were resuspended in 2 mL of MeOH in 4 M of HCl (1:1). The extracts were incubated at 80°C for 1 h to remove the sugar moieties of triterpenoid saponins. The hydrolyzed products were

extracted three times with hexane:EtOAc (1:1) and dried completely. The obtained pellet was resuspended in 500 µL of MeOH:chloroform (1:1). A portion of the solution was dried in a gas chromatography–mass spectrometry (GC-MS) vial. Additionally, 100 µL of the solution was evaporated and trimethylsilylated using a mixture of 50 µL of *N*,*N*-dimethylformamide (Kishida Chemical Co., Ltd., Osaka, Japan) and 50 µL of BSTFA : TMCS (99:1) (TCI) at 80°C for 30 min. For semi-quantitative analysis, an asiatic acid authentic standard was applied to the plant tissue powder before extraction.

2.10 GC-MS analysis

GC-MS analyses were performed as reported previously (Suzuki et al., 2019) on a 5977A MSD mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) connected to a 7890B gas chromatograph (Agilent Technologies) with an HP-5MS UI (30 m × 0.25 mm, 0.25-µm film thickness; Agilent Technologies) capillary column for qualitative analysis. The injection temperature was set at 250°C. The column temperature program was as follows: 80°C for 1 min, increase to 300°C at a rate of 20°C/min, and hold for 28 min. The carrier gas was helium at a flow rate of 1.0 mL/min. The ion source temperature was 230°C, and the quadrupole temperature was 150°C. The derivatized sample (1 µL) was injected in splitless injection mode. Peaks were identified by comparing their Rt and mass spectra with those of authentic standards (Supplementary Figure S4). Samples were analyzed in selected ion monitoring (SIM) mode for relative quantification by extracting the mass chromatogram in respective extracted ion chromatogram (EIC) for each metabolite as listed in Supplementary Table S7.

2.11 Statistical analysis

The significance of differences was determined using a one-way single-factor analysis of variance (ANOVA). The significance of the means was separated using Tukey's test. p-Values less than 0.05 ($p < 0.05$) were considered significant in this study. All the statistical analyses were performed with JASP (JASP Team 2020).

3 Result

3.1 Phylogenetic and gene co-expression analyses of *L. japonicus* CPRs

The two ecotypes of *L. japonicus* most commonly used for research are Miyakojima MG-20 and Gifu B-129 (Hashiguchi et al., 2011). The most recent Miyakojima MG-20 transcripts submitted by Li et al. (2020) showed that *L. japonicus* has one copy of CPR class I gene and two copies of CPR class II genes (Supplementary Figure S5-A). This latest genome version of Miyakojima MG-20 was obtained using the Illumina HiSeq 2500 platform and PacBio sequencing system (Li et al., 2020). Interestingly, based on a high-quality *L. japonicus* Gifu v1.2 genome database constructed using 100× PacBio read coverage and RNA-seq analysis (Kamal et al., 2020), seven CPR genes were annotated. One CPR gene belonged to CPR class I, while six CPR genes belonged to CPR class II (Supplementary Figure S5-B). In both ecotype genomes, CPR class I gene was located on chromosome 1, while several copies of CPR class II genes were located close to each other on the same locus of chromosome 4 (Supplementary Figure S5).

Based on phylogenetic tree analysis, LjCPR class I genes from Miyakojima MG-20 and Gifu B-129 were in the same clade, while the LjCPR class II genes from Miyakojima MG-20 and Gifu B-129 branched into two different clades, namely, LjCPR2-1 and LjCPR2-2 (Supplementary Figure S6). Two copies of the CPR class II in Miyakojima MG-20 genome belonged to two different LjCPR2 clades. However, from the six LjCPR class II genes found in the Gifu genome, three genes belonged to the LjCPR2-1 clade, and three genes belonged to the LjCPR2-2 clade. A multiple sequence alignment analysis and amino acid and nucleotide sequence identity matrix were constructed to evaluate homology among *LjCPR* genes (Supplementary Figure S7; Supplementary Tables S2A, B). LjCPRs showed that LjCPR class I from both ecotypes had 100% nucleotide similarity (Supplementary Tables S2A, B). All *LjCPR2-2* genes from Gifu showed 100% nucleotide similarity, but they are not identical to Miyakojima MG-20 *LjCPR2-2*, with 99.1% amino acid similarity (Supplementary Tables S2A, B). In contrast, one isoform of LjCPR2-1 from Gifu (LotjaGi4g1v0301400.3) showed a sequence identical to that of LjCPR2-1 from Miyakojima (Supplementary Table S2B).

Transcriptomic data were retrieved from the database to analyze the gene expression level of different CPR classes in *L. japonicus*. Unfortunately, the transcriptomic database was unavailable for the *L. japonicus* Miyakojima MG-20 genome (Li et al., 2020). Therefore, in this study, only the expression level of *LjCPR* genes from the Gifu v1.2 transcriptomic database was analyzed. *LjCPR1* was mapped to probe ID LotjaGi1g1v0345200. *LjCPR2-1* was mapped to probe ID LotjaGi4g1v0301400. The identical sequence of Gifu *LjCPR2-2s* was mapped to the LotjaGi4g1v0301300 probe, labeled as *LjCPR2-2* in this study. CPR class I and II genes of *L. japonicus* showed distinct expression patterns (Figure 1). As reported in our study before, based on Miyakojima MG-20 v3.0 transcriptomic database, CPR class I was found to be constitutively expressed with lower and more stable expression levels. In contrast, CPR class II was generally found to have higher expression levels than CPR class I, which varied depending on tissues and treatments (Istiandari et al., 2021). Similarly, in the *L. japonicus* Gifu v1.2 transcriptomic database, both *LjCPR2-1* and *LjCPR2-2* showed higher inducible expression than *LjCPR1* (Figure 1). However, *LjCPR2-2* expression was even lower than *LjCPR1* expression in some of the samples but seemed to be tissue-specific. Interestingly, *LjCPR2-2* was predominantly expressed in immature flowers and showed the highest expression level among the three CPR genes. Nevertheless, *LjCPR2-1* expressions were generally higher in almost all samples than those of *LjCPR2-2* and *LjCPR1*, implying *LjCPR2-1* is the dominant LjCPR class II in *L. japonicus*.

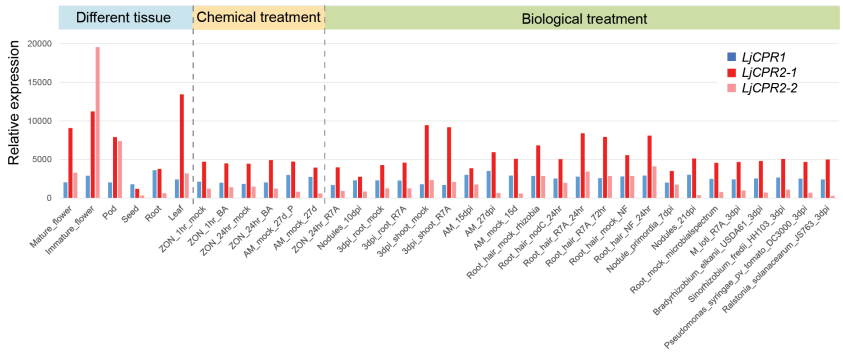


FIGURE 1
Gene co-expression analysis of *L. japonicus* using transcriptomic data from 35 different samples from different tissues, chemical treatments, or biological treatments. The transcriptomic data were obtained from the Gifu genome assembly v1.2 (https://lotus.au.dk/ of).

3.2 Effect of MeJA treatment on triterpenoid biosynthesis in *L. japonicus* hairy roots

Triterpenoid biosynthetic pathway in *L. japonicus* is shown in Figure 2. The major pentacyclic triterpene backbones, β-amyrin, α-amyrin, and lupeol, are converted from 2,3-oxidosqualene by bAS, aAS, and LUS, respectively (Figure 2). The carboxylation of β-amyrin, α-amyrin, and lupeol at the C-28 position by CYP716A51 leads to oleanolic acid, ursolic acid, and betulinic acid production (Suzuki et al., 2019). CYP93E1 catalyzes hydroxylation at the C-24 position of β-amyrin to produce 24-hydroxy β-amyrin (Shibuya et al., 2006; Seki et al., 2008). CYP72A61 converts β-amyrin into sophoradiol by adding a hydroxyl group at the C-22 position (Fukushima et al., 2011). 24-Hydroxy β-amyrin is further oxidized by CYP72A61 that catalyzes hydroxylation at the C-22 position, producing the major soyasaponin aglycone soyasapogenol B (Fukushima et al., 2011).

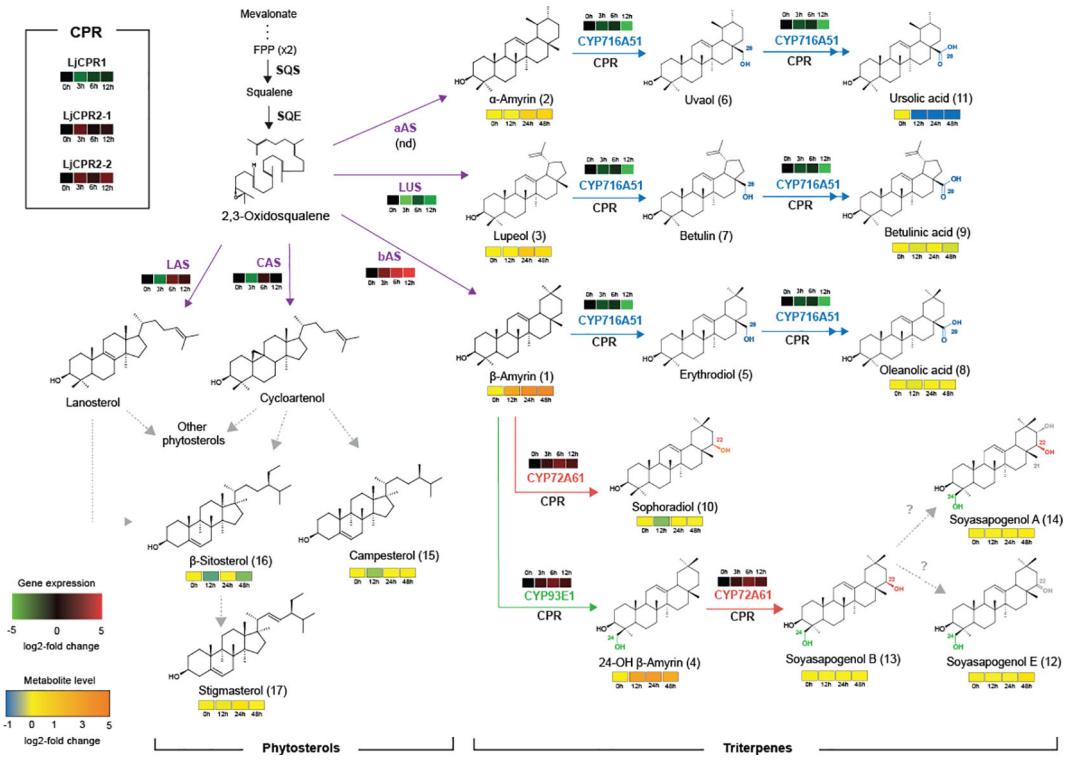


FIGURE 2
Effect of the addition of 100 μM MeJA sampled at different time points on gene expression and metabolite level on triterpenoids and phytosterol biosynthesis in *L. japonicus* hairy root. Single and double arrows indicate one and two oxidation steps, respectively. Dashed arrows indicate multiple steps. The number in parentheses next to the metabolite name refers to the chromatogram peak in Supplementary Figure S8. CYP, cytochrome P450; FPP, farnesyl pyrophosphate; SQS, squalene synthase; SQE, squalene epoxidase; bAS, β-amyrin synthase; aAS, α-amyrin synthase; LUS, lupeol synthase; LAS, lanosterol synthase; CAS, cycloartenol synthase. nd, not detected.

MeJA (100 µM) was added to 1-month-old *L. japonicus* hairy roots that were sampled at different time points to elucidate the effect of phytohormone elicitation on the *L. japonicus* triterpenoid biosynthetic pathway (Figure 2). *L. japonicus* accession Gifu B-129 was used in this study. The qRT-PCR result of extracted RNA from these treated hairy roots showed differences in the regulation of some triterpenoid biosynthetic genes (Figure 3). While *LjCPR1* gene was quickly downregulated by MeJA addition, *LjCPR2-1* and *LjCPR2-2* expressions were significantly upregulated up to four times when compared to the mock sample 3 h after MeJA addition. Interestingly, similar to the *LjCPR1* expression pattern, *CYP716A51* and *LUS* genes also showed downregulation by MeJA addition even 12 h after the treatment. However, very high and quick upregulation was observed in *bAS*, *CYP93E1*, and *CYP72A61* expression levels. The expression of *bAS* was upregulated more than 20 times when compared to the control 12 h after the treatment, while *CYP72A61* and *CYP93E1* upregulation was the highest 6 h after the treatment at approximately five times higher than those of the control. Another triterpene *OSC* gene, *aAS*, was not detected in all samples. To observe the effect of MeJA on primary metabolites such as phytosterols, *CAS* and *LAS* expressions were also analyzed. *CAS* and *LAS* are cycloartenol and lanosterol synthase, respectively, which represent the branch-off entry of phytosterol biosynthesis after 2,3-oxidosqualene cyclization (Figure 2). Both *CAS* and *LAS* expressions were instantly downregulated 3 h after MeJA addition, but then the expressions increased after 6 h and returned to levels similar to those of the control 12 h after the treatment (Figure 3).

Based on GC-MS analysis, the change in expression levels of triterpenoid biosynthetic genes due to MeJA treatment affected the triterpenoid production in *L. japonicus* hairy roots (Figures 2, 4). The triterpenoids analyzed in this study were annotated with numbers corresponding to the chromatogram peaks of MeJA-treated hairy root extracts shown in Supplementary Figure S8. Consistent with the significant upregulation of *bAS* and *CYP93E1* expressions (Figure 3), MeJA addition resulted in a significant increase in β-amyrin and 24-OH β-amyrin levels (Figure 4). The β-amyrin level was increased 10 times 12 h after the treatment and continued to increase until it reached 20 times higher than that of the control even 48 h after the treatment. Similar to β-amyrin level, the 24-OH β-amyrin level also increased rapidly after 12 h to 10 times higher than that of the control and reached a maximum of 24 h after the treatment at 15 times higher than that of the control. Sophoradiol production level also increased 24 h after the treatment. In addition, soyasapogenols showed gradually increasing production levels even 48 h after the treatment. However, no change in oleanolic acid and betulinic acid levels, and even a significant reduction of the ursolic acid level, was observed upon MeJA treatment. However, the lupeol level showed a significant increase after 24 h up to four times that of the control, similar to α-amyrin production. All triterpenoids' peaks were

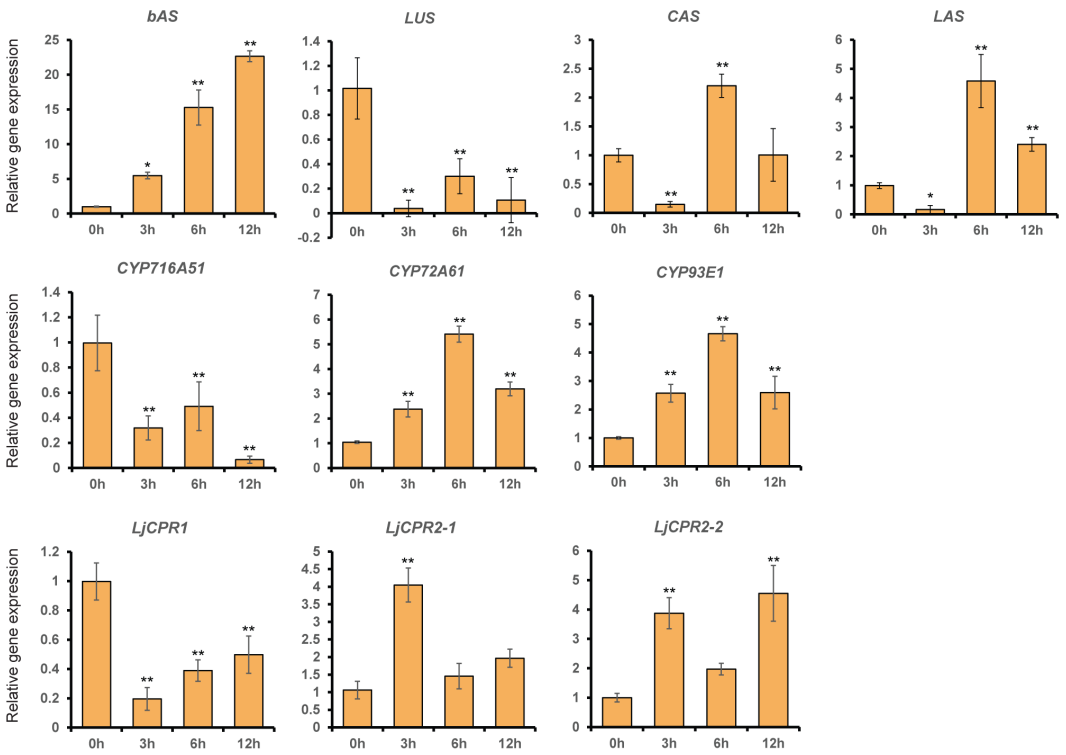


FIGURE 3
The relative expression of triterpenoid biosynthetic genes in *Lotus japonicus* hairy roots at different time periods after methyl jasmonate (MeJA) treatment. Transcript levels of *LjCPR1*, *LjCPR2-1*, *LjCPR2-2*, *bAS*, *LUS*, *CAS*, *LAS*, *CYP716A51*, *CYP72A61*, and *CYP93E1* were analyzed by qRT-PCR in *L. japonicus* hairy roots treated with 100 µM of MeJA for 0, 3, 6, and 12 h after the treatment. Relative expression levels were normalized to those of ubiquitin and are presented as fold induction relative to the control. Data represent the mean of three independent replicates ± SD. Single-factor ANOVA with Tukey's *post-hoc* test was used for statistical comparison with the control sample (0 h). Values were considered statistically significant at * $p < 0.05$ and ** $p < 0.01$. SD, standard deviation.

10.3389/fpls.2023.1214602

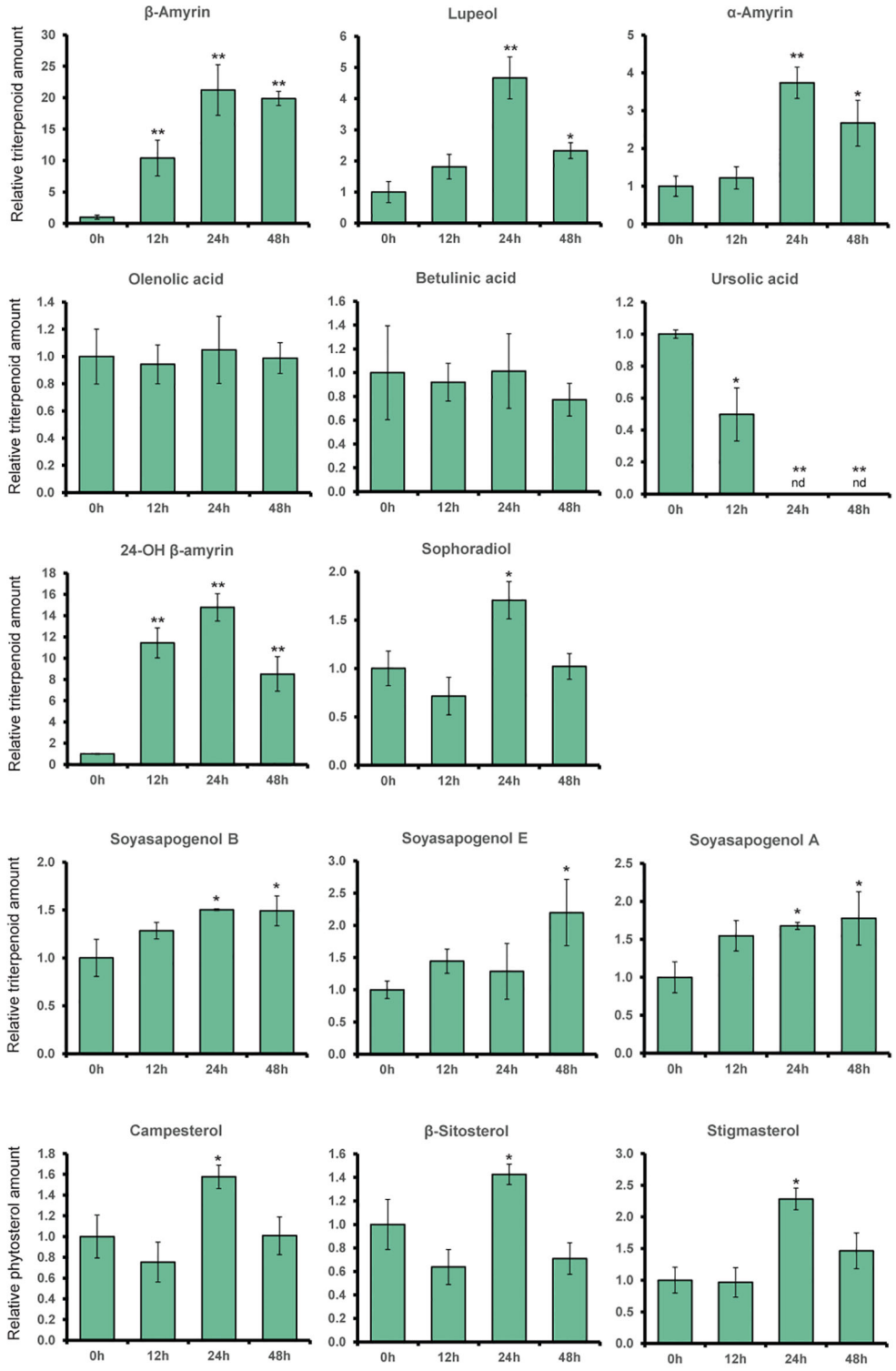


FIGURE 4
The relative amount of triterpenoids and phytosterol content of *L. japonicus* hairy roots treated with methyl jasmonate (MeJA) on different time periods. Production levels of α-amyrin, β-amyrin, lupeol, ursolic acid, oleanolic acid, betulinic acid, sophoradiol, 24-OH β-amyrin, soyasapogenols, and phytosterols were analyzed by GC-MS in *L. japonicus* hairy roots treated with 100 μM MeJA for 0, 12, 24, and 48 h after the treatment. Relative triterpenoids and phytosterol amounts were normalized to that of asiatic acid as the internal standard and are presented as fold induction relative to the control. Data represent the mean of three biological replicates ± SD. Single-factor ANOVA with Tukey's post-hoc test was used for statistical comparison with the control sample (0h). Values were considered statistically significant at *P < 0.05 and **P < 0.01. SD, standard deviation. nd, not detected.

annotated, and the chromatogram area was measured by comparing the retention time and mass spectrum with their authentic standards. The analyzed triterpenoids were the major triterpenoid constituents in *L. japonicus* (Suzuki et al., 2019), which served as the representative for triterpenoids in this study. The levels of other minor triterpenoids were too low to be detected by GC-MS. Campesterol, β-sitosterol, and stigmasterol, as the three major sterols in plants, were analyzed to observe the effect of MeJA on the phytosterol pathway. Due to the lack of standard compounds, the phytosterol peak was annotated based on similarity with a mass spectrum from the NIST library (Supplementary Figure S9). All phytosterols showed only a small increase after 24 h and returned to levels similar to those of the control after 48 h (Figure 4).

3.3 *Ljcpr1* and *Ljcpr2-1* loss-of-function mutant plants

The triterpenoid profiles of *LORE1* insertion mutant lines of *L. japonicus* accession Gifu B-129 (Fukai et al., 2012; Urbański et al., 2012) were analyzed to investigate the effect of loss-of-function of either *LjCPR1* or *LjCPR2* genes. Based on a previous study, *LUS* and *CYP716A*s genes that are responsible for lupeol and betulinic acid production, respectively, were upregulated in secondary aerenchyma of hydroponic-cultured *L. japonicus* (Suzuki et al., 2022) and soybean plants under flooded conditions (Takahashi et al., 2022). Therefore, to investigate if plant culture condition affects LjCPRs' involvement with specific CYPs, the *Ljcpr1* and *Ljcpr2-1* loss-of-function mutant lines were both cultured in soil and a hydroponic system. The seeds of two independent homozygous mutant lines, 30003941 (L1-A) and 30059903 (L1-B), which contain a single *LORE1* insertion into the first and fourth exons of *LjCPR1*, were obtained (Figure 5A). Other seeds of two independent homozygous mutant lines 30037476 (L2-1A) and 30065390 (L2-1B), which contain a non-single insertion *LORE1* insertion into the first and fourth exons of *LjCPR2-1*, were obtained (Figure 5B).

While both *Ljcpr1* mutants are single insertion homozygous mutants (Supplementary Figure S1), PCR genotyping results showed that there were heterozygous insertions on other genes other than *LjCPR2-1* in these mutant lines (Supplementary Figure S1). Therefore, the effect of other gene mutations on both *Ljcpr2-1* mutant lines should not be ruled out. Other seeds of two independent mutant lines, which contain a *LORE1* insertion on the *LjCPR2-2* exon, were also screened; however, no homozygous mutant was obtained. Therefore, only *Ljcpr1* and *Ljcpr2-1* loss-of-function mutants were analyzed in this study. *Ljcpr2-1* mutant lines might serve as a representative mutant line for LjCPR class II since they were more dominantly expressed than *LjCPR2-2* (Figure 1). Some WT allele-specific primer sets (F and R) and insertion allele-specific primer sets (F and P2) were designed for *LjCPR1* and *LjCPR2-1* genes and other genes that may have *LORE1* insertions in their exons (Supplementary Figure S1; Supplementary Table S2) according to a previous report (Urbański et al., 2012).

No significant difference in the physiology of the *Ljcpr* mutants could be observed in the plant leaves, stems, or roots. However, a notable change was observed in the seed pods (Figures 5C–E). The collected seed pods from soil-cultured mutants were then measured and counted to observe the effect of *Ljcpr* loss on the seed physiology (Figures 5C, D). The number of pods and pod length of two homozygous *Ljcpr1* mutant lines showed a more significant reduction compared to *Ljcpr2-1* mutant seed pods (Figure 5D). The physiological change of the seed pod can be observed in Figure 5E.

The triterpenoid sapogenin profile of each soil-cultured mutant plant was analyzed using GC-MS, and the result showed different profiles for *Ljcpr1* and *Ljcpr2-1* loss-of-function mutant lines (Figure 6), similar to that of the hydroponic-cultured mutants (Supplementary Figure S10). A significant difference was shown both in soil-cultured and hydroponic-cultured mutant plants. β-Amyrin, oleanolic acid, lupeol, 24-OH β-amyrin, and sophoradiol levels were significantly decreased in *Ljcpr2-1* mutant lines but showed little or no change in *Ljcpr1* mutants (Figure 6 and Supplementary Figure S10). Interestingly, both *Ljcpr1* and *Ljcpr2-1* mutant lines showed significant decreases in betulinic acid and ursolic acid and no change in soyasapogenol E and A (Figure 6 and Supplementary Figure S10). Both mutants showed lower levels of soyasapogenol B. However, the decrease in soyasapogenol B was more significant in hydroponic-cultured (Supplementary Figure S10) than in the soil-cultured mutant roots (Figure 6). As the three major phytosterols, the campesterol, β-sitosterol, and stigmasterol levels were analyzed and annotated based on mass spectra from the NIST library to investigate the effect of *Ljcpr* loss-of-function on the primary metabolisms (Supplementary Figure S9). However, due to a shift in retention time, campesterol could not be detected in hydroponic-cultured samples. Campesterol, β-sitosterol, and stigmasterol were shown to be significantly reduced in both *Ljcpr1* and *Ljcpr2-1* loss-of-function soil-cultured mutants (Figure 6).

3,4 Knockout of *Ljcpr1* gene in transgenic hairy roots

To directly confirm the involvement of different LjCPR classes on the triterpenoid biosynthetic pathway, we used the CRISPR-Cas9 system. We previously described a CRISPR-Cas9 vector, pMgP237-2A-GFP, which has been used to generate *L. japonicus* hairy root knockout mutants (Suzuki et al., 2019). Since *LjCPR2-1* and *LjCPR2-2* genes have very similar sequence identities, it was very difficult to obtain single or complete null-mutant of double-knockout *Ljcpr2*s due to a lower probability of removing all the intact sequences. Thus, in this study, only *Ljcpr1* knockout mutants were successfully obtained and analyzed. Two target sequences were simultaneously integrated into the vector to generate double tgRNA-pMgP237. Transgenic hairy roots were induced by *A. rhizogenes* ATCC15834 harboring double tgRNA-pMgP237 or the empty vector as a control. A total of nine target sequences on *LjCPR1* (Supplementary Figure S2) were selected using CRISPRdirect software (Naito et al., 2015).

Putative *Ljcpr1* hairy root mutant lines were selected using PCR and electrophoresis (HMA). Extra bands were observed in nine of the putative mutant hairy root lines but not in the control hairy root

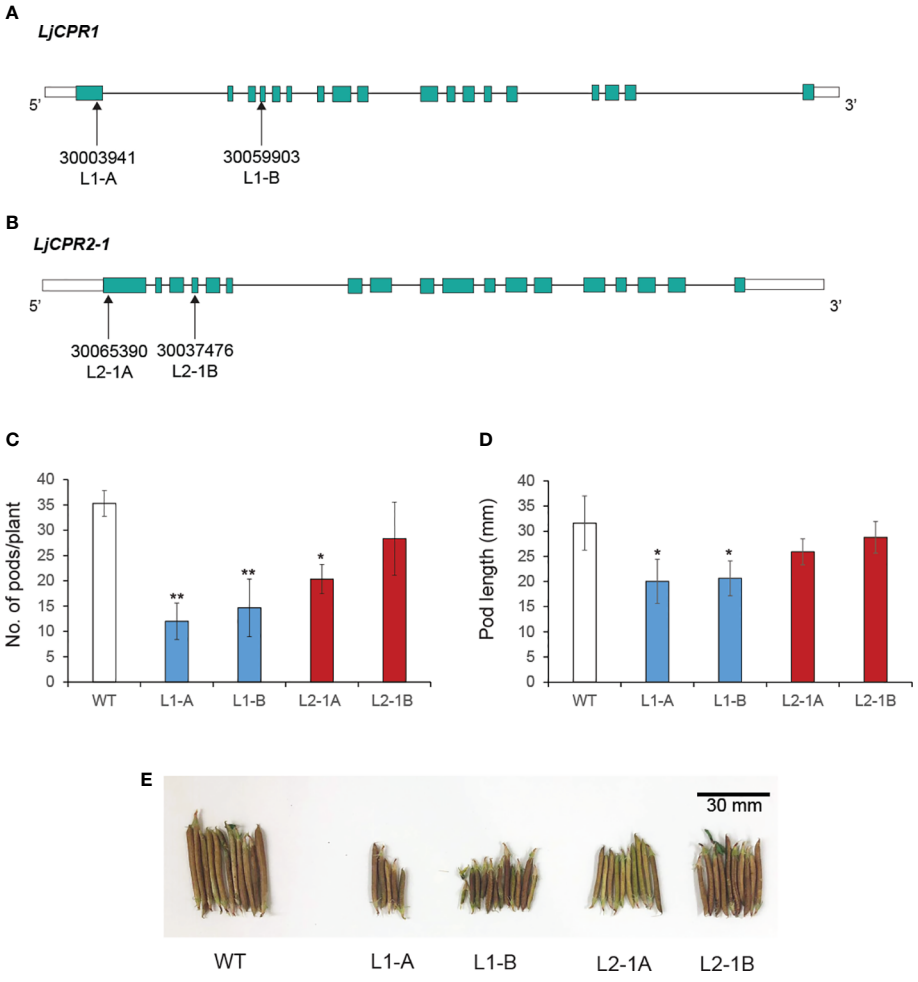


FIGURE 5
The *LORE1* insertion mutant lines selected for this study. Two homozygous *LORE1* insertion mutant lines were chosen as **(A)** *Ljcpr1* and **(B)** *Ljcpr2-1* loss-of-function mutants. The **(C)** pod numbers and **(D)** length of each mutant were quantified. The photo of representative mutant pods is shown in panel **(E)** Data represent the mean of three biological replicates in pod count **(C)** and N = 26 randomly selected pods from each mutant plant for the pod length measurement **(D)**, and both are presented as ± SD. Single-factor ANOVA with Tukey's *post-hoc* test was used for statistical comparisons. Values were considered statistically significant at * $p < 0.05$ and ** $p < 0.01$. SD, standard deviation.

lines (Supplementary Figure S3), suggesting that mutations occurred in *LjCPR1* gene and produced heteroduplex PCR fragments. Genomic DNA fragments around the target sites were cloned and sequenced to confirm mutations. Mutated alleles were not found in the control lines, EV-1 (Supplementary Figure S3). No WT sequences were detected in all obtained mutants (Supplementary Figure S3). Four mutant hairy root lines with longer nucleotide deletions were chosen for further analysis (Figure 7). These mutant hairy root lines were generated by three gRNAs (gRNA 4A, 4B, and 5B) that successfully cut the target nucleotides on the FAD and FMN/FAD hinge domain of *LjCPR1* gene (Figure 7A), which resulted in three frameshift mutant lines (L1-4.1, L1-5.1, and L1-5.2) and one non-frameshift mutant line (L1-4.2) (Figure 7B).

Additionally, we analyzed the amino acid sequences of all *LjCPR1* mutant proteins from the mutant hairy root lines. All frameshift mutations in the hairy root mutant lines L1-4.1, L1-5.1, and L1-5.2 caused premature termination, resulting in a lower number of amino acids (approximately 200) when compared to those in the WT; however, the deletion of 72-bp allele in the non-frameshift mutant L1-4.2 did not cause early termination of the protein (Supplementary File 1). The non-frameshifting mutation resulted in the absence of amino acids 222nd to 246th in the WT *LjCPR1*. Structural analysis of the mutant protein indicated that the deletion of 72-bp nucleotides encoded one strand of α-helix near the FMN-binding domain (Supplementary Figure S11A). Nevertheless, comparing the WT and mutant LjCPR1 protein structure models using two different templates revealed that the loss of 24 amino acids in the *LjCPR1* mutant line L1-4.2 did not alter the overall conformation of the protein (Supplementary Figures S11A, B). The amino acid path in the mutant, where one α-helix was missing, was directly connected to amino acid chains and continued into a β-sheet structure with the same configuration as the WT *LjCPR1* (Supplementary Figures S11A, B).

The triterpenoid profiles of control and *Ljcpr1* mutant hairy roots were then analyzed using GC-MS (Figure 8). All the frameshift mutants from different target regions showed significantly lower betulinic acid and ursolic acid content when

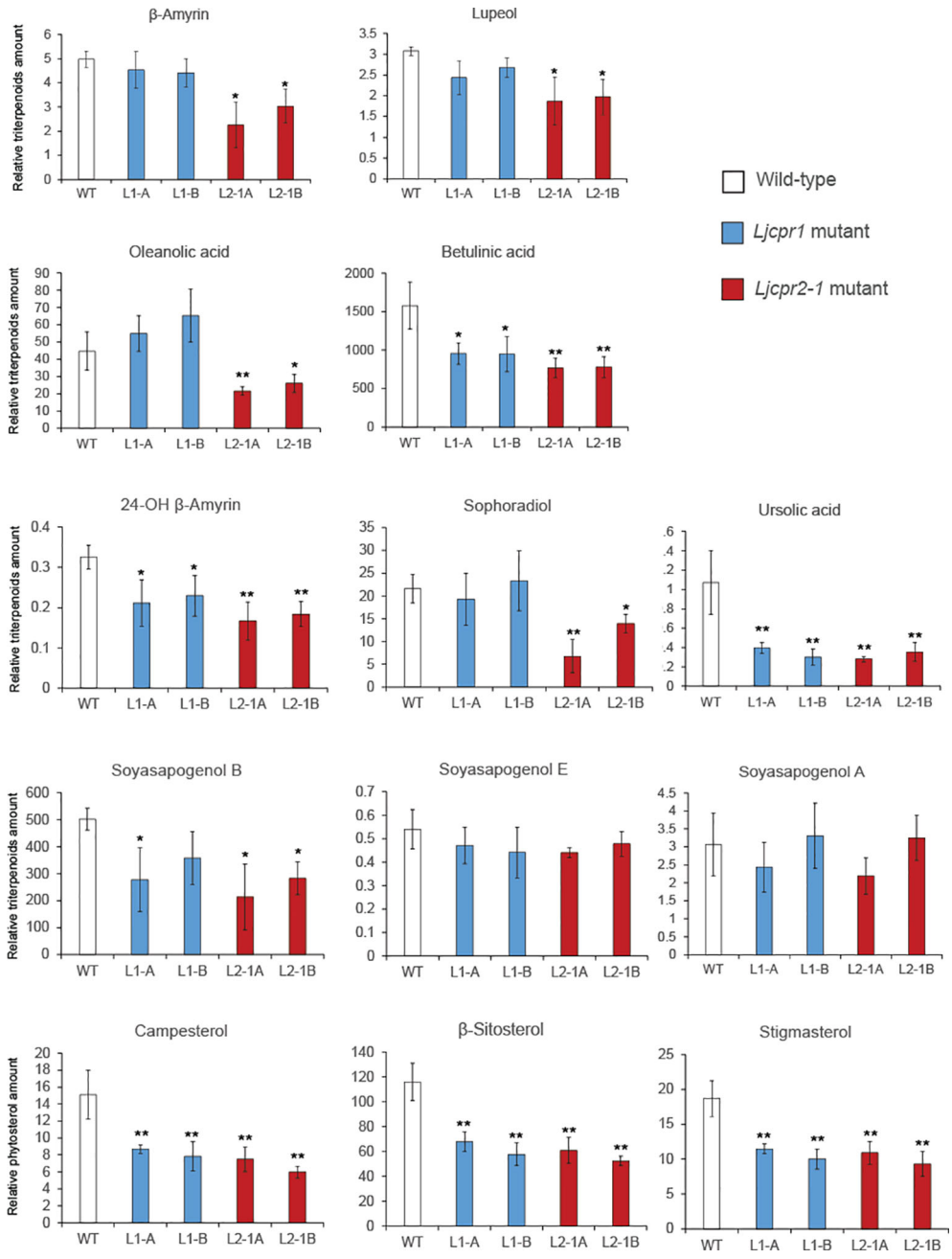


FIGURE 6
The relative amount of triterpenoids and phytosterols of soil-cultured *Ljcpr LORE1* insertion mutant roots analyzed by GC-MS. Relative triterpenoid and phytosterol amounts were normalized to those of asiatic acid as internal standard and are presented as fold induction relative to the wild-type control (WT). Data represent the mean of three biological replicates ± SD. Single-factor ANOVA with Tukey's *post-hoc* test was used for statistical comparison to control (WT). Values were considered statistically significant at * $p < 0.05$ and ** $p < 0.01$. SD, standard deviation; GC-MS, gas chromatography–mass spectrometry.

compared to the control. Frameshift mutant lines L1-5.1 and L1-5.2 showed similar levels of lupeol, α-amyrin, and soyasapogenols compared to the control, while L1-4.1 showed significantly lower levels of those triterpenes compared to the control. All frameshift mutant lines showed similar levels of β-amyrin, 24-OH β-amyrin, and sophoradiol compared to the control. Interestingly, the non-frameshift L1-4.2 mutant line showed significantly higher oleanolic acid and betulinic acid than the control. In the phytosterol biosynthesis pathway, the knockout of *Ljcpr1* genes showed significantly lower β-sitosterol and stigmasterol amounts than the control. A similar experiment was repeated in another target region of *LjCPR1* gene to generate other *Ljcpr1*-KO hairy root mutant lines (Supplementary Figure S12). Similarly, the GC-MS result on these *Ljcpr1*-KO hairy root mutants showed significantly lower betulinic and ursolic acid levels and no effect on β-amyrin, 24-OH β-amyrin, sophoradiol, and soyasapogenol contents (Supplementary Figure

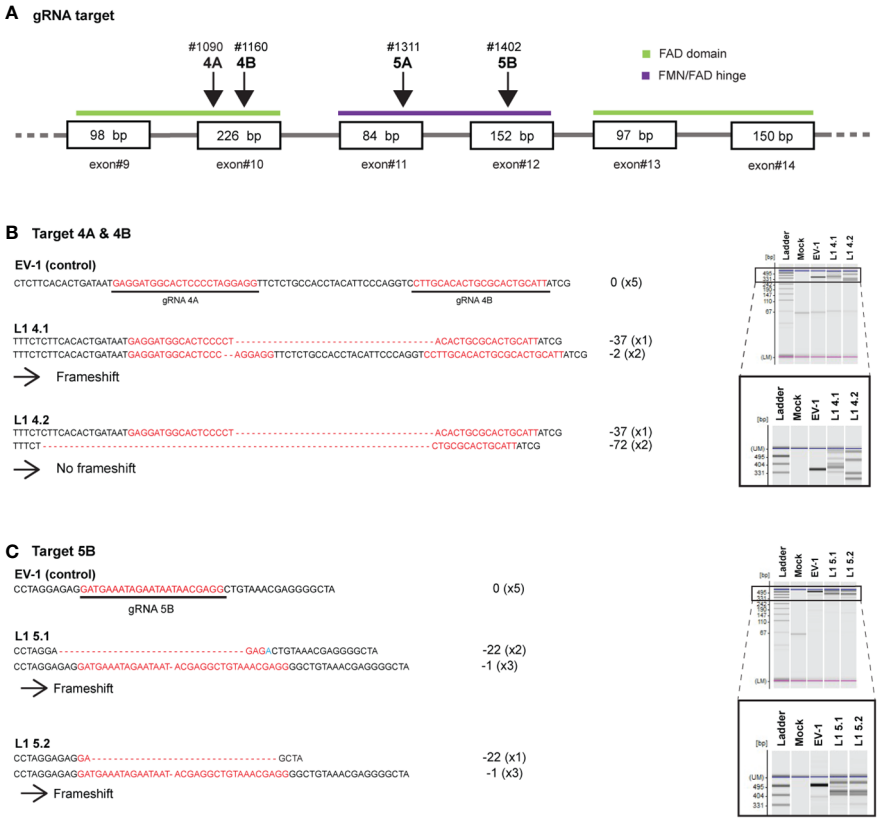


FIGURE 7
Disruption of *LjCPR1* gene in transgenic *Lotus japonicus* hairy roots by CRISPR/Cas9 system. Two sets of four gRNAs were designed to target *LjCPR1* gene in different domain regions **(A)**. The gRNA set nos. 4A and 4B **(B)** and 5B **(C)** successfully cut *LjCPR1* gene as confirmed by heteroduplex mobility assay (HMA) and sequencing, resulting in four *Ljcpr1*-KO hairy root mutant lines **(B, C)**.

S12). The total ion chromatogram (TIC) (Supplementary Figure S13) showed a significant decrease in the betulinic acid level of *Ljcpr1* knockout mutant hairy roots.

4 Discussion

Different structures of *LjCPR* genes were found in Miyakojima MG-20 and Gifu ecotype of *L. japonicus*. While both ecotypes of *L. japonicus* possess a single CPR class I gene and identical sequences, a minimum of two isoforms were found in LjCPR class II genes, branching into LjCPR2-1 and LjCPR2-2 clades in the phylogenetic tree. Only a single gene of each *LjCPR2-1* and *LjCPR2-2* gene were found in the Miyakojima MG-20 genome. However, multiple isoforms of *LjCPR2-1* and *LjCPR2-2* genes were found in the Gifu v1.2 genome. Interestingly, while Miyakojima MG-20 *LjmCPR2-1* showed an identical sequence to Gifu *LjgCPR2-1a*, *LjmCPR2-2* has no genes identical to any of the *LjgCPR2-2* isoforms. Based on expression analysis of *L. japonicus* Gifu v1.2 transcriptomic database, *LjCPR2-2* showed uniquely highest expression in immature flowers, compared to *LjCPR2-1* and *LjCPR1*. One of the main differences between the Miyakojima MG-20 and Gifu B-129 accessions is their flowering ability. Miyakojima MG-20 is known for its early and abundant flowering phenotypes, while Gifu B-129 is a late flowering phenotype with reduced flowering under fluorescent light. Therefore, the difference in the *LjCPR2-2* sequence between Miyakojima MG-20 and Gifu might contribute to explaining the phenotypic differences in the flowering ability of these ecotypes.

The MeJA treatment on *L. japonicus* hairy roots revealed segregation between LjCPR class I and II regulatory mechanisms. The results suggested that *LjCPR1* show regulatory mechanisms similar to betulinic acid biosynthetic genes *LUS* and *CYP716A51*, while *LjCPR2*s show a regulatory mechanism similar to soyasaponin biosynthetic genes *bAS*, *CYP93E1*, and *CYP72A61* regarding MeJA elicitation (Figure 2). However, even though *LUS* showed downregulated expression, the lupeol levels showed a significant increase of up to four times after 24 h when compared to control, similar to α-amyrin production (Figure 4). The increased lupeol and α-amyrin levels were possibly due to the accumulation of unconverted lupeol or α-amyrin into betulinic acid and ursolic acid, respectively, as the *CYP716A51* expression was very low. Interestingly, these triterpenoid biosynthetic pathways and CPR classes were supported by gene co-expression analysis showing that *LjCPR1* has a stronger correlation value with *CYP716A51* than *LjCPR2*, while *LjCPR2* has a stronger correlation value with *bAS*, *CYP93E1*, and *CYP72A61* than *LjCPR1*, as previously reported in Istiandari et al. (2021). The same correlation of LjCPR classes toward different CYPs is not a coincidence, suggesting that different CPR classes might have specific regulatory mechanisms

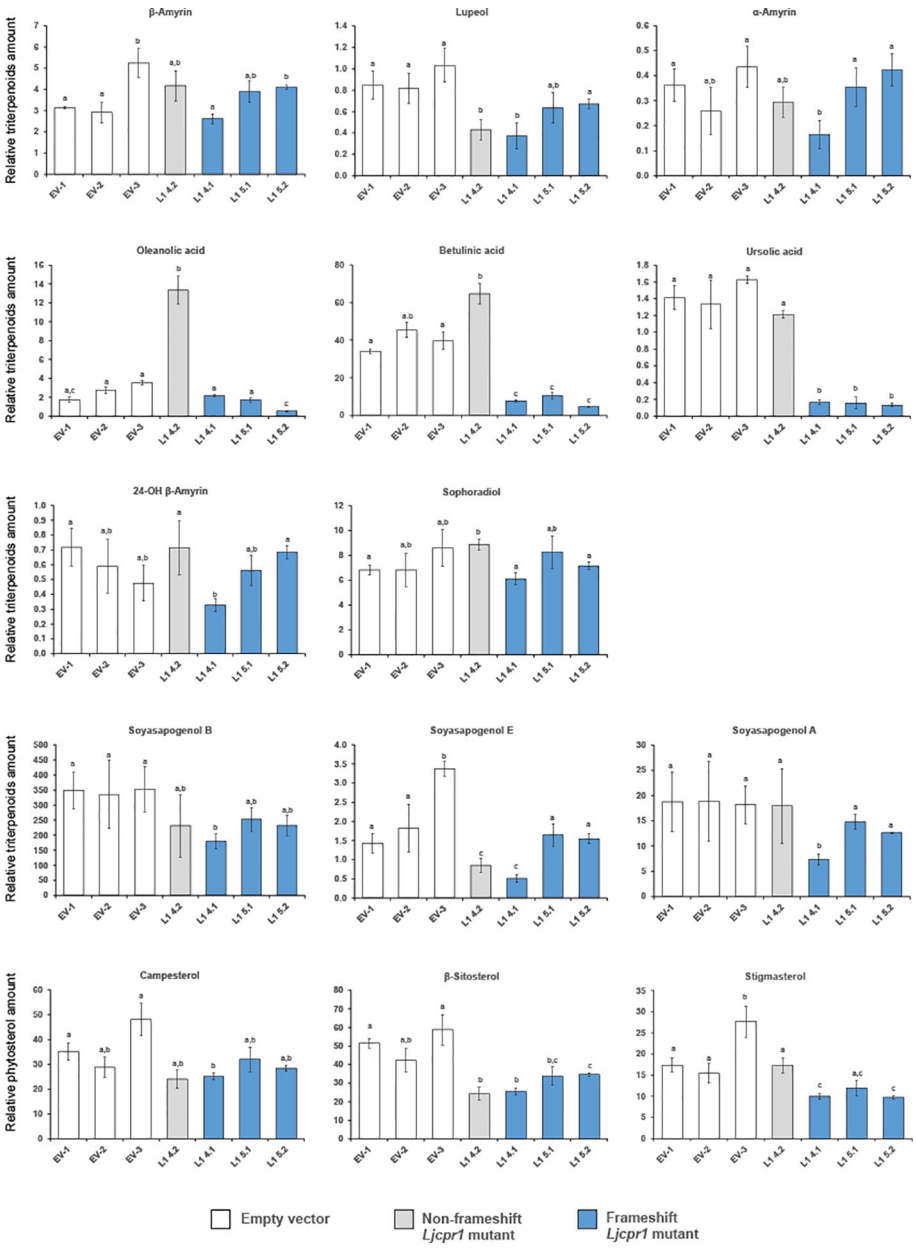


FIGURE 8
The relative triterpenoid and phytosterol contents of hairy root *Ljcpr-1* mutants analyzed using GC-MS. Relative triterpenoid and phytosterol amounts were normalized to those of asiatic acid as the internal standard and are presented as fold induction relative to the empty vector control. Data represent the mean of three technical replicates ± SD. Single-factor ANOVA was used for statistical comparisons. The different letters indicate significant differences ($p < 0.05$, one-way ANOVA followed by Tukey's test). GC-MS, gas chromatography–mass spectrometry.

with different CYPs involved in triterpenoid biosynthesis in *L. japonicus*.

Ljcpr1 and *Ljcpr2-1 LORE-1* insertion mutants also exhibited different physiological changes in pod length and number. The loss of *LjCPR1* gene resulted in a significantly reduced pod length and number than the loss of *LjCPR2-1* gene. The significant involvement of *LjCPR1* gene in pod development might be related to the higher expression of *LjCPR1* in pods and seeds compared to *LjCPR2* (Figure 1). *LjCPR1* was also strongly correlated with adenylate translocator and abscisic acid (ABA) regulation (PCC value > 0.75, Supplementary Table S8). Adenylate translocator was reported to be responsible for translocating starch for accumulation in maize endosperms (Shannon et al., 1998) to nourish the embryo (Yan et al., 2014). ABA plays a very important role in seed development, dormancy, and germination (Xiong and Zhu, 2003). Loss of *LjCPR1* gene might compromise adenylate translocator function and ABA regulation in seed and pod development in *L. japonicus*. Furthermore, the fact that the loss of *LjCPR1*, but not *LjCPR2*, compromised pod number and length in *L. japonicus* also suggests that *LjCPR1* is crucial to support CYPs and other electron acceptors involved in seed development.

Previous gene co-expression analysis results (Istiandari et al., 2021) and MeJA treatment on *L. japonicus* hairy roots in this study showed that *LjCPR2*s had a strong correlation with *bAS*, CYP93E2, and *CYP72A61* genes. This notion was supported by the analysis of *Ljcpr2-1* loss-of-function mutants in this study that subsequently showed a significant reduction of β-amyrin, 24-OH β-amyrin, and sophoradiol when compared to *Ljcpr1* mutants and WT. However, it was previously suggested that *LjCPR1* had a strong correlation with *CYP716A51* (Istiandari et al., 2021), which was shown by the significant reduction in betulinic acid and ursolic acid levels in *Ljcpr1* loss-of-function mutants. However, betulinic acid and ursolic acid were also reduced in *Ljcpr2-1* mutants, suggesting that *LjCPR1* is not acting alone in supporting *CYP716A51*. There might be synergistic work of *LjCPR1* and *LjCPR2-1* in supporting *CYP716A51*. However, in the case of *CYP93E1* and *CYP72A61*, the vital role of *LjCPR2-1* cannot be complemented by the presence of *LjCPR1*. Also interestingly, whereas both *Ljcpr1* and *Ljcpr2-1* mutants showed reductions in β-sitosterol content, only *Ljcpr1* mutants showed a lower level of stigmasterol in hydroponic-cultured mutants (Supplementary Figure S10). In *A. thaliana*, stigmasterol is synthesized from β-sitosterol by CYP710A, while campesterol and β-sitosterol biosynthesis do not involve CYPs (Morikawa et al., 2006). Therefore, the involvement of CYP710A in the stigmasterol biosynthesis pathway might be correlated stronger with *LjCPR1* than with *LjCPR2*.

Gene editing using CRISPR/Cas9 confirmed the function of a specific CPR class. *Ljcpr1*-KO mutant hairy roots exhibited a varied triterpenoid profile. The premature termination of the frameshift mutations resulted in a shorter *LjCPR1* (approximately 200 amino acids less) than the WT, with the missing 200 amino acids encoding the conserved domain critical for CPR activity, including FAD- and NADPH-binding sites. Consequently, the activity of CYP716A51, which was involved in the conversion of β-amyrin into oleanolic, betulinic, and ursolic acids, significantly reduced (Figure 8). Similar to the *LORE1* loss-of-function mutant line *Ljcpr1*, the knockout of *Ljcpr1* gene reduced the betulinic acid and ursolic acid contents, while β-amyrin, 24-OH β-amyrin, and sophoradiol contents remained unchanged in all mutants.

In contrast, the non-frameshifting mutant line, L1-4.2, displayed a triterpenoid level similar to the WT. The preserved configuration of the native protein in the mutant *LjCPR1* may allow it to function similarly to the protein of the WT (Supplementary Figures S11A, B). Metabolite profile analysis revealed differences in the production levels of lupeol, oleanolic acid, and soyasapogenol E in the *Ljcpr1*-KO hairy root mutant line L1-4.2. The absence of α-helix in the mutant *LjCPR1* may have altered specific protein–protein interactions with CYPs responsible for triterpenoid production, affecting electron transfer rates and CYP activity. Further studies, such as those producing recombinant CPR and conducting *in vitro* assays, are needed to confirm the impact of CPR mutations on CYP activity.

Although the loss of some amino acids may disrupt protein conformation, non-frameshift protein mutants have a higher probability of retaining their functions than frameshift or nonsense mutations. In a study by Bermejo-Das-Neves et al. (2014) on non-frameshifting indels in human genetic variation in Mendelian disease, out of 2,163 NFS-Indels, 757 were disease-causing variants, and 1,406 were neutral or unknown. These data suggest that the mutant *LjCPR1* from the hairy root mutant line L1-4.2 may be a neutral non-frameshifting mutant. Changes in the amino acid sequence also raise the possibility of altering the CYP : CPR binding motifs (Istiandari et al., 2021), suggesting that these non-functional LjCPR1 mutant proteins may still bind to CYPs or be replaced by other redox proteins, such as *LjCPR2s* or cytochrome b_5, due to changes in protein–protein affinity (Esteves et al., 2020).

While each mutant displayed differences in triterpenoid profiles, the result of *Ljcpr1*-KO mutant hairy roots confirmed the findings in *Ljcpr1* loss-of-function *LORE1* insertion mutant plants (Figure 6). These findings highlight the significant involvement of *LjCPR1* in betulinic acid and ursolic acid production while highlighting that *LjCPR1* is not essential for producing β-amyrin, oleanolic acid, 24-OH β-amyrin, and sophoradiol. However, our study only obtained *Ljcpr1*-KO hairy root mutants, and further research is needed to confirm the function of *LjCPR2*s in *L. japonicus*. Complementation assays of *LjCPR2-1* or the generation of *LjCPR2-1* overexpression lines should be performed to confirm the critical role of *LjCPR2-1* in the production of β-amyrin, oleanolic acid, 24-OH β-amyrin, sophoradiol, and soyasapogenols *in planta*. The unaffected oleanolic acid content in the *LORE1 Ljcpr1* mutants raises questions about the involvement of *LjCPR2s* in *bAS* and CYP716A51, which appear to be unaffected by the loss of *Ljcpr1* gene, unlike the effect on betulinic acid and ursolic acid biosynthesis.

The level of oleanolic acid was not changed in the *LORE1 Ljcpr1* mutants, which might be due to the high level of β-amyrin in *Ljcpr1* mutants compared to that of the *Ljcpr2-1* mutants. However, it is known that CPR can only support CYPs, not OSCs. Therefore, this result implies that there might be a more complex regulatory mechanism for how CPR works with CYPs. There might be a possibility that CYP716A51 can make a specific complex with different OSCs (bAS, aAS, and LUS) and create a specific metabolon. Metabolons are temporary multi-protein complexes of sequential enzymes that mediate substrate channeling (Zhang and Fernie, 2021). Metabolons have been found in several primary metabolisms, such as monolignol biosynthesis in *Populus trichocarpa* (Lin et al., 2021), or secondary metabolisms, such as camalexin biosynthesis in *A. thaliana* (Mucha et al., 2019) and dhurrin biosynthesis in *Sorghum bicolor* (Nielsen et al., 2008), which involved membrane-bound CYPs, CPRs, and other enzymes. The significant reduction of betulinic acid and ursolic acid but not in oleanolic acid level suggested that LjCPR1 is specifically involved in the betulinic acid and/or ursolic acid biosynthetic pathways, but not oleanolic acid biosynthesis. Instead, it was suggested that LjCPR2-1 might be involved more closely in the oleanolic acid biosynthetic pathway with bAS as the first enzyme converting 2,3-oxidosqualene into β-amyrin as a substrate for CYP716A51. However, more studies are needed to confirm this hypothesis. Protein–protein interaction analysis using techniques such as bimolecular fluorescence complementation or protoplast two-hybrid assay should be conducted to reveal the

structural proteins of the metabolon scaffold *in planta* (Nielsen et al., 2008; Mucha et al., 2019).

Both CPR classes were reported to be able to support the *in vitro* activities of CYPs involved in specialized metabolism (Rana et al., 2013). However, based on numerous functional analyses *in planta*, CPR class I was believed to be responsible for basal or constitutive metabolisms, while CPR class II was more responsible for adaptation and defense mechanisms, involving numerous specialized metabolisms (Rana et al., 2013; Parage et al., 2016; Huang et al., 2021; Zou et al., 2021). This study showed that LjCPR1 is closely involved with CYP716A51, a C-28 oxidase, which is involved in triterpenoid biosynthesis, one of the specialized metabolisms in *L. japonicus*. Based on the elicitor treatment, the LjCPR class I showed no induction (Figure 3), which is in line with the previous notion that only CPR class II is inducible (Figure 3). The result that *CYP716A51* was not inducible during methyl jasmonate treatment (Figure 3) might suggest a function for C-28 oxidized triterpenes in *L. japonicus* other than as a defense mechanism. In line with phenotypic change observed in the seeds of *Ljcpr1* mutant plants (Figure 5) and expression level changes in mature seeds (Figure 1), LjCPR1 and CYP716A51 might also have physiological roles in the seed development of *L. japonicus*. To validate this hypothesis, further analysis should be conducted on the expression of the β-glucuronidase (GUS) reporter gene driven by the native promoters of *LjCPR1* and *LjCPR2* in response to MeJA.

In *Medicago truncatula*, CYP716A12 is involved in oleanane-derived hemolytic sapogenin biosynthesis. During the developmental stages, such as the reproductive phase, *CYP716A12* showed a significant increase in expression level and an increase in hemolytic sapogenin content. Ten-week-old *M. truncatula cyp716a12* mutant plants showed dwarf phenotype compared to WT. This finding suggested a possible dual role of hemolytic sapogenins in defense and plant developmental growth in *M. truncatula* (Carelli et al., 2011). Functional analysis and phenotypic observation of *cyp716a51* mutant of *L. japonicus* might provide insight into LjCPR1 and CYP716A51 involvement in plant primary metabolism.

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

PI, EF, and TM designed the experiments. PI conducted all of the experiments, analyzed the results, and wrote the whole manuscript. EF, SY, HS, and TM conceived and supervised the study. SY, HS, and TM made the manuscript revisions. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Grant-in-Aid for Scientific Research [JP19H02921 to TM, 20H02913, 21K19082 to HS] and the Monbukagakusho Scholarship from MEXT to PI.

Acknowledgments

We thank Dr. Kiyoshi Ohyama (Tokyo Institute of Technology) for providing sophoradiol and 24-hydroxy-β-amyrin standards. We also thank Miyazaki University for providing seeds of *L. japonicus* Gifu B-129 accession (WT) and Aarhus University for providing seeds of *L. japonicus* Gifu B-129 accession (*LORE1* insertion lines) through the Japanese National BioResource Project. This study will be included in a dissertation submitted by Pramesti Istiandari to Osaka University in partial fulfillment of the requirement for her doctoral degree.

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.2023.1214602/full#supplementary-material

References

Akhgari, A., Laakso, I., Maaheimo, H., Choi, Y. H., Seppänen-Laakso, T., Oksman-Caldentey, K. M., et al. (2019). Methyljasmonate elicitation increases terpenoid indole alkaloid accumulation in Rhazya stricta hairy root cultures. *Plants* 8, 534. doi: 10.3390/plants8120534

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201. doi: 10.1093/bioinformatics/bti770

Bermejo-Das-Neves, C., Nguyen, H. N., Poch, O., and Thompson, J. D. (2014). A comprehensive study of small non-frameshift insertions/deletions in proteins and prediction of their phenotypic effects by a machine learning method (KD4i). *BMC Bioinf.* 15, 111. doi: 10.1186/1471-2105-15-111

Carelli, M., Biazzi, E., Panara, F., Tava, A., Scaramelli, L., Porceddu, A., et al. (2011). Medicago truncatula CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. *Plant Cell* 23, 3070–3081. doi: 10.1105/tpc.111.087312

DeLano, W. L. (2020). The pyMOL molecular graphics system, version 2.3. *Schrödinger LLC.*

Esteves, F., Campelo, D., Gomes, B. C., Urban, P., Bozonnet, S., Lautier, T., et al. (2020). The role of the FMN-domain of human cytochrome P450 oxidoreductase in its promiscuous interactions with structurally diverse redox partners. *Front. Pharmacol.* 11. doi: 10.3389/fphar.2020.00299

Fukai, E., Soyano, T., Umehara, Y., Nakayama, S., Hirakawa, H., Tabata, S., et al. (2012). Establishment of a Lotus japonicus gene tagging population using the exon-targeting endogenous retrotransposon LORE1. *Plant J.* 69, 720–730. doi: 10.1111/j.1365-313X.2011.04826.x

Fukushima, E. O., Seki, H., Ohyama, K., Ono, E., Umemoto, N., Mizutani, M., et al. (2011). CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. *Plant Cell Physiol.* 52, 2050–2061. doi: 10.1093/pcp/pcr146

Hashiguchi, M., Tsuruta, S.-I., and Akashi, R. (2011). Morphological traits of lotus japonicus (Regal) ecotypes collected in Japan. *Interdiscip. Bio Cent.* 3, 4.1–4.7. doi: 10.4051/ibc.2011.3.1.0004

Hashimoto, R., Ueta, R., Abe, C., Osakabe, Y., and Osakabe, K. (2018).). Efficient multiplex genome editing induces precise, and self-ligated type mutations in tomato plants. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00916

Huang, R., Liu, L., He, X., Wang, W., Hou, Y., Chen, J., et al. (2021). Isolation and functional characterization of multiple NADPH-cytochrome P450 reductase genes from camellia sinensis in view of catechin biosynthesis. *J. Agric. Food Chem.* 69, 14926–14937. doi: 10.1021/acs.jafc.1c04255

Istiandari, P., Yasumoto, S., Srisawat, P., Tamura, K., Chikugo, A., Suzuki, H., et al. (2021). Comparative analysis of NADPH-cytochrome P450 reductases from legumes for heterologous production of triterpenoids in transgenic saccharomyces cerevisiae. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.762546

Jensen, K., and Møller, B. L. (2010). Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry* 71, 132–141. doi: 10.1016/j.phytochem.2009.10.017

Kamal, N., Mun, T., Reid, D., Lin, J. S., Akyol, T. Y., Sandal, N., et al. (2020). Insights into the evolution of symbiosis gene copy number and distribution from a chromosome-scale Lotus japonicus Gifu genome sequence. *DNA Res.* 27, dsaa015. doi: 10.1093/dnares/dsaa015

Li, H., Jiang, F., Wu, P., Wang, K., and Cao, Y. (2020). A high-quality genome sequence of model legume lotus japonicus (Mg-20) provides insights into the evolution of root nodule symbiosis. *Genes (Basel)* 11, 483. doi: 10.3390/genes11050483

Lin, C. Y., Sun, Y., Song, J., Chen, H. C., Shi, R., Yang, C., et al. (2021). Enzyme complexes of ptr4CL and ptrHCT modulate co-enzyme A ligation of hydroxycinnamic acids for monolignol biosynthesis in populus trichocarpa. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.727932

Misra, R. C., Maiti, P., Chanotiya, C. S., Shanker, K., and Ghosh, S. (2014). Methyl jasmonate-elicited transcriptional responses and pentacyclic triterpene biosynthesis in sweet basil. *Plant Physiol.* 164, 1028–1044. doi: 10.1104/pp.113.232884

Mizutani, M., and Ohta, D. (1998). Two isoforms of NAPDH:cytochrome p450 reductase in Arabidopsis thaliana gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol.* 116, 357–367. doi: 10.1104/pp.116.1.357

Morikawa, T., Mizutani, M., Aoki, N., Watanabe, B., Saga, H., Saito, S., et al. (2006). Cytochrome P450 CYP710A encodes the sterol C-22 desaturase in Arabidopsis and tomato. *Plant Cell* 18, 1008–1022. doi: 10.1105/tpc.105.036012

Mucha, S., Heinzlmeir, S., Kriechbaumer, V., Strickland, B., Kirchhelle, C., Choudhary, M., et al. (2019). The formation of a camalexin biosynthetic metabolon. *Plant Cell* 31, 2697–2710. doi: 10.1105/tpc.19.00403

Naito, Y., Hino, K., Bono, H., and Ui-Tei, K. (2015). CRISPRdirect: Software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* 31, 1120–1123. doi: 10.1093/bioinformatics/btu743

Nakayasu, M., Akiyama, R., Lee, H. J., Osakabe, K., Osakabe, Y., Watanabe, B., et al. (2018). Generation of α-solanine-free hairy roots of potato by CRISPR/Cas9 mediated genome editing of the St16DOX gene. *Plant Physiol. Biochem.* 131, 70–77. doi: 10.1016/j.plaphy.2018.04.026

Nielsen, K. A., Tattersall, D. B., Jones, P. R., and Møller, B. L. (2008).). Metabolon formation in dhurrin biosynthesis. *Phytochemistry* 69, 88–98. doi: 10.1016/j.phytochem.2007.06.033

Parage, C., Foureau, E., Kellner, F., Burlat, V., Mahroug, S., Lanoue, A., et al. (2016). Class II cytochrome P450 reductase governs the biosynthesis of alkaloids. *Plant Physiol.* 172, 1563–1577. doi: 10.1104/pp.16.00801

Qu, X., Pu, X., Chen, F., Yang, Y., Yang, L., Zhang, G., et al. (2015). Molecular cloning, heterologous expression, and functional characterization of an NADPH-cytochrome P450 reductase gene from Camptotheca acuminata, a camptothecin-producing plant. *PLoS One* 10, 1–19. doi: 10.1371/journal.pone.0135397

Rana, S., Lattoo, S. K., Dhar, N., Razdan, S., Bhat, W. W., Dhar, R. S., et al. (2013). NADPH-cytochrome P450 reductase: molecular cloning and functional characterization of two paralogs from withania somnifera (L.) dunal. *PloS One* 8, e57068. doi: 10.1371/journal.pone.0057068

Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H., et al. (2008). Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14204–14209. doi: 10.1073/pnas.0803876105

Shannon, J. C., Pien, F. M., Cao, H., and Liu, K. C. (1998). Brittle-1, an adenylate translocator, facilitates transfer of extraplastidial synthesized ADP-glucose into amyloplasts of maize endosperms. *Plant Physiol.* 117, 1235–1252. doi: 10.1104/pp.117.4.1235

Shephard, E. A., Phillips, I. R., Bayney, R. M., Pike, S. F., and Rabin, B. R. (1983). Quantification of NADPH:cytochrome P-450 reductase in liver microsomes by a specific radioimmunoassay technique. *Biochem. J.* 211, 333–340. doi: 10.1042/bj2110333

Shibuya, M., Hoshino, M., Katsube, Y., Hayashi, H., Kushiro, T., and Ebizuka, Y. (2006). Identification of β-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. *FEBS J.* 273, 948–959. doi: 10.1111/j.1742-4658.2006.05120.x

Sundin, L., Vanholme, R., Geerinck, J., Goeminne, G., Höfer, R., Kim, H., et al. (2014). Mutation of the inducible ARABIDOPSIS THALIANA CYTOCHROME P450 REDUCTASE2 alters lignin composition and improves saccharification. *Plant Physiol.* 166, 1956–1971. doi: 10.1104/pp.114.245548

Suzuki, H., Fukushima, E. O., Shimizu, Y., Seki, H., Fujisawa, Y., Ishimoto, M., et al. (2019). Lotus japonicus triterpenoid profile and characterization of the CYP716A51 and ljCYP93E1 genes involved in their biosynthesis in planta. *Plant Cell Physiol.* 60, 2496–2509. doi: 10.1093/pcp/pcz145

Suzuki, H., Takahashi, H., Fukushima, E. O., Nakazono, M., Muranaka, T., and Seki, H. (2022). Identification of basic helix-loop-helix transcription factors that activate betulinic acid biosynthesis by RNA-sequencing of hydroponically cultured Lotus japonicus. *bioRxiv* 11, 516519. doi: 10.1101/2022.11.16.516519

Takahashi, H., Abo, C., Suzuki, H., Romsuk, J., Oi, T., Yanagawa, A., et al. (2022). Triterpenoids in aerenchymatous phellem contribute to internal root aeration and waterlogging adaptability in soybean. *Res. Square* [Preprint]. doi: 10.21203/rs.3.rs-2230730/v1 (Accessed November 8, 2022).

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

Urbański, D. F., Małolepszy, A., Stougaard, J., and Andersen, S. U. (2012). Genome-wide LORE1 retrotransposon mutagenesis and high-throughput insertion detection in Lotus japonicus. *Plant J.* 69, 731–741. doi: 10.1111/j.1365-313X.2011.04827.x

Urbański, D. F., Małolepszy, A., Stougaard, J., and Andersen, S. U. (2013). High-throughput and targeted genotyping of Lotus japonicus LORE1 insertion mutants. *Methods Mol. Biol.* 1069, 119–146. doi: 10.1007/978-1-62703-613-9_10

Xiong, L., and Zhu, J.-K. (2003). Regulation of abscisic acid biosynthesis. *Plant Physiol.* 133, 29–36. doi: 10.1104/pp.103.025395

Yan, D., Duermeyer, L., Leoveanu, C., and Nambara, E. (2014). The functions of the endosperm during seed germination. *Plant Cell Physiol.* 55, 1521–1533. doi: 10.1093/pcp/pcu089

Zhang, Y., and Fernie, A. R. (2021). Metabolons, enzyme–enzyme assemblies that mediate substrate channeling, and their roles in plant metabolism. *Plant Commun.* 2, 100081. doi: 10.1016/j.xplc.2020.100081

Zou, X., Zhang, Y., Zeng, X., Liu, T., Li, G., Dai, Y., et al. (2021). Molecular cloning and identification of NADPH cytochrome P450 reductase from panax ginseng. *Molecules* 26, 6654. doi: 10.3390/molecules26216654

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 09 August 2023
DOI 10.3389/fpls.2023.1098987

OPEN ACCESS

EDITED BY
Luan Luong Chu,
Phenikaa University, Vietnam

REVIEWED BY
Charu Chandra Giri,
Osmania University, India
Clément Cuello,
Université de Tours, France
Pankaj Kumar Sharma,
Birla Institute of Technology and Science, India
Qi Tang,
Hunan Agricultural University, China

*CORRESPONDENCE
Patel Mohana Kumara
monapatelpgatti@gmail.com;
mohanakumara.p@uhsbagalkot.edu.in

RECEIVED 15 November 2022
ACCEPTED 05 July 2023
PUBLISHED 09 August 2023

CITATION
Kumara PM, Varun E, Sanjay JR, Madhushree AH and Thimmappa R (2023) *De novo* transcriptome analysis of *Dysoxylum binectariferum* to unravel the biosynthesis of pharmaceutically relevant specialized metabolites.
Front. Plant Sci. 14:1098987.
doi: 10.3389/fpls.2023.1098987

De novo transcriptome analysis of *Dysoxylum binectariferum* to unravel the biosynthesis of pharmaceutically relevant specialized metabolites

Patel Mohana Kumara[1,2]*, Eranna Varun[2], Joshi Renuka Sanjay[2], Anchedoddi Hanumegowda Madhushree[2] and Ramesha Thimmappa[3]

[1]Department of Biotechnology and Crop Improvement, Kittur Rani Chennamma College of Horticulture, Arabhavi, University of Horticultural Sciences, Bagalkot, Karnataka, India, [2]Center for Ayurveda Biology and Holistic Nutrition, The University of Trans-Disciplinary Health Sciences and Technology (TDU), Bengaluru, Karnataka, India, [3]Amity Institute of Genome Engineering, Amity University Uttar Pradesh, Noida, India

The tropical tree, *D. binectariferum*, is a prominent source of chromone alkaloid rohitukine, which is used in the semi-syntheses of anticancer molecules such as flavopiridol and P-276-00. The biosynthetic pathway of rohitukine or its derivatives is currently unknown in plants. Here, we explored chromone alkaloid biosynthesis in *D. binectariferum* through targeted transcriptome sequencing. Illumina sequencing of leaves and roots of a year-old *D. binectariferum* seedling generated, 42.43 and 38.74 million paired-end short reads, respectively. Quality filtering and *de novo* assembly of the transcriptome generated 274,970 contigs and 126,788 unigenes with an N50 contig length of 1560 bp. The assembly generated 117,619 translated unigene protein sequences and 51,598 non-redundant sequences. Nearly 80% of these non-redundant sequences were annotated to publicly available protein and nucleotide databases, suggesting the completeness and effectiveness of the transcriptome assembly. Using the assembly, we identified a chalcone synthase (CHS) and three type III polyketide synthases (PKS-III; non-CHS type) that are likely to be involved in the biosynthesis of chromone ring/noreugenin moiety of rohitukine. We also identified key enzymes like lysine decarboxylase in the piperidine pathway that make the piperidine moiety of rohitukine. Besides these, the upstream enzymes in flavonoid biosynthesis like phenylalanine ammonia-lyase (PAL), trans-cinnamate 4-hydroxylase (C4H),4-coumarate-CoA ligase (4CL), and chalcone isomerase (CHI) have also been identified. Also, terpene synthases that are likely to be involved in the biosynthesis of various terpenoid scaffolds have been identified. Together, the *D. binectariferum* transcriptome resource forms a basis for further exploration of biosynthetic

pathways of these valuable compounds through functional validation of the candidate genes and metabolic engineering in heterologous hosts. Additionally, the transcriptome dataset generated will serve as an important resource for research on functional genomics and enzyme discovery in *D. binectariferum* and comparative analysis with other Meliaceae family members.

KEYWORDS

***Dysoxylum binectariferum*, *de novo* transcriptome, rohitukine, chromone alkaloids biosynthesis, chalcone synthases, polyketide synthases**

1 Introduction

Rohitukine, a prominent chromone alkaloid currently known to occur in five plant species belonging to *Meliaceae* and *Rubiaceae* families (Khadem and Marles, 2012; Varun et al., 2023). Rohitukine is a unique chromone alkaloid having a noreugenin chromone scaffold conjugated to a ring containing one or more nitrogen atoms (Houghton, 2002; Mohanakumara et al., 2010; Mohana Kumara, 2012). Flavopiridol (Sanofi) and P-276-00 (Piramal) two semi-synthetic derivatives of rohitukine are in the advanced stages of clinical trials for various cancer treatments (Jain et al., 2012). Flavopiridol (alvocidib; L868275; HMR-1275; NSC 649890 of Sanofi-Aventis + NCI) is an established cyclin-dependent kinases (CDK) inhibitor with broad specificity to CDK1, CDK2, and CDK4 leading to cell cycle arrest at both G1 and G2 phases (Sedlacek et al., 1996; Stadler et al., 2000; Łukasik et al., 2021). Flavopiridol is also a promising agent in inducing p53-independent apoptosis in Chronic Lymphocytic Leukaemia (CLL) and therefore this has been approved as an orphan drug for treating CLL (Christian et al., 2009; Albert et al., 2014; Mandal et al., 2021). Whereas P-276-00 is currently in phase II clinical studies for advanced refractory neoplasms and multiple myeloma (Christian et al., 2009; Borowczak et al., 2022). In addition to cancer, flavopiridol has also been shown to be effective in the treatment of arthritis and atherosclerotic plaque formation (Sekine et al., 2008; Chen et al., 2021).

Rohitukine was first reported in *Amoora rohituka* and later in *Dysoxylum binectariferum*, *Dysoxylum acutangulum* (*Meliaceae*), *Schumanniophyton magnificum* and *S. problematicum* (*Rubiaceae*) (Harmon et al., 1979; Naik et al., 1988; Ismail et al., 2009; Mohanakumara et al., 2010). Among these species, *D. binectariferum* accumulates the highest amount of rohitukine in stem bark (3-7% by dry weight). Whereas the closest relative of *D. binectariferum, D. malabaricum* does not accumulate rohitukine (Houghton, 2002; Mohanakumara et al., 2010). Also, various rohitukine derivatives such as dysoline, schumaniofioside A and chrotacumines have been reported from *D. binectariferum* (Ismail et al., 2009; Izwan Mohd Lazim et al., 2013; Morita et al., 2014; Mohana Kumara et al., 2016). Besides plants, endophytic fungi associated with *A. rohituka* and *D. binectariferum* have also been shown to produce rohitukine in culture (Mohana Kumara, 2012; Mohana Kumara et al., 2012; Kumara et al., 2014). But the biosynthetic pathway of chromone alkaloids in general has not been elucidated so far (Abe et al., 2005; Morita et al., 2007; Izwan Mohd Lazim et al., 2013). Earlier, desorption electrospray ionization mass spectrometry imaging (DESI-MSI) shows that rohitukine in germinating seedlings is largely restricted to the cotyledonary tissue, followed by the embryo and the seed coat (Mohana Kumara et al., 2015; Mohana Kumara et al., 2016; Varun et al., 2023). Within seedlings, rohitukine was predominantly distributed in the roots, collar region of the stem, and young leaves. In the stem and roots, rohitukine was primarily restricted to the cortex region (Mohana Kumara et al., 2016). DESI-MSI and electrospray ionization (ESI) tandem mass spectrometry (MS/MS) analysis revealed the presence of oxidized, acetylated glycosylated, and methylated derivatives of rohitukine (Mohana Kumara et al., 2015; Mohana Kumara et al., 2016). In addition to chromone alkaloids, *Dysoxylum* is also known to contain as many as 279 triterpenoids belonging to different scaffolds like dammarane, nortriterpenoid, oleanane, lupane, tirucallane, lanostane, cycloartane, glabretal and cyclopropane types (Yan et al., 2021; Naini et al., 2022). With recent advancements in sequencing technologies like genome and transcriptome sequencing of medicinal plants has become an important tool in understanding the biosynthetic pathway of metabolites of therapeutic relevance. For example, the genomes and transcriptomes of medicinal plants such as *Asparagus racemosus*, *Curcuma longa*, *Polygonum cuspidatum*, *Ocimum* spp., and *Azadirachta indica* have helped in establishing the different metabolic pathways (Narnoliya et al., 2014; Rajakani et al., 2014; Krishnan et al., 2016; Pandreka et al., 2021; Joudaki et al., 2023). These sequence resources form a base for further elucidation and functional characterization of the constituent metabolic pathways facilitating metabolic engineering in heterologous systems (Ma et al., 2021; Hu et al., 2023; Kwan et al., 2023). In the current study, we report the *de novo* transcriptome sequencing, assembly of the leaf and root tissues of *D. binectariferum* and annotation of genes in specialized metabolic pathways including chromones, alkaloids, flavonoids, sesquiterpenes and triterpene pathways. We also report differentially expressed genes in leaf and root tissues and study their tissue-specific gene expression. Finally, we identified potential genes involved in the above biosynthetic pathways and showed relative expression of their transcripts in leaves and roots.

2 Materials and methods

2.1 Plant material

D. binectariferum was identified, collected and the voucher specimen was deposited at The University of Transdisciplinary Health Sciences and Technology herbarium, Bangalore (voucher specimen number; 122951-55). During the fruiting season *D. binectariferum* seeds were collected from Jog, Central Western Ghats, India (14^0 13' 65" N and 74^0 48' 35" E). Seeds were sown in polybags and seedlings were kept under shade with continuous watering and maintained in a nursery at the University of Transdisciplinary Health Sciences and Technology, Bengaluru. The leaves and roots of one-year-old seedlings of similar age and size were used in transcriptome sequencing and metabolite analysis (Figure 1A). The sampling was non-invasive with no impact on the natural growth or regeneration of *D. binectariferum* populations in the wild. And the study was conducted following relevant national and institutional guidelines.

2.2 RNA isolation from *D. binectariferum* and differential expression analysis

Total RNA was isolated from the root and leaf tissues of one-year-old *D. binectariferum* seedlings using TRIZOL reagent (Sigma Life Science, USA) (Meng and Feldman, 2010). Each sample included three biological replicates. The quantity and quality of total RNA was determined by NanoDdrop (Thermo Scientific) and agarose gel electrophoresis. The purity of total RNA was estimated using the absorbance ratio at 260/280 and 260/230, and the RNA integrity number (RIN). Samples showing acceptable RNA integrity numbers above 7 were used in library preparation. Sequencing was done from both 5' and 3' ends on the Illumina platform (NovaSeq 6000) according to the manufacturer's instructions (Illumina Inc., San Diego, CA, USA).

About 500ng of total RNA was used in first-strand cDNA synthesis using the Takara cDNA synthesis kit according to the manufacturer's instructions. qRT PCR analysis was carried out in triplicates using SYBR Green Universal Master Mix (Takara) in 98-well optical plates

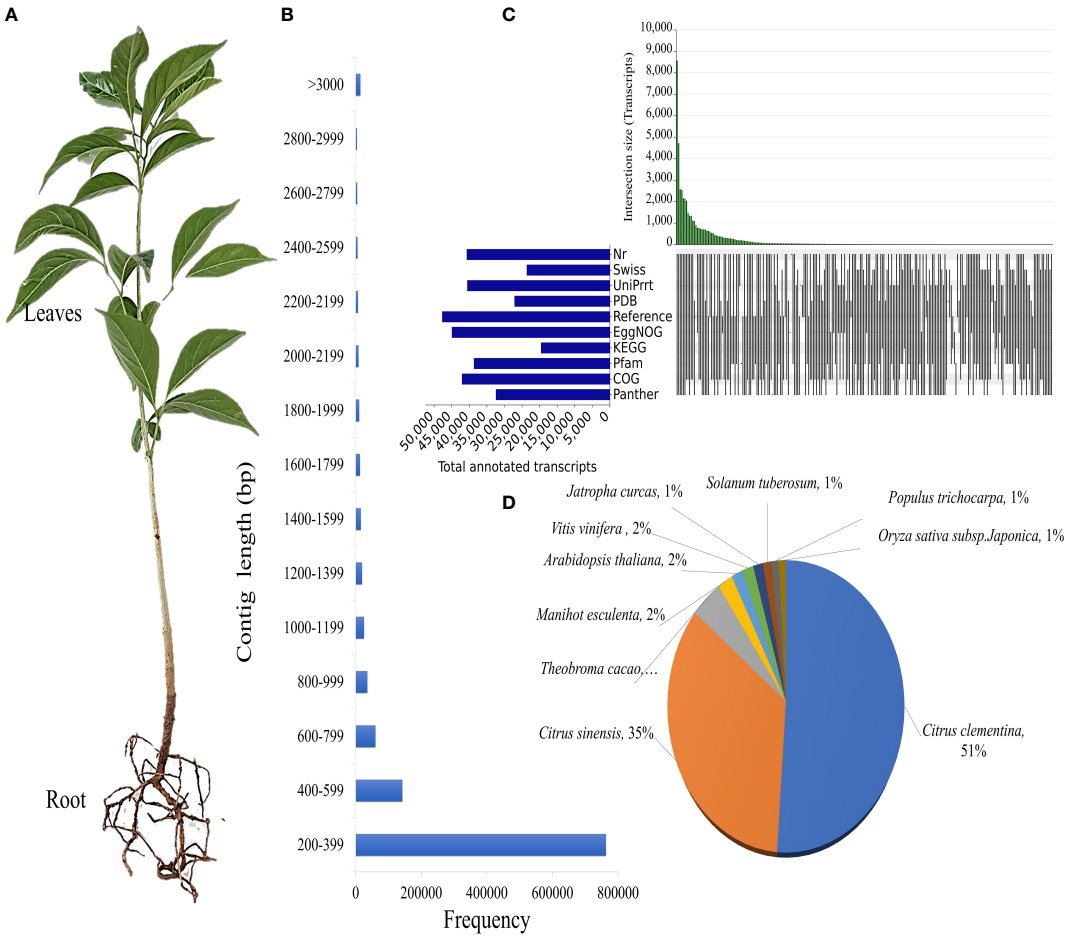


FIGURE 1
Annotation of *D. binectariferum* seedling transcriptome. **(A)** One-year-old *D. binectariferum* seedling used in transcriptome sequencing and metabolite analysis. **(B)** The contig's length distribution of *de novo* assembled transcripts. **(C)** Upset plot showing the number of assembled unigenes, annotated to Nr, Swiss, TrEMBL/UniProt, PDB, Reference sequence, EggNOG, KEGG and Pfam databases, **(D)** Distribution of top hits (species) for the non-redundant unigenes identified in the transcriptome.

using Applied Biosystems, Quantum studio 3 Real-time qPCR system. Each (10µl) reaction contained a 10ng (2µL) cDNA template, 0.4µl of 5 pM each primer, and 5µl SYBR Green mix. Cycling conditions were as follows: 1 cycle of 50°C for 2 min, 95°C for 3 min, 40 cycles of 95°C for 10 sec, 55°C for 30 sec and 1 cycle of 95°C for 15 sec, 55°C for 1 min and 95°C for 15 sec. The EF2 gene (elongation factor 2) was used as a normalization control, and all samples were analysed in triplicates, and a dissociation curve validated the specificity of each primer pair (Xu et al., 2011; Moraes et al., 2015; Linardić and Braybrook, 2021; Xu et al., 2023). Relative quantification for levels of transcripts between the samples was calculated using 2−ΔΔCT method.

2.3 *De novo* transcriptome assembly

D. binectariderum roots and leaves (three biological replicates) were sequenced. The raw data obtained was quality checked by trimming and removing adaptor sequences and other low-quality sequences using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/). The raw reads were also processed using Trimmomatic v0.38.2 to remove low-quality reads using default parameters (Bolger et al., 2014). The clean reads were assembled using Trinity Version 2.9.1(Grabherr et al., 2011) with default parameters. The level of completeness of the final transcript assembly was evaluated using BUSCO v 5.4.4 tool (Seppey et al., 2019). Coding regions of the assembled transcripts were predicted using Transdecoder Version 5.5.0 (Haas et al., 2013). We removed redundant sequences, identified non-redundant or representative protein sequences using CD-HIT version 1.2, and retained the longest sequence with a minimum sequence identity threshold of 0.9 contigs in each cluster (Fu et al., 2012). Non-redundant or representative sequences (>200 amino acids cut off) were annotated based on sequence similarity using blastp against the following databases; NCBI non-redundant (Nr), Swiss-Prot 2018, TrEMBL/UniProt, Protein database, Reference sequence database with e value 10^{-3} (NCBI BLAST+ blastp Galaxy Version 2.10.1+galaxy2; Christiam Camacho et al., 2009; Peter et al., 2015), HMMER/Protein family (Pfam) v3.3.2 (https://www.ebi.ac.uk/Tools/hmmer/), EggNOG (http://eggnog5.embl.de/#/app/home), Clusters of Orthologous Groups of proteins (COGs), Gene Ontology (GO) (http://www.pantherdb.org/), and Kyoto Encyclopaedia of Genes and Genomes (KEGG) (https://www.kegg.jp/) (Figure S1).

2.4 Differential expression analysis

Global differentially expressed genes (DEGs) analysis between root and leaf samples was performed using DESeq2 (ver 2.11.40.7) tool with the Benjamini-Hochberg procedure (Love et al., 2014). The expression levels were calculated and normalized using TPM methods (Lin et al., 2016; Patro et al., 2017). DEGs were identified with adjusted FDR ≤ 0.05 (false discovery rate), log2 (fold change) of > 2 and FPKM value of >11. An online enrichment tool, ShinyGO v0.75, was used to identify different KEGG pathways enriched in DEGs (http://bioinformatics.sdstate.edu/go75/). The significance of KEGG terms was determined by a p-value or q-value of ≤0.05. Further investigation was performed with selected DEGs. Within this list, we concentrated on genes associated with 17 different metabolic pathways comprising terpenoids, flavonoids, piperidine, and chromone alkaloids etc.

2.5 Extraction of rohitukine and quantification

Metabolites were extracted from the leaf and root tissues of *D. binectariferum* using methods described earlier (Mohanakumara et al., 2010). Briefly, the freeze-dried samples were ground to a fine powder. Extraction was carried out using methanol (10 mL). The extracts were vortexed, sonicated (30 min) and centrifuged (8,000 rpm for 10 min). Next, extracts were passed through membrane filters (0.2 µm) and kept in airtight vials at -20°C until further use. Samples were analyzed using reverse-phase HPLC (Shimadzu, LC20AT, Japan), RP-18 column (4.6 x 250 mm, 5µm) with UV absorbance at 254 nm. The standard rohitukine was prepared with a series of concentrations (0.2 – 1.0 mg/ml) using liquid chromatography-mass spectrometry (LC-MS) grade methanol and filtered using 0.2µ syringe filters. Acetonitrile and 0.1% TFA were used in gradient mode as the mobile phase. The linear graph obtained (y = mx) was used in quantification of rohitukine in samples (R^2 = 0.99) (Mohana Kumara et al., 2016). The significance of rohitukine content in the leaf and root of *D. binectariferum* was tested using t-tests (unpaired), F-tests, and Kruskal-Wallis tests using Past 4.11 (Hammer et al., 2001).

2.6 GC-MS analysis

Leaf and root tissues of *D. binectariferum* were also subjected to volatile analysis using GC-MS (Sharma et al., 2021). Leaf and root samples were dried at 40 °C for 8-12 hrs in a hot air oven and 1.0 ± 0.01 g of fine powder was subjected to headspace analysis using GC-MS fitted with RTx-volatiles capillary column (3.0 m × 0.25 mm × 0.25 µm). The analysis was done using a Shimadzu® - Nexis GCMS 2030 coupled to a mass spectrometer with a triple quadrupole TQ8040NX, equipped with an HS-20 auto-sampler (Shimadzu, Tokyo, Japan). The following GC temperature program was used; the column oven temperature was maintained at 80°C for 1 min., followed by two heating ramps of 5 and 10°C/min. until reaching temperatures of 150 °C and 200 °C, respectively. Mass spectra were obtained using electron impact at 70 eV and a start and end mass-to-charge ratio (*m/z*) of 30 and 500, respectively. The compounds were identified by comparison to the mass spectra from library databases (NIST 98; http://www.nist.gov) and by calculating Kovat's indices using alkane standards (C8-C24) RT values.

3 Results and discussion

3.1 *De novo* assembly

The *D. binectariferum* transcriptome generated 42.43 and 38.74 million paired-end short reads (150 bp) for leaves and roots,

respectively. Filtering for quality resulted in 41.52 (97.85%) and 38.09 (98.33%) million clean reads for the leaf and roots, respectively (Table 1). *De novo* assembly of the short reads generated 274,970 contigs and 126,788 unigenes from the whole transcriptome with an N50 length of 1,560 bp (Table 2). The average GC content of the contigs derived from the transcriptome was 42.6% (Table S1). Of the 2,326 BUSCOs in the Eudicots dataset, 2,142 (92.1%) complete BUSCOs were detected in the assembly (Table S2). The results indicated that the assembly was almost complete with an adequate representation of the gene directory. From the assembly, 67.67% of the contigs (763,242) were 200-400 bp in length, 21.0% (237,930) were 400-1,000 bp, 7.43% (83,771) were 1,000-2,000 bp, 2.55% (28,784) were 2,000-3,000 bp, and only 1.26% (14,233) exceeded 3,000 bp (Figure 1B). Using the CD-HIT tool (>200 amino acids cut off), we identified 51,598 nonredundant protein sequences from a total of 117,619 translated unigenes/protein sequences. After removing the redundancy, *De novo* assembly generated short reads of 117,619 contigs and 51,340 unigenes from the whole transcriptome with an N50 length of 1,176bp. The average GC content of the contigs derived from the transcriptome was 46.89% (Table S1) (Table 2).

TABLE 2 Summary of *D. binectariferum* transcriptome final assembly.

Assembly	Contigs/Unigenes
# contigs	117,619
# contigs (>= 0 bp)	117,619
# contigs (>= 1000 bp)	51,340
Largest contig	15,318
Total length	132,769,470
Total length (>= 0 bp)	132,769,470
Total length (>= 1000 bp)	82,035,717
N50	1,176
N90	687
auN	1,481.2
L50	36,872
L90	96,885
GC (%)	46.89
CD-HIT PROTEIN (>200)	51,598

3.2 Functional annotation

To identify putative protein functions, all the assembled unigenes were annotated using the Basic Local Alignment Search Tool (BLAST) against ten publicly available protein databases. Out of 51,598 unigenes, 40,699 (78.88%) were annotated to Nr, 23,618 (45.77%) to Swiss-Prot, 40,552 (78.59%) to TrEMBL/UniProt, 27,101 (52.52%) to Protein database, 47,675 (92.4%) to a Reference sequence database, 44,919 (87.06%) to EggNOG, 19548 (37.89%) to KEGG, 38,297 (74.22%) to Pfam, 42,040 (81.48%) to COG and 36,530 (70.8%) to Panther GO (Figure 1C; Additional File 1). *Citrus clementina* (7,174), *Citrus sinensis* (4,837), *Theobroma cacao* (621), *Manihot esculenta* (296), and *A. thaliana* (224) contributed to the most gene annotations (Figure 1D). Two *Citrus* species were the major hits because both *Citrus* and *D. binectariferum* are related phylogenetically as well as encode shared metabolic pathways (Levitsky and Dembitsky, 2014; Bhambhani et al., 2017; Du et al., 2021; Hou et al., 2022).

TABLE 1 Quality analysis of *Dysoxylum binectariferum* transcriptome.

Samples	Replicates	Forward R1/ Reverse R2	Number of seq. Before trimming	Number of seq. After trimming	GB Before trimming	GB After trimming
DBR (root)	DBR1	R1	43,318,058	42,370,596	14.5	13.7
		R2	43,318,058	42,370,596	14.5	13.7
	DBR2	R1	43,000,082	41,776,589	14.4	13.5
		R2	43,000,082	41,776,589	14.4	13.4
	DBR3	R1	40,989,199	40,431,285	13.8	12.9
		R2	40,989,199	40,431,285	13.8	12.8
Avg.			42,435,779	41,526,156	14.2	13.3
DBL (leaf)	DBL1	R1	45,090,230	44,321,743	15.2	14
		R2	45,090,230	44,321,743	15.2	13.7
	DBL2	R1	39,422,787	38,839,563	13.2	12.3
		R2	39,422,787	38,839,563	13.2	12.2
	DBL3	R1	31,723,969	31,134,812	10.7	9.6
		R2	31,723,969	31,134,812	10.7	9.4
Avg.			38,745,662	38,098,706	13.0	11.9

Further, Nr unigenes were annotated to three major ontologies (GO): molecular function (MF), biological process (BP), and cellular component (CC). BP comprises 41.63% of the total assigned annotations, whereas CC and MF comprised 27.63% and 30.74% respectively. Among biological processes, the unigenes are predominantly annotated to the cellular process (32.6%) and metabolic process (28.1%), followed by biological regulation (7.4%). Similarly, for the CC category, the largest number of unigenes were assigned to cellular anatomical entities (45.2%) and protein-containing complexes (7.4%). While, in the MF category, the catalytic activity (31.1%), and binding (15.6%) were the most annotated (Figure 2A). A total of 42,040 transcripts were assigned to 25 COG classifications with the majority in the category "function unknown" (10,854, 21.04%), followed by "post-translational modification, protein turnover, chaperone functions" (3,351, 6.49%), "transcription" (2,928, 5.67%) and "carbohydrate metabolism and transport" (2,237, 4.34%) (Figure 2B).

The KEGG classifications for the assembled unigenes were used to evaluate the completeness of the transcriptome library as well as the effectiveness of the annotation process for identifying the specialized metabolic pathways. A total of 19,548 assembled unigenes were assigned into six main functional categories (Metabolism, Genetic Information Processing, Environmental Information Processing, Cellular Processes, Organismal Systems, and Human Diseases) and 46 subcategories (Figure 3A) and 431 KEGG pathways. The two most abundant sub-categories were "metabolism" and "human diseases", covering 57.47% and 23.14% of the total annotations, respectively. The rest were covered by the remaining categories of Genetic Information Processing (7.28%), Environmental Information Processing (6.22%), Cellular Processes (6.39%) and Organismal Systems (10.13%). Furthermore, the unigenes coding for specialized metabolite biosynthesis were analyzed. The 17 major specialized metabolic pathways were selected and their respective KO and unigene counts are shown in Figure 3B and Additional File 2). Of these, 213 unigenes were assigned to "Phenylpropanoid biosynthesis", followed by 124 unigenes for "Terpenoid backbone biosynthesis", 102 for steroid biosynthesis, and others. These annotations form a basis for the functional characterization of genes involved in the specialized metabolism and regulation of *D. binectariferum* (Kanehisa and Goto, 2000; Liu et al., 2013; Bhambhani et al., 2017).

3.3 Metabolic pathway analysis

We identified 6,495 DEGs in total, including 3,532 genes that were upregulated in roots and 2,963 genes that were downregulated in leaves. Further, unigenes related to 17 different specialized metabolic pathways were analyzed for their expression levels (log_2fold) in roots and leaves. Of the 736 DEGs that were found to be involved in specialized metabolism, 284 of them were upregulated and 452 of them were downregulated in the root compared to the leaf (Additional file 3 and 4). Based on KEGG pathway enrichment of the bubble diagram, carbon metabolism, stilbenoid, flavonoid, unsaturated fatty acids, and phenylpropanoid biosynthesis were the most dominant pathways, and the majority of KEGG-identified genes were associated with metabolic pathways and secondary metabolite biosynthesis (Figure 3C).

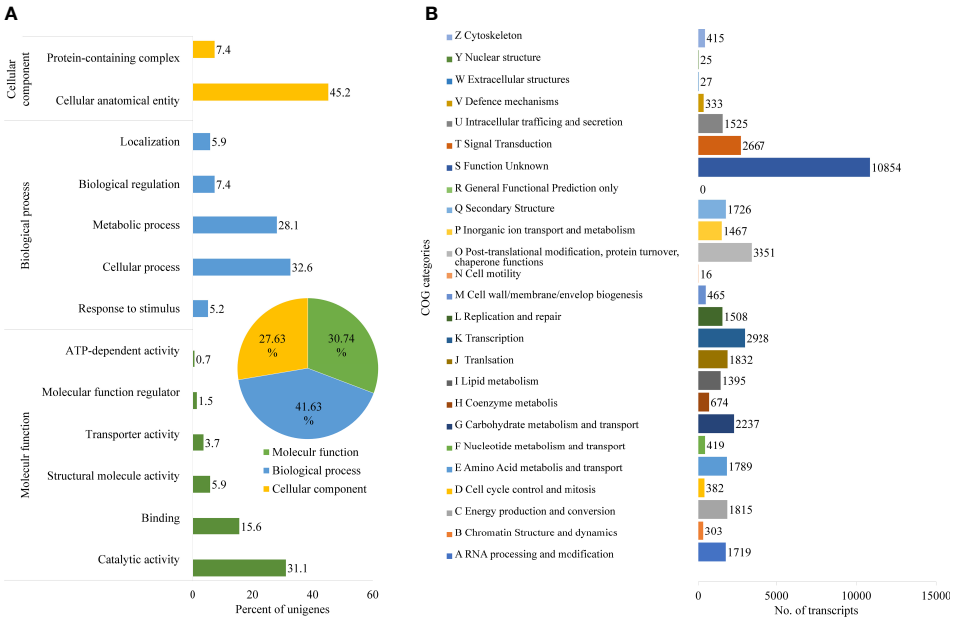


FIGURE 2
Functional annotation of *D. binectariferum* transcriptome based on Gene Ontology (GO). **(A)** GO functional classifications of assembled *D. binectariferum* unigenes. Insert in pie chart showing percentage of annotation to three different classes of gene ontology. **(B)** Clusters of Orthologous Groups (COG) functional classifications of assembled *D. binectariferum* unigenes and associated number of transcripts with COG function categories.

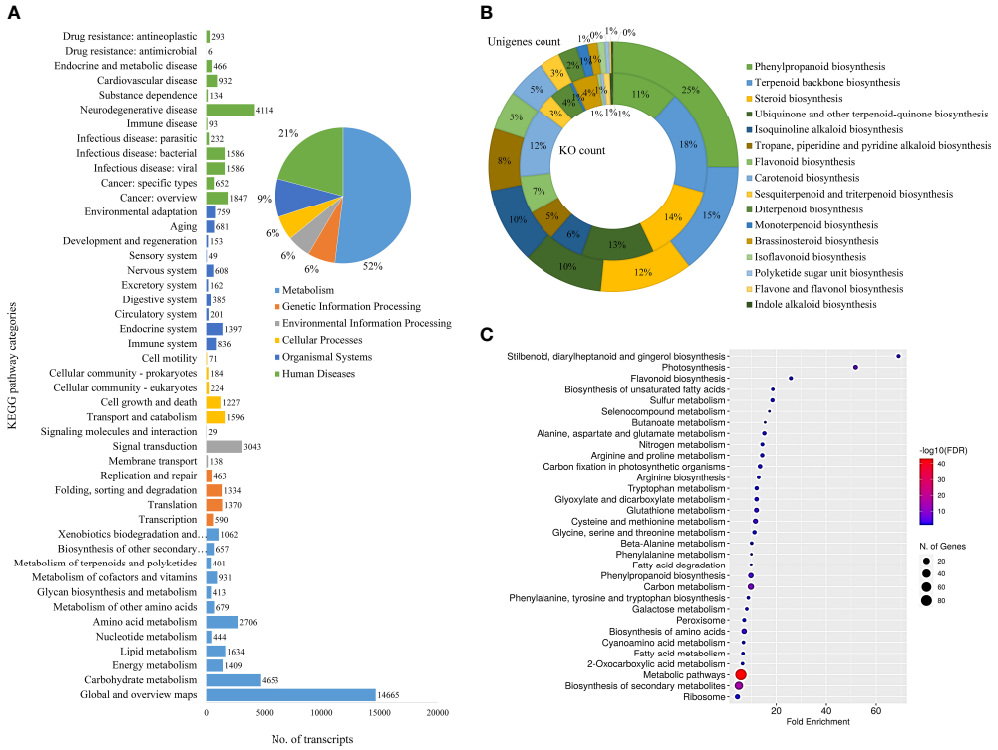


FIGURE 3
Functional annotation of *D. binectariferum* transcriptome based on KEGG. **(A)** Assembled *D. binectariferum* unigenes annotated to the functional classification of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and their associated number of transcripts are shown for each of the KEGG functional category. **(B)** Pie chart showing the selected KEGG pathways related to specialized metabolic pathways and the inner ring represents the KO category while the outer ring represents its respective unigenes count, **(C)** KEGG pathway enrichment from final annotated transcripts of *D. binectariferum*. The vertical axis represents the pathway's name, and the horizontal axis represents the fold enrichment. The size and color of bubbles indicate the number and degree of enrichment of different metabolites, respectively.

3.3.1 Identification of unigenes involved in the terpenoid pathway

Terpenoids comprise the largest group of structurally diverse natural compounds and are known to be biosynthesized *via* two biosynthetic routes; the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway and the mevalonic acid (MVA) pathway (Sandeep and Ghosh, 2020). The isoprene unit (C5) is synthesized from the MEP pathway and is engaged in the formation of mono-(C10), Di-(C20) and other terpenoids. Whereas the isoprene unit from the MVA pathway is used in the synthesis of triterpene (C30) and sesquiterpenes (C15) (Schilmiller et al., 2009; Zhao et al., 2013; Ghissing et al., 2023). In the *D. binectariferum* transcriptome, around 269 unigenes (70 key enzymes) were found to be associated with the terpenoid pathways (Figure 4; Figure S2). Of these, we identified 48 unigenes encoding 6 key enzymes in the mevalonate pathway (MVA) and 29 unigenes encoding 8 key enzymes in the MEP pathway leading to the formation of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (Kuzuyama, 2002). IPP and DMAPP go through a series of condensation reactions by prenyl diphosphate synthases to form prenyl diphosphates like geranyl diphosphate (GPP; C10), farnesyl diphosphate (FPP; C15), geranylgeranyl diphosphate (GGPP; C20) and other diphosphates. These prenyl diphosphate precursors form a basis for the biosynthesis of a diverse classes of terpenoids currently observed in plants (Figure 4; Figure S2) (Lorigooini et al., 2020). The formation of FPP from IPP by the enzyme short-chain Z-isoprenyl diphosphate synthase (K12503, 1 unigene) leads to triterpenoid synthesis *via* squalene. The squalene is oxidized to form 2, 3 oxidosqualene by squalene monooxygenase (SQLE, 9 unigenes) and 2, 3 oxidosqualene is the central precursor in biosynthesis of diverse triterpenoids including dammarane, cabraleadiol, 3-epiocotillol, methyl shoreate, and others (Yan et al., 2014; Fan et al., 2021). The 2, 3 oxidosqualene also leads to the formation of sterols (phyto- and ergosterols) involved in sitosterol and phytosterols biosynthesis in plants. Farnesyl diphosphate synthase (FDPS, 4 unigenes) accepts both DMAPP and IPP from MEP and MVA pathways in the mono and sesquiterpenoids (Szkopińska and Płochocka, 2005). We analyzed the volatile components of the leaf and root tissues of *D. binectariferum* using GC-MS to identify the possible precursors of the terpenoids. The leaf contained 14% monoterpenes, 6.5% sesquiterpenes, 1.5% ketones, and 1.27% cycloalkanes, while the root contained 49.75% sesquiterpenes, 8.86% sesquiterpenoids, 0.2% aldehydes, and 0.11% ketones (Table S3). The farnesyl diphosphate synthase (FDPS, 4 unigene) is involved in the synthesis of the sesquiterpenoids. The enzyme (-)-germacrene D synthase (GERD, 6 unigenes) makes germacren-type sesquiterpenes. In GC-MS we detected several sesquiterpenenes including α- and β-cadinenes, α- and β-copaene α-ylangene, caryophyllene, alloaromadendrene, α-guaiene, and germacrene-D in leaves and α-muurolene in the root. Additionally, monoterpene (+)-4-carene was also detected in *D. binectariferum*

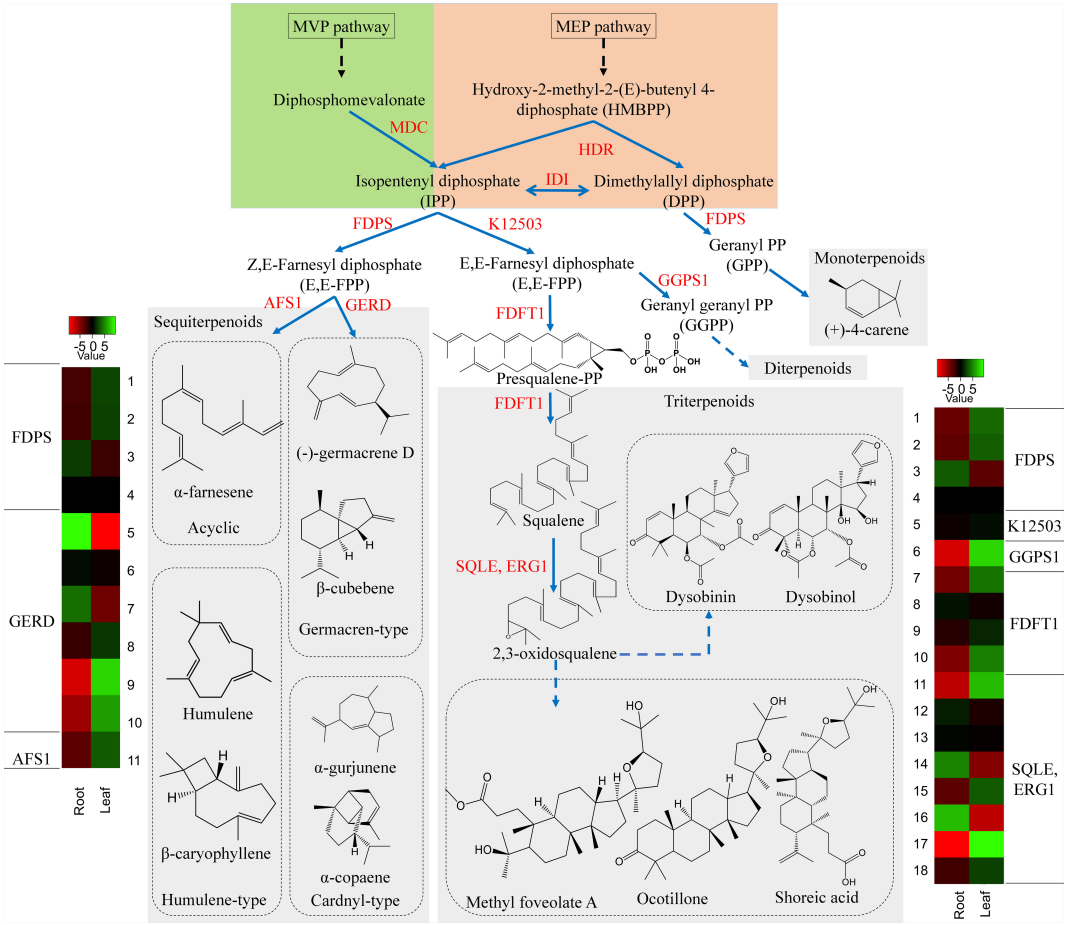


FIGURE 4
D. binectariferum candidates involved in the biosynthetic pathways of terpenoids. Key biosynthetic enzymes (red)) identified from the *D. binectariferum* transcriptome are shown at each step. Heat-map displaying differential expression of unigene transcripts (enzymes) of *D. binectariferum* transcriptome are shown for roots and leaves. Details of numbers labelled in the Heat-map are provided in the additional file 3. MDC-Mevalonate pyrophosphate decarboxylase; HDR- 4-Hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR)/HMBPP reductase; IDI-Isopentenyl diphosphate isomerase; FDPS- Farnesyl diphosphate synthase; FDPS-Farnesyl diphosphate synthase; K12503: short-chain Z-isoprenyl diphosphate synthase; AFS 1-α-farnesene synthase; GERD- (-)-Germacrene D synthase; GGPS1-Geranylgeranyl pyrophosphate synthase1; SQLE/ERG1- Squalene epoxidase/Squalene monooxygenase.

leaves (Table S3; Figure S3). All these sesquiterpenes detected in GC-MS analysis were mapped to different terpenoid biosynthetic routes (Figure 4; Figure S3). The transcriptome showed 12 unigenes involved in monoterpenoid biosynthesis and 21 unigenes encoding 7 enzymes in gibberellin biosynthesis (2 beta-dioxygenase GA2ox, 7 unigenes, ent-kaurenoic acid monooxygenase KAO, 4 unigenes) (Figures 3B, 4). In total, the MEP pathway predominates in the roots and MVA pathway in the leaf which contributes to the generation of many volatile molecules (Inaba and Ito-Inaba, 2010; Tomar et al., 2013; Pérez et al., 2022) and they play key roles in the biosynthesis of various defensive compounds (Yan et al., 2021; Naini et al., 2022; Ghissing et al., 2023).

3.3.2 Identification of unigenes involved in the flavonoid pathway

In flavonoid biosynthesis, twelve key enzymes are involved in the conversion of *p*-coumaroyl CoA to naringenin, and *D. binectariferum* revealed 44 unigenes to be associated with all the 12 key enzymes (Figure 3B). Chalcone synthase (*CHS*, 6 unigenes) is the first key enzyme in flavonoid biosynthesis that converts 4-coumaroyl CoA to naringenin chalcone. Then the isomerization of naringenin chalcone to noreugenin is catalyzed by the enzyme chalcone isomerase (*CHI*, 2 unigenes). Noreugenin forms a central precursor from which all other flavonoids are derived (Liu et al., 2013). The naringenin is converted to dihydrokaempferol by an enzyme naringenin 3-dioxygenase (*F3H*, 1 unigene) and dihydrokaempfero is further converted to leucopelargonidin by a bifunctional enzyme dihydroflavonol 4-reductase/flavanone 4-reductase (*DFR*, 3 unigenes). Leucopelargonidin is a key molecule, where (+)-Afzelechin and pelargonidin are formed by the enzymes leucoanthocyanidin reductase (*LAR*, 4 unigenes) and anthocyanidin synthase (*ANS*, 1 unigene) respectively, and also converted back to the dihydrokaempferol by *ANS* (Zhao et al., 2017; Wang et al., 2021). Further, pelargonidin is reduced to form (-)-Epiafzelechin by an enzyme anthocyanidin reductase (*ANR*, 1 unigene) (Figure 5). The genes that encode these enzymes, *PAL* (3 unigenes), *C4H* (5 unigene),

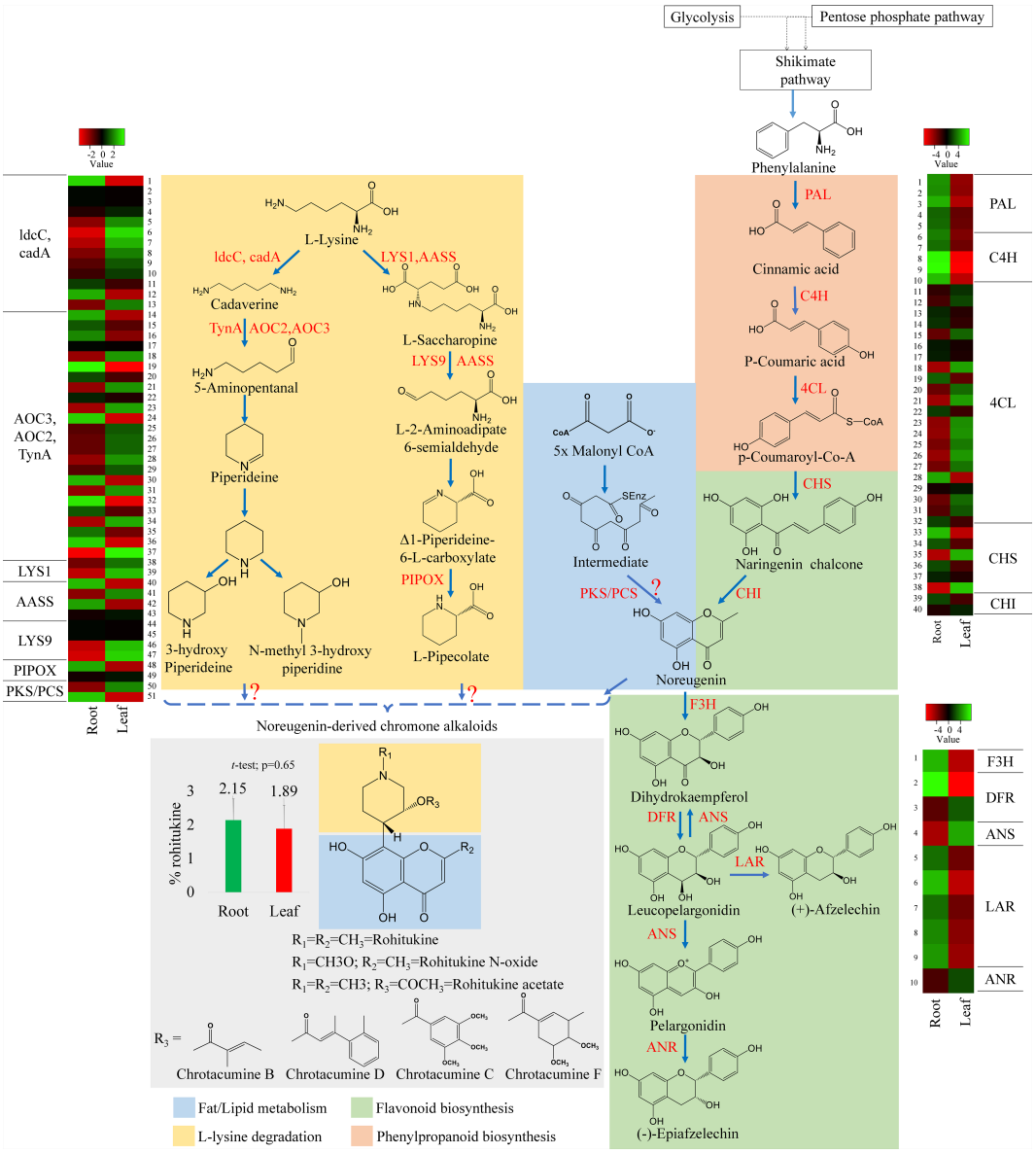


FIGURE 5
Putative biosynthetic pathway of rohitukine. Convergence of piperidine (also called lysine degradation) and shikimate routes (highlighted with different colours) in biosynthesis of chromone alkaloids in *D. binectariferum*. (rohitukine). Key biosynthetic enzymes (red) identified from the *D. binectariferum* transcriptome are shown at each. Steps highlighted with question marks denote unknown enzymatic conversions or candidate enzymes. Dotted arrows represent multiple steps or enzymes. Heat-map displaying differential expression of unigene transcripts (enzymes) *D. binectariferum* transcriptome are shown for roots and leaves. Details of numbers labelled in the Heat-map are provided in the additional file 3. The graph (insert) showing rohitukine content in the root and leaves of a one-year-old *D. binectariferum* seedling that was used in transcriptome sequencing. Mean (percent rohitukine, % of dry weight) and the error bars derived from three independent biological replicates (n =3). Piperidine (Lysine degradation) pathway; ldcC-Lysine decarboxylase; cadA- Lysine decarboxylase; LYS1- Lysine-forming saccharopine dehydrogenase; AASS- Alpha-aminoadipic semialdehyde synthase; AOC2/AOC3/TynA- primary-amine oxidase; LYS9- L-glutamate-forming saccharopine dehydrogenase 9; PIPOX- Sarcosine oxidase/L-pipecolate oxidase. Shikimate/Flavonoid/chromone alkaloid pathways: PAL- Phenylalanine ammonia-lyase; C4H- Trans-cinnamate 4-hydroxylase;4CL-4-coumarate-CoA ligase; PKS/PCS, Type III Polyketide synthase/Pentaketide chromone synthase; CHI, Chalcone isomerase; F3H, flavanone 3-hydroxylase; DFR, Dihydroflavonol 4-reductase/flavanone 4-reductase; LAR, Leucoanthocyanidin reductase; ANS, Anthocyanidin synthase; ANR, Anthocyanidin reductase.

CHS (6 unigenes), *CHI* (2 unigene) and *4CL* (22 unigenes) were found to be significantly upregulated in roots compared to leaves (Figure S4).

In parallel, using HPLC, we also measured rohitukine content in *D. binectariferum* leaf and the root tissues and the data showed that rohitukine content was comparatively more in root (2.15 ± 0.62%) than in the leaf (1.89 ± 0.69%) (Figure 5; $p > 0.05$; not significant).

To test if the expression pattern of flavonoid and other associated pathway genes is correlated with rohitukine content we measured their expression patterns using quantitative Realtime PCR (qRT) in leaves and roots. The upstream genes *PAL, C4H, 4CL*, and *CHI* are involved in the biosynthesis of the key precursor naringenin (Table S4; Figure S4) and all these four genes were highly expressed in the

roots compared to the leaves and this was comparable to DEseq-RNA-seq expression data (Figure S4). Further, using BLASTP with *Arabidopsis CHS* as a query sequence, we identified *CHS* like genes from *D. binectariferum* transcriptome. In total four full-length *CHS*-like genes were identified. Of these, one of them corresponds to the *CHS* (DN149243) and it is likely to be involved in the biosynthesis of noreugenin. The other three unigenes DN1192 (*PKS1*), DN4064 (*PKS2*), and DN567668 belong to type-III polyketide synthases (*PKS-III)* (Figure 6A; Table S5). We subjected these candidates for phylogenetic analysis along with other functionally characterized CHSs and PKS-IIIs from plants together with bacterial PKSs as out groups. These genes were grouped into three clusters; chalcone synthases (*CHS*), plant non *CHS/PKS-III*, and bacterial *PKS* (Figure 6; Table S5). *D. binectariferum PKS-III* candidates DN1192 (PKS1), DN4064 (PKS2), and DN567668 were clustered with known plant *PKS-III's* and it is likely that one of these could be involved in the biosynthesis of chromone alkaloids (Figure 6). The candidates *PKS1* and *PKS2* are highly expressed in the root with low to negligible expression in the leaves of *D. binectariferum* (Figure 6B) corresponding roughly with the rohitukine content in roots. For example, a pentaketide chromone synthase (*PCS*) that makes noreugenin (5,7-dihydroxy-2-methylchromone) by successive condensation of five malonyl-CoA precursor units is known from the plant *Aloe arborescens* (Izwan Mohd Lazim et al., 2013). Therefore, functional characterization of the *PKS-III* like candidates from *D. binectariferum* likely reveal *PCS* like enzyme in the biosynthesis of rohitukine or chromone alkaloids.

3.3.3 Identification of unigenes involved in the piperidine pathway

The L-lysine degradation pathway seem to provide N-containing phenol ring in chromone alkaloids. The amino acid L-lysine is degraded by a known enzyme lysine decarboxylase (*ldcC, cadA*, 13 unigenes) leading to the formation of cadaverine (Tomar et al., 2013). Next, primary-amine oxidase (*AOC3/AOC2/tynA*, 24 unigenes) converts cadaverine to 5-aminopentanal. Cadaverine and 5-aminopentanal are the central precursors in the biosynthesis of L-pipecolate and piperidine (Reimer et al., 2017) (Figure 5). The following enzymes are known to be involved in pipecolate and peridine biosynthesis from cadaverine; L-lysine-forming saccharopine dehydrogenase (*LYS1*, 2 unigenes), alpha-aminoadipic semialdehyde synthase (*AASS*, 5 unigenes), alpha-aminoadipic semialdehyde synthase, saccharopine dehydrogenase (*LYS9*, 3 unigenes) and sarcosine oxidase/L-pipecolate oxidase (*PIPOX*, 2 unigenes) (Figure 5) (Hartmann and Zeier, 2018). Presence of these candidates/unigenes and their expression in leaves suggest that the L-lysine degradation pathway is likely operative in *D. binectariferum* (Figures 5, 6B). In plants with L-lysine degradation pathway the associated pathway genes are normally expressed in leaves and the pathway end products like pyridine and piperidine alkaloids are also present in leaves (defence

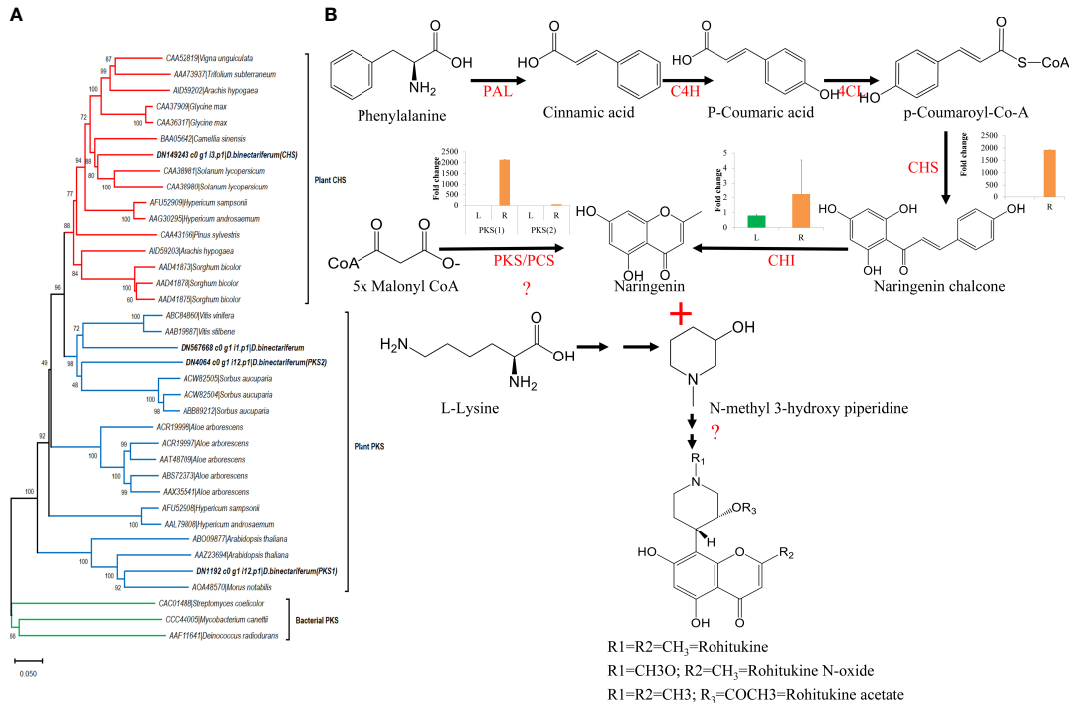


FIGURE 6
Candidate polyketide synthases identified from *D. binectariferum* transcriptome and their likely role in rohitukine/chromone alkaloids biosynthesis. **(A)** Phylogenetic tree of functionally characterized chalcone synthase (CHS) and polyketide synthase IIIs (PKSIII) from diverse plants along with CHS and PKS-III candidates identified from *D. binectariferum* transcriptome. Tree was created using MEGA7 with 1000 bootstrap values. Candidates highlighted in bold represent the candidates from *D. binectariferum*. **(B)** Predicted biosynthetic pathway of rohitukine/chromone alkaloids in *D. binectariferum*. Graphs in the inserts represent quantitative expression of key genes in the chromone alkaloid biosynthetic pathways (qRT mean ± SD, n=3, L; Leaves, R; Roots). PAL, Phenylalanine ammonia-lyase; C4H, Trans-cinnamate 4-hydroxylase; 4CL, 4-coumarate-CoA ligase; PKS/PCS, Type III Polyketide synthase/Pentaketide chromone synthase; CHI, Chalcone isomerase.

related) (Inaba and Ito-Inaba, 2010; Khadem and Marles, 2012; Hartmann and Zeier, 2018).

3.3.4 Putative chromone alkaloid biosynthetic pathway

Rohitukine is a chromone alkaloid consisting of noreugenin or flavone scaffold attached to a nitrogen containing piperidine ring (Harmon et al., 1979; Mohanakumara et al., 2010). Noreugenin chromone scaffold is a central precursor in biosynthesis of diverse chromone alkaloids including rohitukine and its derivatives. Noreugenin is made either through; a) a flavonoid pathway or b) through successive condensation of multiple malonyl co-A units by type-III polyketide synthase-like enzymes. The presence of the unigenes coding for PKS-III candidate enzymes as well as their high expression in roots where rohitukine is highly accumulated suggests that the route 'b' is more plausible (Figure 6B). And the piperidine ring is likely derived from the L-lysine degradation pathway and condensation of piperidine moiety and noreugenin yields rohitukine or chromone alkaloids (Figures 5, 6). The results also highlight the convergence of multiple biosynthetic pathways including the shikimic acid/phenylpropanoid pathway, flavonoids, acetate to pentaketide pathway, and L-lysine degradation pathway in the biosynthesis of complex chromone alkaloids like rohitukine. These results form a base for the further comprehensive investigation of the chromone alkaloid biosynthesis that is required for engineering heterologous hosts to make these valuable molecules and their derivatives.

4 Conclusions

D. binectariferum, an endemic medicinal plant of the Western Ghats, India, is well known to produce a chromone alkaloid called rohitukine and as well as a variety of triterpenoids and flavonoids. Rohitukine is a natural precursor for the semi-synthetic of anticancer drugs flavopiridol and P-276-00. To understand the biosynthetic pathway of rohitukine, we generated a comprehensive transcriptome assembly of leaf and root tissues and identified 51,598 non-redundant protein sequences of more than 200 amino acids. About 78.95% of these unigenes were annotated to the Nr database highlighting the completeness of the assembly. Next, with a combination of metabolite profiling and transcriptome assembly, we presented a biosynthetic route to these diverse compounds including terpenoids, flavonoids, and chromone alkaloids. More specifically, we discover candidate genes in rohitukine biosynthesis, and these enzymes strongly suggest the possibility of involvement of noreugenin pathway in the production of rohitukine and these biosynthetic routes have not been described previously. Therefore, these results pave the way for further functional characterization of these genes and clarify the biosynthesis pathway of chromone alkaloids, specifically rohitukine in *D. binectariferum*.

Data availability statement

The data presented in the study are deposited found in the NCBI repository, BioProject-PRJNA943416, and SRA accession numbers, SRR23901936 and SRR23901937.

Author contributions

PK and RT designed the research. PK and EV carried out experiments and analyzed the data. EV, JS and AM participated in sample collection and carried out the qRT-PCR assays. PK, RT and EV wrote the manuscript. The authors read and approved the final manuscript.

Funding

The current study was financially supported by SERB Early Career Grant ECR/2017/002686.

Acknowledgments

MP thanks SERB Early Career Grant ECR/2017/002686 for the financial support. RT acknowledges DBT-Ramalingaswami fellowship. We acknowledge usegalaxy.org for the server support. We acknowledge Prof. Uma Shaanker R, UAS Bangalore for his constructive suggestions while developing 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.2023.1098987/full#supplementary-material

ADDITIONAL FILE 1
Functional annotation of final transcriptome of *D. binectariferum* with different databases.

ADDITIONAL FILE 2
Annotated (17 selected) KEGG metabolite pathways of *D. binectariferum*.

ADDITIONAL FILE 3
Differential expression of unigenes in leaves and roots of *D. binectariferum*.

ADDITIONAL FILE 4
Differential expression of unigenes (DEGs) in root and leaf tissues of *D. binectariferum* of selected 17 specialized metabolic pathways.

References

Abe, I., Utsumi, Y., Oguro, S., Morita, H., Sano, Y., and Noguchi, H. (2005). A plant type III polyketide synthase that produces pentaketide chromone. *J. Am. Chem. Soc.* 127, 1362–1363. doi: 10.1021/ja0431206

Albert, T. K., Rigault, C., Eickhoff, J., Baumgart, K., Antrecht, C., Klebl, B., et al. (2014). Characterization of molecular and cellular functions of the cyclin-dependent kinase CDK9 using a novel specific inhibitor. *Br. J. Pharmacol.* 171, 55–68. doi: 10.1111/bph.12408

Bhambhani, S., Lakhwani, D., Gupta, P., Pandey, A., Dhar, Y. V., Kumar Bag, S., et al. (2017). Transcriptome and metabolite analyses in Azadirachta indica: identification of genes involved in biosynthesis of bioactive triterpenoids. *Sci. Rep.* 7, 5043. doi: 10.1038/s41598-017-05291-3

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

Borowczak, J., Szczerbowski, K., Ahmadi, N., and Szylberg, Ł. (2022). CDK9 inhibitors in multiple myeloma: a review of progress and perspectives. *Med. Oncol.* 39, 39. doi: 10.1007/s12032-021-01636-1

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 15 (10), 421. doi: 10.1186/1471-2105-10-421

Chen, K. T. J., Militao, G. G. C., Anantha, M., Witzigmann, D., Leung, A. W. Y., and Bally, M. B. (2021). Development and characterization of a novel flavopiridol formulation for treatment of acute myeloid leukemia. *J. Control Release* 333, 246–257. doi: 10.1016/j.jconrel.2021.03.042

Christian, B. A., Grever, M. R., Byrd, J. C., and Lin, T. S. (2009). Flavopiridol in chronic lymphocytic leukemia: a concise review. *Clin. Lymphoma Myeloma* 9 (Suppl 3), S179–S185. doi: 10.3816/CLM.2009.s.009

Du, Y., Song, W., Yin, Z., Wu, S., Liu, J., Wang, N., et al. (2021). Genomic analysis based on chromosome-level genome assembly reveals an expansion of terpene biosynthesis of Azadirachta indica. *bioRxiv* 13, 853861. doi: 10.1101/2021.11.11.468207

Fan, W., Fan, L., Wang, Z., and Yang, L. (2021). Limonoids from the genus melia (Meliaceae): phytochemistry, synthesis, bioactivities, pharmacokinetics, and toxicology. *Front. Pharmacol.* 12. doi: 10.3389/fphar.2021.795565

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Ghissing, U., Kutty, N. N., Bimolata, W., Samanta, T., and Mitra, A. (2023). Comparative transcriptome analysis reveals an insight into the candidate genes involved in anthocyanin and scent volatiles biosynthesis in colour changing flowers of *Combretum indicum*. *Plant Biol.* 25 (1), 85–95. doi: 10.1111/plb.13481

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

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

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4, 9.

Harmon, A. D., Weiss, U., and Silverton, J. v (1979). The structure of rohitukine, the main alkaloid of Amoora rohituka (Syn. Aphanamixis polystachya) (meliaceae). *Tetrahedron Lett.* 20, 721–724. doi: 10.1016/S0040-4039(01)93556-7

Hartmann, M., and Zeier, J. (2018). l-lysine metabolism to N-hydroxypipecolic acid: an integral immune-activating pathway in plants. *Plant J.* 96, 5–21. doi: 10.1111/tpj.14037

Hou, Q., Li, S., Shang, C., Wen, Z., Cai, X., Hong, Y., et al. (2022). Genome-wide characterization of chalcone synthase genes in sweet cherry and functional characterization of CpCHS1 under drought stress. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.989959

Houghton, P. J. (2002). Chromatography of the chromone and flavonoid alkaloids. *J. Chromatogr A* 967, 75–84. doi: 10.1016/S0021-9673(01)01555-2

Hu, Y., Zhang, Y., Šmarda, P., Bureš, P., and Guo, Q. (2023). Transcriptome and proteome associated analysis of flavonoid metabolism in haploid Ginkgo biloba. *Int. J. Biol. Macromol.* 224, 306–318. doi: 10.1016/j.ijbiomac.2022.10.125

Inaba, T., and Ito-Inaba, Y. (2010). Versatile roles of plastids in plant growth and development. *Plant Cell Physiol.* 51, 1847–1853. doi: 10.1093/pcp/pcq147

Ismail, I. S., Nagakura, Y., Hirasawa, Y., Hosoya, T., Lazim, M. I. M., Lajis, N. H., et al. (2009). Chrotacumines A–D, chromone alkaloids from Dysoxylum acutangulum. *J. Nat. Prod* 72, 1879–1883. doi: 10.1021/np9003849

Izwan Mohd Lazim, M., Safinar Ismail, I., Shaari, K., Abd Latip, J., Ali Al-Mekhlafi, N., and Morita, H. (2013). Chrotacumines E and F, two new chromone-alkaloid analogs from Dysoxylum acutangulum (Meliaceae) leaves. *Chem. Biodivers* 10, 1589–1596. doi: 10.1002/cbdv.201200391

Jain, S. K., Bharate, S. B., and Vishwakarma, R. A. (2012). Cyclin-dependent kinase inhibition by flavoalkaloids. *Mini Rev. Med. Chem.* 12, 632–649. doi: 10.2174/138955712800626683

Joudaki, F., Ismaili, A., Sohrabi, S. S., Hosseini, S. Z., Kahrizi, D., and Ahmadi, H. (2023). Transcriptome analysis of gall oak (Quercus infectoria): *De novo* assembly, functional annotation and metabolic pathways analysis. *Genomics* 115 (2), 110588. doi: 10.1016/j.ygeno.2023.110588

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Khadem, S., and Marles, R. J. (2012). Chromone and flavonoid alkaloids: occurrence and bioactivity. *Molecules* 17, 191–206. doi: 10.3390/molecules17010191

Krishnan, N. M., Jain, P., Gupta, S., Hariharan, A. K., and Panda, B. (2016). An improved genome assembly of Azadirachta indica A. Juss. *G3: Genes Genomes Genet.* 6, 1835–1840. doi: 10.1534/g3.116.030056

Kumara, P. M., Soujanya, K. N., Ravikanth, G., Vasudeva, R., Ganeshaiah, K. N., and Shaanker, R. U. (2014). Rohitukine, a chromone alkaloid and a precursor of flavopiridol, is produced by endophytic fungi isolated from Dysoxylum binectariferum Hook.f and Amoora rohituka (Roxb).Wight & Arn. *Phytomedicine* 21, 541–546. doi: 10.1016/j.phymed.2013.09.019

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

Kwan, B. D., Seligmann, B., Nguyen, T.-D., Franke, J., and Dang, T.-T. T. (2023). Leveraging synthetic biology and metabolic engineering to overcome obstacles in plant pathway elucidation. *Curr. Opin. Plant Biol.* 71, 102330. doi: 10.1016/j.pbi.2022.102330

Levitsky, D. O., and Dembitsky, V. M. (2014). Anti-breast cancer agents derived from plants. *Nat. Prod. Bioprospect* 5, 1–16. doi: 10.1007/s13659-014-0048-9

Lin, Y., Golovnina, K., Chen, Z.-X., Lee, H. N., Negron, Y. L. S., Sultana, H., et al. (2016). Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual Drosophila melanogaster. *BMC Genomics* 17, 28. doi: 10.1186/s12864-015-2353-z

Linardić, M., and Braybrook, S. A. (2021). Identification and selection of optimal reference genes for qPCR-based gene expression analysis in Fucus distichus under various abiotic stresses. *PloS One* 16 (4), e0233249. doi: 10.1371/journal.pone.0233249

Liu, X., Lu, Y., Yuan, Y., Liu, S., Guan, C., Chen, S., et al. (2013). *De novo* transcriptome of Brassica juncea seed coat and identification of genes for the biosynthesis of flavonoids. *PloS One* 8, e71110. doi: 10.1371/journal.pone.0071110

Lorigooini, Z., Jamshidi-kia, F., and Dodman, S. (2020). "Analysis of sesquiterpenes and sesquiterpenoids," in *Recent Advances in Natural Products Analysis* (Elsevier), 289–312.

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

Łukasik, P., Baranowska-Bosiacka, I., Kulczycka, K., and Gutowska, I. (2021). Inhibitors of cyclin-dependent kinases: types and their mechanism of action. *Int. J. Mol. Sci.* 22 (6), 2806. doi: 10.3390/ijms22062806

Ma, C., Zhang, K., Zhang, X., Liu, G., Zhu, T., Che, Q., et al. (2021). Heterologous expression and metabolic engineering tools for improving terpenoids production. *Curr. Opin. Biotechnol.* 69, 281–289. doi: 10.1016/j.copbio.2021.02.008

Mandal, R., Becker, S., and Strebhardt, K. (2021). Targeting CDK9 for anti-cancer therapeutics. *Cancers (Basel)* 13 (9), 2181. doi: 10.3390/cancers13092181

Meng, L., and Feldman, L. (2010). A rapid TRIzol-based two-step method for DNA-free RNA extraction from Arabidopsis siliques and dry seeds. *Biotechnol. J.* 5, 183–186. doi: 10.1002/biot.200900211

Mohana Kumara, P. (2012). *Assessing the genetic diversity of Dysoxylum binectariferum and prospecting for anticancer metabolite rohitukine from its phylogenetically related species in the Western Ghats, India* (Doctoral dissertation, University of Agricultural Science, GKVK, Bangalore, India).

Mohanakumara, P., Sreejayan, N., Priti, V., Ramesha, B. T., Ravikanth, G., Ganeshaiah, K. N., et al. (2010). Dysoxylum binectariferum Hook.f (Meliaceae), a rich source of rohitukine. *Fitoterapia* 81, 145–148. doi: 10.1016/j.fitote.2009.08.010

Mohana Kumara, P., Srimany, A., Arunan, S., Ravikanth, G., Uma Shaanker, R., and Pradeep, T. (2016). Desorption electrospray ionization (DESI) mass spectrometric imaging of the distribution of rohitukine in the seedling of Dysoxylum binectariferum hook. F. *PloS One* 11, e0158099. doi: 10.1371/journal.pone.0158099

Mohana Kumara, P., Srimany, A., Ravikanth, G., Uma Shaanker, R., and Pradeep, T. (2015). Ambient ionization mass spectrometry imaging of rohitukine, a chromone anti-cancer alkaloid, during seed development in Dysoxylum binectariferum Hook.f (Meliaceae). *Phytochemistry* 116, 104–110. doi: 10.1016/j.phytochem.2015.02.031

Mohana Kumara, P., Zuehlke, S., Priti, V., Ramesha, B. T., Shweta, S., Ravikanth, G., et al. (2012). Fusarium proliferatum, an endophytic fungus from Dysoxylum binectariferum Hook.f, produces rohitukine, a chromane alkaloid possessing anti-cancer activity. *Antonie van Leeuwenhoek Int. J. Gen. Mol. Microbiol.* 101, 323–329. doi: 10.1007/s10482-011-9638-2

Moraes, G. P., Benitez, L. C., do Amaral, M. N., Vighi, I. L., Auler, P. A., da Maia, L. C., et al. (2015). Evaluation of reference genes for RT-qPCR studies in the leaves of rice seedlings under salt stress. *Genet. Mol. Res.* 14, 2384–2398. doi: 10.4238/2015.March.27.24

Morita, H., Kondo, S., Oguro, S., Noguchi, H., Sugio, S., Abe, I., et al. (2007). Structural Insight into Chain-Length Control and Product Specificity of Pentaketide Chromone Synthase from Aloe arborescens. *Chem. Biol.* 14, 359–369. doi: 10.1016/j.chembiol.2007.02.003

Morita, H., Nugroho, A. E., Nagakura, Y., Hirasawa, Y., Yoshida, H., Kaneda, T., et al. (2014). Chrotacumines G-J, chromone alkaloids from Dysoxylum acutangulum with osteoclast differentiation inhibitory activity. *Bioorg Med. Chem. Lett.* 24, 2437–2439. doi: 10.1016/j.bmcl.2014.04.020

Naik, R. G., Kattige, S. L., Bhat, S., Alreja, B., de Souza, N. J., and Rupp, R. H. (1988). An antiinflammatory cum immunomodulatory piperidinylbenzopyranone from Dysoxylum binectariferum: isolation, structure and total synthesis. *Tetrahedron* 44, 2081–2086. doi: 10.1016/S0040-4020(01)90352-7

Naini, A. A., Mayanti, T., and Supratman, U. (2022). Triterpenoids from Dysoxylum genus and their biological activities. *Arch. Pharm. Res.* 45, 63–89. doi: 10.1007/s12272-022-01371-9

Narnoliya, L. K., Rajakani, R., Sangwan, N. S., Gupta, V., and Sangwan, R. S. (2014). Comparative transcripts profiling of fruit mesocarp and endocarp relevant to secondary metabolism by suppression subtractive hybridization in Azadirachta indica (neem). *Mol. Biol. Rep.* 41, 3147–3162. doi: 10.1007/s11033-014-3174-x

Pandreka, A., Chaya, P. S., Kumar, A., Aarthy, T., Mulani, F. A., Jennifer, C., et al. (2021). Limonoid biosynthesis 3: Functional characterization of crucial genes involved in neem limonoid biosynthesis. *Phytochemistry* 184, 112669. doi: 10.1016/j.phytochem.2021.112669

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

Peter, J. A., Cock, J. M., Grüning, C. B, Johnson, J. E., and Soranzo, N. (2015). NCBI BLAST+ integrated into Galaxy. *GigaScience* 4, 39. doi: 10.1186/s13742-015-0080-7

Pérez, L., Alves, R., Perez-Fons, L., Albacete, A., Farré, G., Soto, E., et al. (2022). Multilevel interactions between native and ectopic isoprenoid pathways affect global metabolism in rice. *Transgenic Res.* 31, 249–268. doi: 10.1007/s11248-022-00299-6

Rajakani, R., Narnoliya, L., Sangwan, N. S., Sangwan, R. S., and Gupta, V. (2014). Subtractive transcriptomes of fruit and leaf reveal differential representation of transcripts in Azadirachta indica. *Tree Genet. Genomes* 10, 1331–1351. doi: 10.1007/s11295-014-0764-7

Reimer, L. C., Will, S. E., and Schomburg, D. (2017). The fate of lysine: Non-targeted stable isotope analysis reveals parallel ways for lysine catabolization in Phaeobacter inhibens. *PloS One* 12, e0186395. doi: 10.1371/journal.pone.0186395

Sandeep,, and Ghosh, S. (2020). "Chapter 12 - Triterpenoids: Structural diversity, biosynthetic pathway, and bioactivity," in *Bioactive Natural Products Studies in Natural Products Chemistry*. Ed. Atta-ur-Rahman, (Elsevier), 411–461. doi: 10.1016/B978-0-12-819483-6.00012-6

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. USA* 106, 10865–10870. doi: 10.1073/pnas.0904113106

Sedlacek, H., Czech, J., Naik, R., Kaur, G., Worland, P., Losiewicz, M., et al. (1996). Flavopiridol (L86 8275; NSC 649890), a new kinase inhibitor for tumor therapy. *Int. J. Oncol.* 9, 1143–1168. doi: 10.3892/ijo.9.6.1143

Sekine, C., Sugihara, T., Miyake, S., Hirai, H., Yoshida, M., Miyasaka, N., et al. (2008). Successful treatment of animal models of rheumatoid arthritis with small-molecule cyclin-dependent kinase inhibitors. *J. Immunol.* 180, 1954–1961. doi: 10.4049/jimmunol.180.3.1954

Seppey, M., Manni, M., and Zdobnov, E. M. (2019). *BUSCO: Assessing Genome Assembly and Annotation Completeness*. 227–245. doi: 10.1007/978-1-4939-9173-0_14

Sharma, A., Bhardwaj, G., and Cannoo, D. S. (2021). Antioxidant potential, GC/MS and headspace GC/MS analysis of essential oils isolated from the roots, stems and aerial parts of Nepeta leucophylla. *Biocatal Agric. Biotechnol.* 32, 101950. doi: 10.1016/j.bcab.2021.101950

Stadler, W. M., Vogelzang, N. J., Amato, R., Sosman, J., Taber, D., Liebowitz, D., et al. (2000). Flavopiridol, A novel cyclin-dependent kinase inhibitor, in metastatic renal cancer: A University of Chicago phase II consortium study. *J. Clin. Oncol.* 18, 371. doi: 10.1200/JCO.2000.18.2.371

Szkopińska, A., and Płochocka, D. (2005). Farnesyl diphosphate synthase; regulation of product specificity. *Acta Biochim. Pol.* 52, 45–55. doi: 10.18388/abp.2005_3485

Tomar, P. C., Lakra, N., and Mishra, S. N. (2013). Cadaverine: a lysine catabolite involved in plant growth and development. *Plant Signal Behav.* 8, e25850. doi: 10.4161/psb.25850

Varun, E., Bhakti, K., Aishwarya, K., Suraj, R. H., Jagadish, M. R., and Mohana Kumara, P. (2023). Rohitukine content across the geographical distribution of Dysoxylum binectariferum Hook F. and its natural derivatives are potential sources of CDK inhibitors. *Heliyon* 9 (2), e13469. doi: 10.1016/j.heliyon.2023.e13469

Wang, X., Hu, H., Wu, Z., Fan, H., Wang, G., Chai, T., et al. (2021). Tissue-specific transcriptome analyses reveal candidate genes for stilbene, flavonoid and anthraquinone biosynthesis in the medicinal plant Polygonum cuspidatum. *BMC Genomics* 22, 353. doi: 10.1186/s12864-021-07658-3

Xu, B., Liu, L., and Song, G. (2023). Functions and regulation of translation elongation factors. *Front. Mol. Biosci.* 19 (8). doi: 10.3389/fmolb.2021.816398

Xu, M., Zhang, B., Su, X., Zhang, S., and Huang, M. (2011). Reference gene selection for quantitative real-time polymerase chain reaction in Populus. *Anal. Biochem.* 408 (2), 337–339. doi: 10.1016/j.ab.2010.08.044

Yan, H., Si, H., Zhao, H., Chen, L., Yu, J., Zhao, H., et al. (2021). Four new cycloartane triterpenoids from the leaves of Dysoxylum binectariferum. *Phytochem. Lett.* 41, 101–105. doi: 10.1016/j.phytol.2020.11.013

Yan, H.-J., Wang, J.-S., and Kong, L.-Y. (2014). Cytotoxic steroids from the leaves of Dysoxylum binectariferum. *Steroids* 86, 26–31. doi: 10.1016/j.steroids.2014.04.014

Zhao, L., Chang, W., Xiao, Y., Liu, H., and Liu, P. (2013). Methylerythritol phosphate pathway of isoprenoid biosynthesis. *Annu. Rev. Biochem.* 82, 497–530. doi: 10.1146/annurev-biochem-052010-100934

Zhao, H., Ren, L., Fan, X., Tang, K., and Li, B. (2017). Identification of putative flavonoid-biosynthetic genes through transcriptome analysis of Taihe Toona sinensis bud. *Acta Physiol. Plant* 39, 1–13. doi: 10.1007/s11738-017-2422-9

10.3389/fpls.2023.1098987

Glossary

CDK	Cyclin-dependent kinase
CLL	Chronic Lymphocytic Leukemia
DESI MS	Desorption Electrospray Ionization Mass Spectrometry Imaging
ESI	Electrospray Ionization
LC	Liquid Chromatography
HPLC	High-performance liquid chromatography
RP-18	Reversed Phase –
MeOH	Methanol
GC-MS	Gas chromatography- Mass Spectrometry
NIST	National Institute of Standards and Technology
RIN	RNA integrity number
NCBI	National Center for Biotechnology Information
Nr	Non-redundant
Pfam	Protein family
GO	Gene Ontology
COGs	Clusters of Orthologous Groups of proteins
KEGG	Kyoto Encyclopaedia of Genes and Genomes
DEGs	Differentially expressed genes
TFs	Transcription factors
DEGs	Differentially expressed genes
FDR	False-discovery rate
FPKM	Fragments per kilobase per million mapped
KO	KEGG Orthology
EF2	Elongation factor 2
MVA	Mevalonic acid
MEP	2-C-methyl-D-erythritol 4-phosphate
IPP	Isopentenyl diphosphate
DPP	Dimethylallyl diphosphate
GPP	Geranyl diphosphate
FPP	Farnesyl diphosphate
HDR	HDR-4-Hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR)/ HMBPP reductase
IDI	IPP isomerase
K12503	short-chain Z-isoprenyl diphosphate synthase
AFS 1	α-farnesene synthase
GGPP	Geranylgeranyl diphosphate
SQLE	Squalene epoxidase
ERG1-	Squalene monooxygenase
FDPS	Farnesyl diphosphate synthase
GERD	(-)-Germacrene D synthase
GA2ox	Gibberellin 2 beta-dioxygenase
KAO	Ent-kaurenoic acid monooxygenase
PAL	Phenylalanine ammonia-lyase
C4H	Trans-cinnamate 4-hydroxylase
4CL	4-coumarate-CoA ligase
CHS	Chalcone synthase
CHI	Chalcone isomerase
DFR	Dihydroflavonol 4-reductase/flavanone 4-reductase
F3H	Flavanone 3-hydroxylase
LAR	Leucoanthocyanidin reductase
ANS	Anthocyanidin synthase
ANR	Anthocyanidin reductase
ldc-C	Lysine decarboxylase
PatA	Putrescine aminotransferase
LYS1 L	Lysine-forming saccharopine dehydrogenase
AASS	Alpha-aminoadipic semialdehyde synthase
LYS9	L-glutamate-forming saccharopine dehydrogenase
PIPOX	Sarcosine oxidase/L-pipecolate oxidase
PCS	Pentaketide chromone synthase
PKS-III	Type III Polyketide synthase.

(Continued)

Continued

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 05 October 2023
DOI 10.3389/fpls.2023.1218914

OPEN ACCESS

EDITED BY
Rajesh Chandra Misra,
John Innes Centre, United Kingdom

REVIEWED BY
Rowan Andrew Craig Mitchell,
Rothamsted Research, United Kingdom
Laura E. Bartley,
Washington State University, United States

*CORRESPONDENCE
Tao Wang
johnwt@cnbg.net
Yinfeng Xie
xxyyff@njfu.edu.cn

RECEIVED 08 May 2023
ACCEPTED 14 September 2023
PUBLISHED 05 October 2023

CITATION
Xu D, Wang Z, Zhuang W, Wang T and Xie Y (2023) Family characteristics, phylogenetic reconstruction, and potential applications of the plant BAHD acyltransferase family.
Front. Plant Sci. 14:1218914.
doi: 10.3389/fpls.2023.1218914

Family characteristics, phylogenetic reconstruction, and potential applications of the plant BAHD acyltransferase family

Donghuan Xu[1,2], Zhong Wang[2], Weibing Zhuang[2], Tao Wang[2]* and Yinfeng Xie[1]*

[1]Co-Innovation Center for Sustainable Forestry in Southern China, College of Life Sciences, Nanjing Forestry University, Nanjing, China, [2]Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (Nanjing Botanical Garden Mem. Sun Yat-Sen), Nanjing, China

The BAHD acyltransferase family is a class of proteins in plants that can acylate a variety of primary and specialized secondary metabolites. The typically acylated products have greatly improved stability, lipid solubility, and bioavailability and thus show significant differences in their physicochemical properties and pharmacological activities. Here, we review the protein structure, catalytic mechanism, and phylogenetic reconstruction of plant BAHD acyltransferases to describe their family characteristics, acylation reactions, and the processes of potential functional differentiation. Moreover, the potential applications of the BAHD family in human activities are discussed from the perspectives of improving the quality of economic plants, enhancing the efficacy of medicinal plants, improving plant biomass for use in biofuel, and promoting stress resistance of land plants. This review provides a reference for the research and production of plant BAHD acyltransferases.

KEYWORDS

BAHD acyltransferase, protein structure, catalytic mechanism, phylogenetic reconstruction, acylation reactions

1 Introduction

Acylation is a common chemical reaction in living organisms that catalyzes a series of oxygenated and nitrogenous compounds to synthesize corresponding ester and amide products (Wang et al., 2022). The BAHD acyltransferase family is a group of proteins that acylate primary and specialized secondary metabolites in plants. Members of this family mainly use acyl-coenzyme A as the acyl donor to produce various volatile lipids, modified anthocyanins, and compounds related to plant resistance to pathogenic microorganisms,

thus playing important roles in signal transduction, stress defense, and metabolic homeostasis (Rosa and Neish, 1968; Suzuki et al., 2004b; D'Auria, 2006).

The BAHD acyltransferase family was named according to the first letter of each of the first four biochemically characterized enzymes within this family: benzylalcohol O-acetyltransferase (BEAT), anthocyanin O-hydroxycinnamoyltransferase (AHCT), anthranilate N-hydroxycinnamoyl/benzoyltransferase (HCBT), and deacetylvindoline 4-O-acetyltransferase (DAT) (D'Auria, 2006). Members of the BAHD family have been reported in model plants, such as Arabidopsis, Barley, Rice, and Poplar, as well as in important medicinal and economic plants, including Pear, Chinese staff vine (*Celastrus angulatus*), Jasmine, Tea, and Taxus (Aktar et al., 2022; Grienenberger et al., 2009; Zhang and Xu, 2018; Liu et al., 2020; Yamane et al., 2020; Yan et al., 2020; Wang et al., 2021a; Wang et al., 2021b; Zhao et al., 2021; Yuan et al., 2022). These BAHD proteins are involved in the formation of a variety of plant-derived active acylated natural products and their precursors, such as anthocyanins, alkaloids, aromatic alcohols/amines, aliphatic alcohols/amines, terpenoids and sugar detivatives. Based on the clade relationships of BAHD family, a series of representative compounds were shown in Figure 1, including cyanidin 3-O malonylglucoside, vinorine (a precursor to vincristine), geranyl acetate, coumaroyl-agmatine, caffeoyl quinic acid, and paclitaxel (taxol). A deeper understanding of these modifications may provide new opportunities for metabolic engineering and synthetic biology of such compounds.

Here, we review plant BAHD acyltransferases and the characteristics, including the protein structure and catalytic mechanisms. We describe the types of acylation reactions mediated by BAHD acyltransferases, as well as the potential functional differentiation, which has been studied by phylogenetic reconstruction. Moreover, the BAHD family has broad potential applications in human activities and four key aspects will be discussed, including improving the quality of economic plants, enhancing the efficacy of medicinal plants, improving plant biomass for use in biofuel, and promoting the stress resistance of land plants. This review provides a theoretical and practical basis for further research on the functions of plant BAHD acyltransferases and their potential applications.

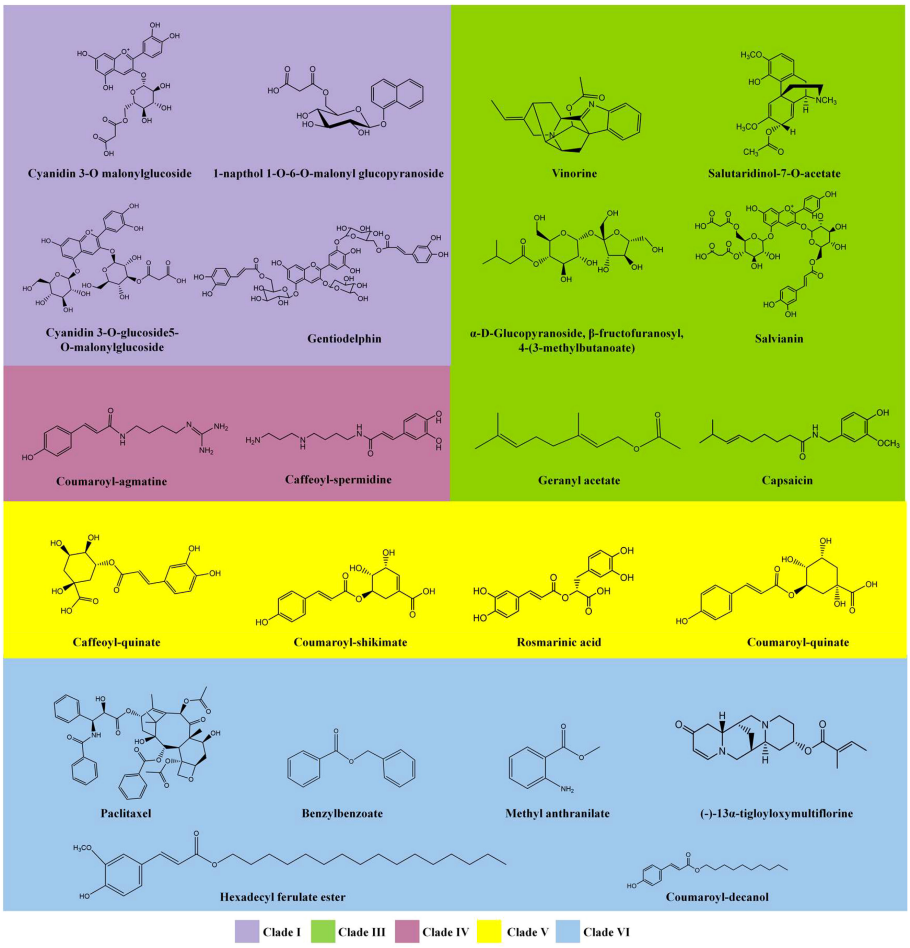


FIGURE 1
Representative metabolites that are acylated by BAHD family members. Compounds are categorized based on the clade relationships of the respective BAHD family members (Figure 3). More structures could be acquired from Supplemental Table S1.

2 Characteristics of BAHD acyltransferases

2.1 Protein structure

The BAHD members are globular proteins that are mostly localized in the cytosol, while a few are localized in the nucleus, such as *Medicago truncatula* anthocyanin 5-O-glucoside-6'''-O-malonyltransferase (MtMaT1) (Yu et al., 2008). They have a molecular mass ranging between 48 and 55 kDa, and the average number of amino acids is 445 (Suzuki et al., 2004b). The primary structure of plant BAHDs is varied, and some sequences with different clades only show 10–30% similarity at the amino acid level (Rosa and Neish, 1968); while the similarity within some pairwise comparisons of functionally equivalent members from different species, such as proteins that synthesize benzoic acid methyl esters, is as high as 90% (Nakayama et al., 2003). Despite the varied similarity between BAHD proteins, the typically sequences of all proteins in this family share two conserved motifs: HXXXD and DFGWG (Unno et al., 2007). The HXXXD motif is located near the center of each enzyme, which is essential for catalysis, and is absolutely conserved in BAHD acyltransferases. The DFGWG motif is located near the C-terminal of the protein and might play an important role in the catalytic process and binding of CoA. The DFGWG motif is not entirely conserved and, for example, is DFGFG, DFGWA, or DFGWK in poplars (Liu et al., 2016a). Moreover, the BAHD family members responsible for the synthesis of anthocyanins often contain an additional conserved motif, i.e., YFGNC (Nakayama et al., 2003).

Despite the sequence differences among BAHD members, their spatial structures are similar (Unno et al., 2007). The first characterized crystal structure of a BAHD member was of *Rauvolfia serpentina* vinorine synthase (*Rs*VS) (PDB ID: 2BGH), a globular protein consisting of two nearly equal-sized domains connected by a crossover loop (amino acids 201–213) consisting of 14 β-strands (β1–β14) and 13 α-helixes (α1–α13) (Figure 2A). The HXXXD motif is located at the active center between the two domains, while the DFGWG motif is located at the intersection between β11 and β13, far from the active site. Both domains play an important role in maintaining the catalytic function and binding to the donors and substrates (Ma et al., 2005). The first characterized crystal structure of N-acyltransferase (refers to the acyltransferases using nitrogenous metabolites as substrates) in the BAHD family was of *Hordeum vulgare* agmatine coumaroyltransferase (*Hv*ACT) (PDB ID: 7CYS). The structure shares some commonality with *Rs*VS, that is, they both consist of two domains connected by a long and large crossover loop, but the *Hv*ACT contains 18 β-strands (β1–β18) and 13 α-helixes (α1–α13) (Yamane et al., 2020) (Figure 2B). The crystal structures of BAHD members have been published one after another, and the number of described structures is currently 26 (Garvey et al., 2008; Garvey et al., 2009; Lallemand et al., 2012; Manjasetty et al., 2012; Walker et al., 2013; Eudes et al., 2016b; Levsh et al., 2016; Chiang et al., 2018). The clarity of the crystal structures of BAHD acyltransferases contributes to the understanding of the conserved domains shared by the BAHD family and provides a basis for the exploration of catalytic mechanisms.

2.2 Catalytic mechanisms

The acylation mediated by BAHD acyltransferases involves CoA thioesters as the acyl donor, including acetyl-CoA, malonyl-CoA, succinyl-CoA, benzoyl-CoA, cinnamoyl-CoA, feruloyl-CoA, caffeoyl-CoA, sinapoyl-CoA, and coumaroyl-CoA. Moreover, these BAHD acyltransferases typically utilize alcohols as acceptors to generate the corresponding esters, including flavonoids, anthocyanins, and terpenoids, or use amines as acceptors to produce amide compounds, such as polyamines and alkaloids. Interestingly, the BAHD members differ greatly in their use of donors and acceptors. For example, alcohol acyltransferases are

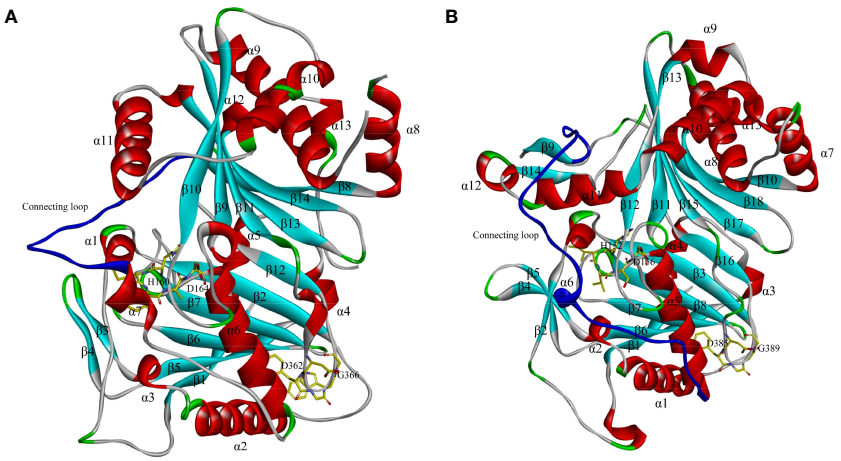


FIGURE 2
Crystal structures of *Rauvolfia serpentina* vinorine synthase (*Rs*VS) **(A)** and *Hordeum vulgare* agmatine coumaroyltransferase (*Hv*ACT) **(B)**. The structure of *Rs*VS is reproduced according to reference reported by Ma et al. (2005), and the structure of is *Hv*ACT reproduced according to reference reported Yamane et al. (2020).

ubiquitous in plants, which accept diverse substrates for acylation, thus producing a variety of volatile lipids, including aromatic acid esters, short-chain fatty acid esters, and monoterpene esters (Hampel et al., 2009). In contrast, the acyltransferases involved in paclitaxel biosynthesis are found only in *Taxus* species, which form an independent phylogenetic branch and use specific substrates to produce taxanes (Kusano et al., 2019). Acylation enriches the structural diversity of the metabolites and improves their stability, lipid solubility, and bioavailability, which promotes their multiple functions in plant growth, as well as their potential applications in human healthcare (Liu et al., 2016a).

The research on crystal structures has improved the understanding of the acylation mechanism, and homology modeling and molecular docking have been widely used to study the acylation process and predict the potential functions of the products. According to the structural characteristics of *Rs*VS, Ma et al. (2005) proposed a catalytic mechanism in which the reaction is triggered by the histidine (His) residue of the HXXXD motif, which deprotonates the -NH_2 and -OH on the substrates, allowing a nucleophilic attack on the carbonyl carbon of the CoA thioester. Next, a tetrahedral intermediate is formed between the thioesterase and the corresponding substrate, which re-protonates to produce the acylated ester or amide. As the most critical amino acid residue in acylation reaction, the His residue is absolutely conserved in all family member sequences, which was confirmed by site-directed mutagenesis of BAHD acyltransferases (Unno et al., 2007; Navarro-Retamal et al., 2016). Moreover, the aspartate (Asp) residue of the HXXXD motif seems to play a structural role, since it points away from the active site. Site-directed mutagenesis of converting Asp residue to amino acids with poor nucleophilic ability in the HXXXD motif of *Taxus cuspidata* 10-deacetylbaccatin III-10β-O-acetyltransferase (*Tc*DBAT) led to over 90% loss of enzyme activity (Wang et al., 2021a). A similar case was also reported by Morales-Quintana et al. (2013) in *Vasconcellea pubescens* alcohol acyltransferase (*Vp*AAT1), while *in vitro* site-directed mutagenesis of H166 and D170 in the HXXXD motif of *Vp*AAT1 induced a loss of activity, confirming the role of these amino acid residues that were directly involved in catalysis. An early mutagenesis study provided evidence that the DFGWG motif is necessary for catalysis and appears to have a structural function, while later studies used similar techniques suggested that the contribution of DFGWG motif was primarily binding to substrates rather than affecting catalytic efficiency (Morales-Quintana et al., 2015).

3 Phylogenetic reconstruction

Phylogenetic analyses of the BAHD family have shown different results, possibly due to the different software and criteria used. D'Auria (2006) was the first to divide the 46 family members into five clades according to the type of substrate used or the conditions under which the genes and enzymes are active. Tuominen et al. (2011) reconstructed phylogenetic relationships using 69 known functional BAHD acyltransferases and other not functionally characterized putative members from Poplar, Arabidopsis, Oryza, Medicago, and Vitis, and obtained results wherein, although consistent with those reported by D'Auria, the family members were divided into eight main clades. Wang et al. (2022) used the amino acid sequences of 129 acyltransferases characterized between 1997 and 2020 to construct a BAHD phylogenetic tree using the maximum-likelihood method and proposed that the type of acceptors should be the main basis for the phylogeny of BAHDs. Moreover, Moghe et al. (2022) listed hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase (HST) and hydroxycinnamoyl-CoA:quinate hydroxycinnamoyl transferase (HQT) as an independent clade based on the reports by D'Auria (2006). Another clade was added for algae acyltransferases and a loosely defined clade with coniferyl alcohol acetyltransferases as the main members, thus further describing the evolutionary relationships of these enzymes. Plant evolution is thought to have resulted in neofunctionalization by enzyme promiscuity and substrate permissiveness, which led to the synthesis of new metabolites and the establishment of related biosynthetic pathways (Fani and Fondi, 2009), while taxon-specific BAHD family expansion via gene duplication could be an evolutionary process contributing to metabolic diversity across plant taxa (Tuominen et al., 2011). This could explain the rapid expansion of metabolic pathways based on different substrates mediated by the BAHD family. In this review, an overview of the acylation reactions involving different family members is presented through different clades of the BAHD family, and representative amino acid sequences were selected for phylogenetic tree construction, which was used to reveal the functional evolution of BAHD family members and predict the function and activity of unidentified proteins (Figure 3).

3.1 Clade I – Synthesis of flavonoids/anthocyanins/phenolic glucosides

The members of Clade I are mostly involved in the acylation of flavonoids, anthocyanins, and phenolic glucosides, in particular, the acylation of anthocyanins. The proteins contain the YFGNC motif, which is a signature sequence of anthocyanin acyltransferases (Yu et al., 2009). The acylation of flavonoids generally occurs on the C6-OH of the glucosyl group, which possess region-specificity of acyl transfer (D'Auria, 2006). Based on the specificity of the donors, these mediated BAHDs are often classified as aromatic and aliphatic acyltransferases (Nakayama et al., 2003). Moreover, most flavonoid acyltransferases exhibit a wide range of acceptor availability that using glycosylated anthocyanins, isoflavons, flavanones, flavanols, and flavones as acceptors, and resulting in a huge diversity of potential products (Saigo et al., 2020). For example, *Arabidopsis thaliana* acyltransferase 1 (*At*3AT1) and acyltransferase 2 (*At*3AT2), which can utilize more than seven flavonoid 3/7-O-glucosides as substrates for acylation (Luo et al., 2007). Structural analysis of *Gentiana triflora* anthocyanin 5,3-aromatic acyltransferase (Gt5,3'AT) (PDB ID: 7DEV) reveals the residues responsible for acyl acceptor specificity and indicates that the selectivity for acyl transfer is determined by the C-terminal lobe of the protein (Murayama et al., 2021).

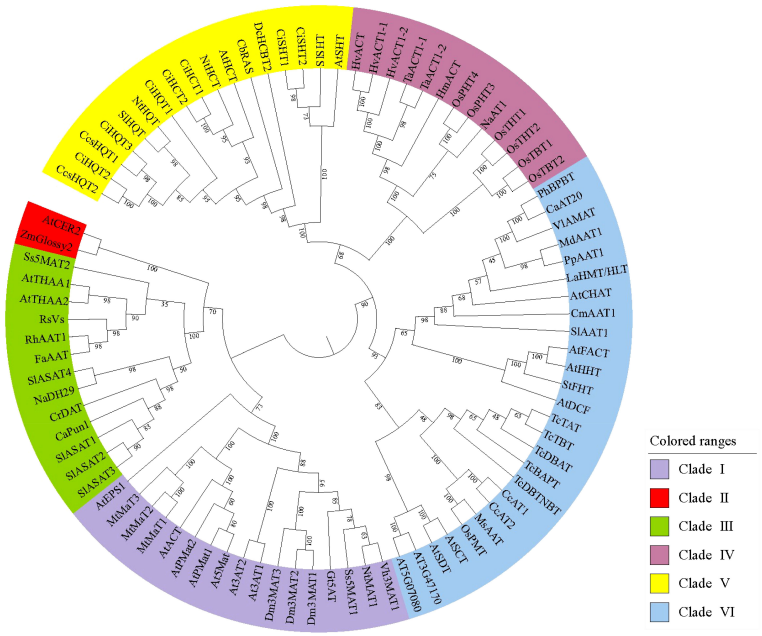


FIGURE 3
Maximum-likelihood phylogenetic tree of plant acyltransferases. These sequences were aligned using ClustalW as implemented in MEGA 11. Evolutionary analyses were conducted in MEGA 11 with 1000 bootstrap replicates, and the evolutionary tree was visualized by iTOL. Details of these sequences are list in Supplemental Table S1.

3.2 Clade II – Elongation of epicuticular waxes

Clade II mainly contains *Arabidopsis thaliana* Eceriferum (*At*CER2), *Zea mays* Glossy2 (*Zm*Glossy2), and *Zm*Glossy2-like, which are involved in the elongation of epicuticular waxes for preventing tissue moisture loss and resisting pathogen attack (Yu et al., 2009). However, Clade II proteins could not meet the original criteria of containing both HXXXD and DFGWG motif, and their functional activities need further investigation, so it is still unclear whether they belong to the BAHD family. Glossy2-like and Glossy2 functionally complement the *At*CER2 mutation, indicating a conserved function, while differences in the utilization of longer alkyl-chain acyl lipids indicate functional differentiation of these two maize enzymes (Alexander et al., 2020). In addition, based on mutation studies, a series of CER2-like proteins essential for epidermal wax production were characterized in *Arabidopsis thaliana* and *Zea mays*. Interestingly, CER2-like proteins possess sequence similarity with CER2, and make a major contribution to the chain-length specificity of fatty acid elongation (Haslam et al., 2017), but they seem to not rely on acyl transfer activity for their biological function (Haslam and Kunst, 2020). Further studies are expected to explore ligand binding and thiolipase activity, which will establish or rule out their most plausible functions for the protein family.

3.3 Clade III – A pluralistic clade

Clade III contains a series of alcohol acyltransferases involved in the biosynthesis of volatile lipids in flowers and mature fruits. Most of them use acetyl-CoA as the donor and accept different kinds of alcohol compounds as substrates, such as geraniol and *n*-octanol (Shalit et al., 2003). Interestingly, the alcohol acyltransferases often use various substrates and produce a wide range of products and, as a result, substrate promiscuity is considered a hallmark for this subset of BAHD members (Moghe et al., 2022). Moreover, this clade also contains proteins involved in the modification of alkaloids, such as vindoline, dimethylmorphine, and capsaicin (De Luca et al., 1985; Grothe et al., 2001; Milde et al., 2022). The alkaloids are a class of nitrogenous compounds, most of which possess a complex ring structure containing nitrogen and show significant biological activity (Mondal et al., 2019). The acyl transfer of the alkaloid-associated BAHD members depends on modification of the hydroxyl or amine moieties within the alkaloids. Recently, a series of acylsugar acyltransferases (ASATs) were classified into clade III. The ASATs are produced in type I/IV trichome tip cells and use acetyl or C2–C12 short-chain acyl groups to produce O-acylsugars (Schilmiller et al., 2012; Moghe et al., 2017). In theory, ASATs can produce various compounds, but the ultimate distribution of acylsugars within a species depends on the preferences of ASATs for substrates, which effectively narrows the

acylsugar structural space (Taguchi et al., 2005; Unno et al., 2007). In addition, Clade III also contains various enzymes involved in the modification of anthocyanin glycoside, terpene, and aliphatic amine, which have been characterized in different angiosperm species, but their functional association was less clear due to the small size of subclades in this branch.

3.4 Clade IV – Agmatine coumaroyltransferases and putrescine hydroxycinnamoyltransferases

Several members of Clade IV have been reported, such as agmatine coumaroyltransferase (ACT) in barley (*Hv*ACT) and wheat (*Triticum aestivum* ACT, *Ta*ACT) (Kage et al., 2017; Yamane et al., 2020; Yamane et al., 2021). These enzymes contain the conserved HXXXD motif, but the DFGWG motif, which is shared by all members of the family, is slightly altered with glycine instead of tryptophan (Burhenne et al., 2003). The putrescine hydroxycinnamoyl- transferases (PHTs) have also been described in tobacco (*Nicotiana attenuata* acyltransferase, *Na*AT1), rice (*Oryza sativa* PHT 1-4, *Os*PHT1-4), and tomato (*Solanum lycopersicum* ACT, *Sl*ACT1) (Onkokesung et al., 2012; Campos et al., 2014; Fang et al., 2022). Both ACT and PHT enzymes can use a variety of amines as acceptors, including agmatine, putrescine, and spermidine, as well as donors such as 4-coumaryl-CoA, caffeoyl-CoA, feruloyl-CoA, and cinnamoyl-CoA, to generate corresponding amide compounds (Kruse et al., 2022). However, several N-acyltransferases have relatively high donor specificity. For example, *Os*TBT2 and *Os*PHT3 only use benzoyl-CoA and *p*-coumaryl-CoA as the donors, respectively; while *Hv*ACT exhibits donor preferences to *p*-coumaroyl-CoA (Burhenne et al., 2003).

3.5 Clade V – Hydroxycinnamoyl-CoA: shikimate acid hydroxycinnamoyl transferase and hydroxycinnamoyl-CoA: quinate acid hydroxycinnamoyl transferase

Clade V mainly contains hydroxycinnamoyl-CoA:shikimate acid hydroxycinnamoyl transferase (HST) and hydroxycinnamoyl-CoA:quinate acid hydroxycinnamoyl transferase (HQT). HST and HQT have hydroxycinnamoyl transferase (HCT) activity and use phenolic-CoA as donors, such as hydroxycinnamoyl-CoA, *p*-coumaryl-CoA, and caffeyl-CoA, to catalyze the acylation reaction with shikimic acid and quinic acid as substrates (Moglia et al., 2016). HST and HQT are involved in the phenylpropane pathway in plants, and their products are important intermediates in lignin synthesis, mediating plant growth and development (Kriegshauser et al., 2021). *In vitro* experiments have reported differences in the substrate selection of these enzymes, such as *Solanum lycopersicum* HQT (*Sl*HQT) and *Cynara cardunculus* HQT (CcsHQT) (Comino et al., 2007; Moglia et al., 2014), which only accept shikimic acid and/or quinic acid, while others do not strictly distinguish between these two substrates. Based on the crystal structure of HCT and mutagenesis experiments, this substrate specificity could be explained by the so-called "arginine handle" and the surrounding amino acids (Chiang et al., 2018). Interestingly, some enzymes, such as *Coleus blumei* rosmarinic acid synthase (*Cb*RAS) and *Coleus blumei* HST (*Cb*HST), can accept other nonrelated substrates, including 3-hydroxyanthranilate, 5-hydroxyanthranilate, 3-hydroxybenzoate, 2,3-dihydroxybenzoate, 3-aminobenzoate, gentisate, catechol, and protocatechuate (Sander and Petersen, 2011; Eudes et al., 2016a). Certainly, the reverse reactions catalyzed by these members also contribute to produce hydroxycinnamoyl-CoA and the former acceptor substrate or split hydroxycinnamoyl-CoA into free hydroxycinnamic acid and CoA (Berger et al., 2006).

3.6 Clade VI – A multicomponent clade

Members in Clade VI show diverse activities. They utilize substrates ranging from terpenoids to medium-chain alcohols, in association with major phylogenetic branches within this clade. Several enzymes are associated with the biosynthesis of volatile esters, including *Petunia hybrida* benzoyl-CoA:8-debenzoylpaeoniflorin 8-O-benzoyltransferase (*Ph*BPBT), *Arabidopsis thaliana* acetyl-CoA:(Z)-3-hexen-1-ol acetyltransferase (*At*CHAT), and *Solanum lycopersicum* alcohol acyltransferase 1 (*Sl*AAT1) (Okada et al., 2005; Molina and Kosma, 2015; Liu et al., 2016b). These enzymes are characterized by the ability to acylate aliphatic and/or aromatic alcohols using aliphatic and/or aromatic CoA-activated donors. Similar to the alcohol acyltransferases in Clade III, these enzymes show a wide substrate diversity, but aliphatic and aromatic acyl acceptors are not the main basis for their clade division. Specialized taxane acyltransferase-mediated acylation reactions are thought to be important steps in modifying the core backbone of paclitaxel, an important anticancer drug. To date, five key acyltransferases have been characterized as being involved in the paclitaxel pathway (Wang et al., 2021a), and these enzymes are clustered into an individual group within Clade VI. With the publication of the whole genome of *Taxus chinensis* var. *mairei*, as many as 53 candidate genes encoding taxane acyltransferases have been reported, which potentially enlarged the size of branch for taxane acyltransferases. It would be worthwhile to investigate whether these candidates function in paclitaxel biosynthesis in the future (Xiong et al., 2021). Members of Clade VI also include lipid-related acyltransferases, such as *Marchantia emarginata* ω-hydroxyacid/fatty alcohol hydroxycinnamoyltransferase (*Me*HFT) and *Marchantia polymorpha* feruloyltransferase (*Mp*FHT), which are involved in the biosynthesis of cutin monomers (Wang et al., 2017). These enzymes use aroyl-CoA as donors, including feruloyl-CoA, *p*-coumaroyl-CoA, caffeoyl-CoA, and sinapoyl-CoA, and accept fatty acids or fatty alcohols as substrates to produce alkyl hydroxycinnamates. Functionally, they are similar to those of Clade II, but orthologs of the characterized members of Clade II are found only in angiosperms. Moreover, the clade also includes the first anthraniloyl-CoA:methanol acyltransferase (AMAT) found in the BAHD family with O-aminobenzoylkylase A as the acceptor for synthesis of aromatic esters, as well as the *Lupinus albus* tigloyl-CoA: (−)- 13α-hydroxymultiflorine/(+)- 13α-hydroxylupanine O-tigloyltransferase (*La*HMT/HLT), which is related to the biosynthesis of quinolicidine alkaloids (Goulet et al., 2015; Huang et al., 2022).

3.7 Other members

A few acyltransferases show a specific distribution in different clades due to functional diversity and complex evolutionary dynamics of the BAHD family. For example, *Salvia splendens* anthocyanin 5-O-glucoside-4'''-O-malonyltransferase 2 (*Ss*5MaT2), although associated with the modification of anthocyanins, was classified in Clade III because it does not contain the conserved motifs that are common in Clade I (Suzuki et al., 2004b). This also reflects different routes for anthocyanin acyltransferase activity. Due to the diversity of available substrates, members with HCT activity are found in different clades, except for Clade II, and differences are also observed in the number of hydroxycinnamoyl residues carried by the products (Roumani et al., 2021). The different groups that are modified in the alkaloids, as well as the wide diversity in the sources of precursors for nitrogenous heterocyclic compounds, have resulted in a scattered distribution of alkaloid acyltransferases throughout the BAHD family. Moreover, Moghe et al. (2022) reported an algal acyltransferase clade based on HQT activity and an undefined clade consisting of coniferyl alcohol acetyltransferases. Wang et al. (2022) reported a class of lipid-related acyltransferases that use long-chain fatty acids as donors and accepting glycerol derivatives as acceptors to form glycerolipids. Since their functions have not been fully determined using *in vitro* experiments, they were not classified in Figure 3.

4 Potential applications of BAHD acyltransferases

As we catalog below, BAHD members mediate diverse ecological interactions in plants to ensure their normal growth and development. For example, the BAHD family of proteins is involved in the following: improving plant pollination by forming brilliant colors and scented volatiles; supporting plant morphology by mediating lignin synthesis; protecting plant reproduction by mediating pollen wall formation; and resisting different biotic and abiotic stresses by forming a variety of secondary metabolites in different tissues. Given the enzyme promiscuity of the BAHD family and the rapid evolutionary neofunctionalization, understanding the BAHD family will enhance our knowledge of plant ecology. Moreover, these extensive acylated modifications may have applications in, for example, economic development and human healthcare. Here, we focus on the application prospects of the BAHD family in the development of human activities and summarize the functions of the BAHD family mediating the diverse plant traits.

4.1 Improving the quality of economic plants

Volatile esters play an important role in the formation of aromas in the plant leaf, flower, and fruit. The BAHD members involved in the synthesis of volatile esters are mainly from Clades III and VI, which possess the potential to enhance the quality of economic plants. For example, *Malus domestica* alcohol acyltransferase 2 (*Md*AAT2) is a key enzyme in the last step of apple volatile ester biosynthesis, which is a key factor in guaranteeing fruit quality (Li et al., 2006). Since alcohol acyltransferases mediate the synthesis of specific esters in different plants, which could be an important reason that causes different fruits to exude unique smells. For example, short-chain esters derived from fatty acids contribute to the characteristic flavor of pear, while its alcohol acyltransferase, *Pa*AAT1, is thought to be closely associated with the production of C6 esters (Zhou et al., 2021). The odor in peaches is mainly determined by γ-decalactone, whose synthesis is mainly catalyzed by *Prunus persica* alcohol acyltransferase (*Pp*AAT) (Song et al., 2021). The alcohol acyltransferases also demonstrate broad substrate diversity for driving the delicate development of characteristic odors. For example, overexpression of the gene encoding *Fragaria vesca* alcohol acyltransferase (*Fv*AAT) increases the proportion of different esters to volatiles, including octyl acetate, ethyl caproate, octyl hexanoate, and ethyl caprylate, which indicates that *Fv*AAT can influence the volatile ester composition of strawberry fruit (Dong et al., 2018). The *Vitis vinifera* alcohol acyltransferase (*Vv*AAT)-mediated ester components in grapes are important markers to distinguish different cultivars (Ji et al., 2022).

The formation of leaf and floral aromas has also received attention. The composition of *Cymbopogon winterianus* leaf oil mainly consists of acyclic monoterpenols (geraniol and citronellol) and their esters (geranyl acetate and citronellyl acetate), and the synthesis of these compounds is associated with citronellol alcohol acyltransferase (CAAT) (Kumar, 2020). The function of O*smanthus fragrans* alcohol acyltransferase1 (*Of*AAT1) was predicted to be similar to that of *Petunia hybrida* coniferyl alcohol acyltransferase (*Ph*CFAT), which can catalyze coniferol to produce coniferol diester that is further degraded to isoeugenol. However, it has also been documented that *Of*AAT1 might be involved in the synthesis of fatty acid esters. The mechanism is the same as that of *Glycine max* cholineacetyltransferase (*Gm*CHAT) and *Solanum tuberosum* cholineacetyltransferase (*St*CHAT) (Liu et al., 2016b). Benzylalcohol *O*-acetyltransferase uses benzyl alcohol and acetyl-CoA to produce benzyl acetate, a floral volatile found in a few plants, including *Clarkia pulchella*, *Prunus mume*, and *Chimonanthus praecox*. Interestingly, BEAT is also capable of accepting other alcohols as substrates, but the highest catalytic efficiency is achieved when benzyl alcohol is the substrate (Dudareva et al., 1998). These studies provide important clues to the potential phylogeny and functional diversity of alcohol acyltransferases.

Anthocyanins are natural water-soluble pigments in plants, including delphinidin, cyanidin, and pelargonidin, which are catalyzed by glycosyltransferases to form a variety of anthocyanin glycosides with different colors stored in vacuoles; while the latter are further modified by methyltransferases and acyltransferases, and through molecular superposition and interaction effects, ultimately produce different colors in organs and tissues of different plants (Zhao and Tao, 2015). Florio et al. (2021) reported that the expression level of the *Solanum melongena* anthocyanin acyltransferase

(*SmelAAT*) gene correlated with the accumulation of lycopene in the pericarp, which provides key evidence for the involvement of *Smel*AAT in anthocyanin modification. Anthocyanin acyltransferases also exhibit diverse preferences for the combination of donors and substrates. Among a host of characterized anthocyanin acyltransferases, the *Vitis vinifera* anthocyanin acyltransferases (*Vv*3AT), reported by Rinaldo et al. (2015), showed a preference for anthocyanin monoglycoside molecules, allowing the use of aromatic and aliphatic acyl-CoA thioesters as substrates, which was different from the previously characterized natural BAHD acyltransferase. Suzuki et al. (2004a) reported that *Dendranthema morifolium* malonyl-CoA:anthocyanidin 3-O-glucoside-6″-O-malonyltransferase (*Dm*3MaT2) could catalyze the malonylation of anthocyanidin 3-O-glucoside to produce anthocyanidin 3-O-6″-O-malonylglucoside and subsequently catalyze the latter to produce anthocyanidin 3-O-3″,6″-O-dimalonylglucoside. Extensive studies have shown that aromatic acylation of anthocyanins through intramolecular condensation leads to a more stable molecular structure and coloring, while fatty acid acylation increases the water solubility of the compounds, protects the glycosides from enzymatic hydrolysis, stabilizes the structure, and contributes to the accumulation of the compounds in the vesicles (Bloor and Abrahams, 2002; Luo et al., 2007; Zhang et al., 2014). Therefore, in-depth studies on the catalytic pattern of anthocyanin acyltransferases in plants will help to improve the external and internal qualities and promote bioengineering to produce anthocyanins.

4.2 Enhancing the efficacy of medicinal plants

Acylation plays an important role in the structural modification and pharmacological activity of plant secondary metabolites and is crucial for obtaining structural diversity and active medicinal lead compounds. A series of BAHD members participate in the synthesis of alkaloids and terpenoids, which are very important components of herbal medicines, with a broad spectrum of antitumor, antiviral, antibacterial, antimalarial, and analgesic biological and pharmacological activities (Yu et al., 2009). Vinblastine is a clinically important antitumor chemotherapeutic agent. Vincenzo et al. (1985) first reported that acetyl coenzyme A: deacetylvincristine O-acetyltransferase is involved in the formation of vindoline, a precursor of vincristine, which initiated the research on the synthesis and modification of alkaloids mediated by BAHD acyltransferases (Grothe et al., 2001). *Rs*VS is a key enzyme involved in the biosynthesis of the antiarrhythmic drug ajmaline that can reversibly catalyze the synthesis of the sarpagan alkaloid 16-epi-vellosimine with acetyl-CoA to obtain the ajmalan-like alkaloid vinoline. As the first reported crystal structure of the BAHD family, *Rs*VS confirms the shared key properties of the BAHD family. Grothe et al. (2001) found that *Papaver somniferum* salutaridinol-7-O-acetyltransferase (*Ps*SALAT) catalyzes the phenanthrene alkaloid salutaridinol to produce salutaradinol-7-O-acetate, which is an immediate precursor of morphine. Capsaicin has antioxidant and anticancer effects, and the final step of its biosynthetic pathway involves the N-acylation of vanillin with a variable-length fatty acyl-CoA donor. However, the capsaicin synthase Pun1 is highly insoluble, and its role in capsaicin accumulation is based exclusively on genetic studies and no biochemical activity has been reported. The pharmacologically important anticancer drug paclitaxel is a complex diterpene alkaloid whose biosynthesis involves five acyl modifications. Among them, taxadiene-5α-ol-O-acetyl transferase (TAT) catalyzes the first acylation reaction of the paclitaxel pathway, converting taxadiene-5-α-ol to taxadiene-5-α-acetate, which is considered a slow step of the downstream hydroxylation reaction, while 10-deacetylbaccatin III-10-O-acetyl transferase (DABT) is the key rate-limiting enzyme of the paclitaxel pathway, acetylating the C-10 hydroxyl group of 10-deacetylbaccatin III (10-DAB) to produce baccatin III, an important intermediate of taxol (Wang et al., 2021a).

Due to the low content of natural medicinal active ingredients in plants, the overexpression of the corresponding genes by means of genetic transformation techniques or the reorganization of synthetic pathways using different plants or microorganisms is considered the most promising method to enhance the yield and quality of the products. However, the existing genetic transformation systems, such as the suspension cell line of Taxus and the hairy root pathway mediated by *Agrobacterium perfringens* in periwinkle, are difficult to apply at industrial scales due to factors such as the slow cell growth and unstable production capacity. Moreover, transient acylation and deacylation occur in the synthetic pathways of the above-mentioned medicinal components for flux regulation of organelle targeting, but the promiscuity and preferences of the corresponding acyltransferases have not been fully investigated (McElroy and Jennewein, 2018). Furthermore, most acyltransferases for terpenoids or alkaloids that are heterologously expressed in different engineered strains, such as *Escherichia coli* and *Saccharomyces cerevisiae*, still have drawbacks such as low expression levels, incompatible solubility, and poor stability (Wang et al., 2022). Together with the long synthetic pathways, these factors make it hard to obtain desirable products from heterologous synthesis, or the purification of products cannot meet the actual demand. There is still a long way to go before BAHD members can be efficiently utilized in synthetic biology and metabolic engineering.

4.3 Improving plant biomass for use in biofuel

The manipulation of cell wall polymers can produce plants that are useful for biofuel production. Lignin is a principal structural component of cell walls and is formed from the polymerization of single lignin alcohols (coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol) to produce different lignin units. Coumaryl alcohol produces 4-hydroxyphenyl (H) units, and coniferyl alcohol and sinapyl alcohol produce guaiacyl (G) and syringyl (S) units, respectively (Vanholme et al., 2019). Lignin levels affect plant quality, and excessive lignin levels can cause reduced digestibility. HCT is one of the key enzymes affecting the biosynthesis of lignin G/S units, and research on HCT has focused on its association with lignin (Kriegshauser et al., 2021). A low level of HCT leads to a

decrease in lignin content and results in a significant increase in biomass saccharification efficiency (Serrani-Yarce et al., 2021). However, it also leads to drastic changes in the plant phenotype. Based on the potential functions of HSTs in the pathways of phenylpropanoid metabolism and hormone response, Dong et al. (2022) suggested that HCTs may influence plant lignin composition and development by altering hormone content. Notably, some BAHD members involve the modification of monolignol monomers, which are then incorporated into lignin polymers as acyl conjugates. For example, *Angelica sinensis* feruloyl-CoA monolignol transferase (*As*FMT), which specifically accepts Feruloyl-CoA as the acyl donor but can accept all three monolignol monomers as substrates, shows an increase in saccharification yield in poplar strains with increased levels of feruloylated conjugates mediated by BAHD (Wilkerson et al., 2014). Liu et al. (2022b) reported the crystal structure of *As*FMT, and revealed the key action sites concerning its mediated acylation reaction, which provides insights into the formation of monolignol ferulate conjugates. These data may help in the design of strategies to optimize the lignin composition and amount in biorefining.

BAHD family also plays important roles in the addition of phenolic acids, such as ferulic acid (FA) and *p*-coumaric acid (pCA) to form ester-linked moieties on the xylan backbone of glucuronoarabinoxylan (GAX) (Chandrakanth et al., 2023). In the complex structure of plant cell walls, cellulose is protected by a network of xylan chains that cross-link via FA or lignin. The role of pCA on GAX is less obvious than that of FA, because the oxidative coupling of pCA is much weaker than that of FA. One possibility is that pCA-GAX participates in radical transfer and thus catalyzes the oxidative coupling of neighboring FA on GAX. The degree of ferulic acidification in the cell wall is related to biomass digestibility, as high FA increases the resistance of biomass conversion to ethanol (de Oliveira et al., 2015). *Oryza sativa* acyltransferase 10 (*Os*AT10) was the first reported putative *p*-coumaroyl CoA arabinofuranose transferase (PAT), since overexpression of *OsAT10* induced a 5-fold increase in pCA levels in young green tissues with a concomitant 50% decrease in FA linked to GAX (Bartley et al., 2013). Interestingly, overexpression of *OsAT10* in *Sorghum bicolor* and *Panicum virgatum* and of Sugarcane *AT10* (*ScAT10*) in maize also significantly improve the biomass saccharification efficiency with increased *p*-CA in transgenic strains, but not universally accompanied by a decrease in FA-GAX (Li et al., 2018; Fanelli et al., 2021; Tian et al., 2021). Silencing of the putative FAT-encoding *Setaria viridis* acyltransferase 01 (*Sv*BAHD01/*Sv*AT9) gene decreased arabinoxylan feruloylation by 60% and increased biomass saccharification efficiency by 40–60% in the stems without changing the amount of lignin (de Souza et al., 2018). Similarly, suppression of the ortholog in Sugarcane (*ScBAHD01/ScAT9*) improved the digestibility of sugarcane straw by approximately 20% compared to non-transformed plants (de Souza et al., 2019). These results further indicate that the exploitation of BAHD acyltransferases for plant modification to improve biomass digestibility and to generate optimal biofuel plants is an important challenge for the sustainable production of advanced biofuels.

4.4 Promoting the stress resistance of land plants

Plant BAHD acyltransferases are involved in the synthesis and modification of a wide range of primary and secondary metabolites, thus enhancing the resistance of plants to different biotic and abiotic stresses in different dimensions, which is essential to guarantee their survival and yield. On the one hand, the BAHD proteins could mediate the synthesis of many chemical substances to increase tolerance to various environmental stresses. For example, the resistance of plants to UV radiation is achieved by the BAHD members that mediate the synthesis of anthocyanins, and the resistance benefits from the formation of the so-called bridge-piled structure, leading to an intramolecular co-pigmentation-like effect (de Oliveira et al., 2015; Tohge et al., 2016). Overexpression of the gene that encodes *Taraxacum mongolicum* HQT (*Ta*HQT) increases the chlorogenic acid (CGA) content, which is used to enhance disease resistance and salt tolerance, in the transgenic plants (Liu et al., 2018). On the other hand, The BAHD acyltransferases could also activate multiple metabolic pathways simultaneously to improve plant stress resistance. For example, *Oryza sativa* acyltransferase (*Os*At10) can promote the growth rate and viability of cells, and it improves cold tolerance in *Arabidopsis*, cotton, and poplar by increasing the antioxidative enzyme activity and polyamine level. This provides a valuable genetic resource for breeding cultivars with high yield and high resistance (Tang and Thompson, 2022). Moreover, the *Camellia sinensis HCT* (*CsHCT*) gene is expressed under low temperature, drought, and high salinity environments and is involved in the response to abiotic stresses (Sun et al., 2018). The formation of epidermal waxes mediated by CER2-like proteins may mediate water rejection, particle adhesion, light reflection, and herbivore resistance of plant cells. There could also be crosstalk among different environmental stresses, that is, enhancing resistance to a single environmental stress increases resistance to other stresses at the same time (Haslam et al., 2015).

Pests and diseases are important limiting factors affecting crop yield and their ecological roles. Traditional chemical control enhances pest and disease resistance but affects the ecological balance and human health. Among the characterized BAHD acyltransferases, some clade members specifically mediate plant responses to pests and diseases and are the focus of research in the field of biological control. As an important chemical defense compound in plants, hydroxycinnamic acid amide exhibits a pivotal role in plant–pathogen interactions, and its antimicrobial activity and the mechanism involved in plant immune response have been fully elaborated (Liu et al., 2022a). Hydroxycinnamic acid amides are synthesized through the condensation of various biogenic amines with hydroxycinnamic acids via BAHD acyltransferases. ACT was first characterized in barley, catalyzing agmatine and hydroxycinnamoyl-CoA to produce hydroxycinnamoyl agmatine, the precursor of Hordatine, which is an antifungal compound that is abundant in yellowing barley seedlings (Burhenne et al., 2003; Ube et al., 2017). Homologous genes of ACT in Arabidopsis, Rice, and Wheat have been reported (Onkokesung et al., 2012; Peng et al., 2016; Kage et al., 2017; Yamane et al., 2021), and numerous studies have confirmed that ACT can mediate the synthesis

of neurotoxins similar to those found in spider and wasp venom to promote immunity against pathogenic bacteria and herbivorous pests (Dobritzsch et al., 2016; Bai et al., 2022). Hydroxycinnamoyl-coenzyme A (CoA): malate hydroxycinnamoyl transferase (HMT) and hydroxycinnamoyl-CoA: L-DOPA hydroxycinnamoyl transferase (HDT) were successively characterized in red clover leaves, which mediated the synthesis of phaselic acid and clovamide, protecting the plant from biotic and abiotic stresses. Notably, they share a 72% degree of identity, but differ substantially in substrate specificity, which provides valuable and new directions for rational design of BAHD enzymes with specific and desirable activities (Sullivan and Knollenberg, 2021). O-Acyl sugars produced by glandular trichomes are defense molecules against fungal pathogens and a chemical glue for trapping small insects. *Nicotiana tabacum* acyl-glycoacyltransferase 1 (*Nt*ASAT1) and acyl-glycoacyltransferase 2 (*Nt*ASAT2) are key enzymes involved in acylsugar assembly in tobacco and mediate a post-ingestive odor-tagging indirect defense mechanism, which provides new pathways and target genes for plant resistance against pathogens and insects (Chang et al., 2022). Moreover, members of Clade II, as well as lipid-related acyltransferases in Clade VI, such as *Arabidopsis thaliana* aliphatic suberin feruloyl transferase (*At*ASFT), are involved in the synthesis of plant cell membranes, keratin, and suberin. These substances provide a protective layer on the surface of plants to maintain water homeostasis and provide protection against pathogens. In summary, plant BAHD acyltransferases regulate the accumulation of different metabolites via acylation, thus protecting plants from external environmental changes, which provides new directions for plant ecology and genetic improvement.

5 Summary and prospects

BAHD acyltransferases play an important role in plant growth and development, stress responses, and synthesis and modification of secondary metabolites. In recent years, with the completion and improvement of plant genome sequencing and assembly, and the development of molecular biology techniques, such as molecular docking, homology modeling, RNA interference, targeted mutagenesis, and molecular dynamics simulation, an increasing number of BAHD family members have been discovered and characterized, laying a solid theoretical foundation for further elucidation of the biological functions of the plant BAHD family. Given the extensive potential applications of plant BAHD acyltransferases in human activities, there is still ample room for further exploration of BAHD acyltransferases. Currently, despite the increasing numbers of annotation of BAHD members, studies on their biochemical functions are limited. From a phylogenetic perspective, the role of a series of subfamily members in the evolution of plant geographic lineages remains to be further explored. Moreover, the biochemical properties of BAHD acyltransferases should be the focus of attention. The selection and preferences of different members for substrates and donors, as well as the construction of suitable *in vivo* and *in vitro* reaction systems, could contribute to the rapid development of BAHD acyltransferases in synthetic biology and metabolic engineering and promote their potential application.

Author contributions

DX contributed to conception and design of the study and wrote the manuscript. WZ and ZW organized the database. TW and YX read, and approved the submitted version. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research was funded by the National Natural Science Foundation of China (32101551), the Jiangsu Province Forestry Science and Technology Innovation and Promotion Project (LYKJ[2020]24), and the NBG Fund for Connotation Construction (NBGF202307).

Acknowledgments

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1218914/full#supplementary-material

References

Aktar, S., Bai, P., Wang, L., Xun, H., Zhang, R., Wu, L., et al. (2022). Identification of a BAHD acyltransferase gene involved in plant growth and secondary metabolism in tea plants. *Plants* 11 (19), 2483. doi: 10.3390/plants11192483

Alexander, L. E., Okazaki, Y., Schelling, M. A., Davis, A., Zheng, X., Rizhsky, L., et al. (2020). Maize glossy2 and glossy2-like genes have overlapping and distinct functions in cuticular lipid deposition. *Plant Physiol.* 183, 840–853. doi: 10.1104/pp.20.00241

Bai, Y., Yang, C., Halitschke, R., Paetz, C., Kessler, D., Burkard, K., et al. (2022). Natural history-guided omics reveals plant defensive chemistry against leafhopper pests. *Science* 375, eabm2948. doi: 10.1126/science.abm2948

Bartley, L. E., Peck, M. L., Kim, S.-R., Ebert, B., Manisseri, C., Chiniquy, D. M., et al. (2013). Overexpression of a BAHD acyltransferase, osAt10, alters rice cell wall hydroxycinnamic acid content and saccharification. *Plant Physiol.* 161, 1615–1633. doi: 10.1104/pp.112.208694

Berger, A., Meinhard, J., and Petersen, M. (2006). Rosmarinic acid synthase is a new member of the superfamily of BAHD acyltransferases. *Planta* 224, 1503–1510. doi: 10.1007/s00425-006-0393-y

Bloor, S. J., and Abrahams, S. (2002). The structure of the major anthocyanin in Arabidopsis thaliana. *Phytochemistry* 59, 343–346. doi: 10.1016/S0031-9422(01)00460-5

Burhenne, K., Kristensen, B. K., and Rasmussen, S. K. (2003). A new class of N-hydroxycinnamoyltransferases. Purification, cloning, and expression of a barley agmatine coumaroyltransferase (EC 2.3.1.64). *J. Biol. Chem.* 278, 13919–13927. doi: 10.1074/jbc.M213041200

Campos, L., Lisón, P., López-Gresa, M. P., Rodrigo, I., Zacarés, L., Conejero, V., et al. (2014). Transgenic tomato plants overexpressing tyramine N-hydroxycinnamoyltransferase exhibit elevated hydroxycinnamic acid amide levels and enhanced resistance to Pseudomonas syringae. *Mol. Plant Microbe Interact.* 27, 1159–1169. doi: 10.1094/MPMI-04-14-0104-R

Chandrakanth, N. N., Zhang, C., Freeman, J., de Souza, W. R., Bartley, L. E., and Mitchell, R. A. C. (2023). Modification of plant cell walls with hydroxycinnamic acids by BAHD acyltransferases. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1088879

Chang, A., Hu, Z., Chen, B., Vanderschuren, H., Chen, M., Qu, Y., et al. (2022). Characterization of trichome-specific BAHD acyltransferases involved in acylsugar biosynthesis in Nicotiana tabacum. *J. Exp. Bot.* 73, 3913–3928. doi: 10.1093/jxb/erac095

Chiang, Y. C., Levsh, O., Lam, C. K., Weng, J. K., and Wang, Y. (2018). Structural and dynamic basis of substrate permissiveness in hydroxycinnamoyltransferase (HCT). *PloS Comput. Biol.* 14, e1006511. doi: 10.1371/journal.pcbi.1006511

Comino, C., Lanteri, S., Portis, E., Acquadro, A., Romani, A., Hehn, A., et al. (2007). Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in Cynara cardunculus L. *BMC Plant Biol.* .7, 14. doi: 10.1186/1471-2229-7-14

D'Auria, J. C. (2006). Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol.* 9, 331–340. doi: 10.1016/j.pbi.2006.03.016

De Luca, V., Balsevich, J., and Kurz, W. G. W. (1985). Acetyl coenzyme A: deacetylvindoline O-acetyltransferase, A novel enzyme from catharanthus. *J. Plant Physiol.* 121, 417–428. doi: 10.1016/S0176-1617(85)80078-X

de Oliveira, D. M., Finger-Teixeira, A., Mota, T. R., Salvador, V. H., Moreira-Vilar, F. C., Molinari, H. B., et al. (2015). Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. *J. Plant Biotechnol.* 13, 1224–1232. doi: 10.1111/pbi.12292

de Souza, W. R., Martins, P. K., Freeman, J., Pellny, T. K., Michaelson, L. V., Sampaio, B. L., et al. (2018). Suppression of a single BAHD gene in Setaria viridis causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. *New Phytol.* 218, 81–93. doi: 10.1111/nph.14970

de Souza, W. R., Pacheco, T. F., Duarte, K. E., Sampaio, B. L., de Oliveira Molinari, P. A., Martins, P. K., et al. (2019). Silencing of a BAHD acyltransferase in sugarcane increases biomass digestibility. *Biotechnol. Biofuels* 12, 111. doi: 10.1186/s13068-019-1450-7

Dobritzsch, M., Lübken, T., Eschen-Lippold, L., Gorzolka, K., Blum, E., Matern, A., et al. (2016). MATE transporter-dependent export of hydroxycinnamic acid amides. *Plant Cell* 28, 583–596. doi: 10.1105/tpc.15.00706

Dong, J., Wang, G., Zhong, C., Chang, L., Sun, J., Zhang, H., et al. (2018). Studying function of alcohol acyltransferase gene fvAATW2 of fragaria vesca by overexpressing in tobacco and cultivated strawberry. *J. Horticult.* 45 (01), 41–50. doi: 10.3390/IJMS23169500

Dong, D., Yang, Z., Ma, Y., Li, S., Wang, M., Li, Y., et al. (2022). Expression of a Hydroxycinnamoyl-CoA Shikimate/Quinate Hydroxycinnamoyl Transferase 4 Gene from Zoysia japonica (ZjHCT4) Causes Excessive Elongation and Lignin Composition Changes in Agrostis stolonifera. *Int. J. Mol. Sci.* 23(16), 9500. doi: 10.3390/ijms23169500

Dudareva, N., D'Auria, J. C., Nam, K. H., Raguso, R. A., and Pichersky, E. (1998). Acetyl-CoA:benzylalcohol acetyltransferase–an enzyme involved in floral scent production in Clarkia breweri. *Plant J. Cell Mol. Biol.* 14, 297–304. doi: 10.1046/j.1365-313X.1998.00121.x

Eudes, A., Mouille, M., Robinson, D. S., Benites, V. T., Wang, G., Roux, L., et al. (2016a). Exploiting members of the BAHD acyltransferase family to synthesize multiple hydroxycinnamate and benzoate conjugates in yeast. *Microb. Cell Factories* 15, 198. doi: 10.1186/s12934-016-0593-5

Eudes, A., Pereira, J. H., Yogiswara, S., Wang, G., Teixeira Benites, V., Baidoo, E. E., et al. (2016b). Exploiting the substrate promiscuity of hydroxycinnamoyl-coA: Shikimate hydroxycinnamoyl transferase to reduce lignin. *Plant Cell Physiol.* 57, 568–579. doi: 10.1093/pcp/pcw016

Fanelli, A., Rancour, D. M., Sullivan, M., Karlen, S. D., Ralph, J., Riaño-Pachón,, et al. (2021). Overexpression of a sugarcane BAHD acyltransferase alters hydroxycinnamate content in maize cell wall. *Front. Plant Sci.* 12, 626168. doi: 10.3389/fpls.2021.626168

Fang, H., Zhang, F., Zhang, C., Wang, D., Shen, S., He, F., et al. (2022). Function of hydroxycinnamoyl transferases for the biosynthesis of phenolamides in rice resistance to Magnaporthe oryzae. *J. Genet. Genomics* 49, 776–786. doi: 10.1016/j.jgg.2022.02.008

Fani, R., and Fondi, M. (2009). Origin and evolution of metabolic pathways. *Phys. Life Rev.* 6, 23–52. doi: 10.1016/j.plrev.2008.12.003

Florio, F. E., Gattolin, S., Toppino, L., Bassolino, L., Fibiani, M., Lo Scalzo, R., et al. (2021). A smelAAT acyltransferase variant causes a major difference in eggplant (*Solanum melongena* L.) peel anthocyanin composition. *Int. J. Mol. Sci.* 22, 9174. doi: 10.3390/ijms22179174

Garvey, G. S., McCormick, S. P., Alexander, N. J., and Rayment, I. (2009). Structural and functional characterization of TRI3 trichothecene 15-O-acetyltransferase from Fusarium sporotrichioides. *Protein Sci.* 18, 747–761. doi: 10.1002/pro.80

Garvey, G. S., McCormick, S. P., and Rayment, I. (2008). Structural and functional characterization of the TRI101 trichothecene 3-O-acetyltransferase from Fusarium sporotrichioides and Fusarium graminearum: kinetic insights to combating Fusarium head blight. *J. Biol. Chem.* 283, 1660–1669. doi: 10.1074/jbc.M705752200

Goulet, C., Kamiyoshihara, Y., Lam, N. B., Richard, T., Taylor, M. G., Tieman, D. M., et al. (2015). Divergence in the enzymatic activities of a tomato and Solanum pennellii alcohol acyltransferase impacts fruit volatile ester composition. *Mol. Plant* 8, 153–162. doi: 10.1016/j.molp.2014.11.007

Grienenberger, S., Besseau, P., Geoffroy, D., Debayle, D., Heintz, C., Lapierre, B., et al. (2009). A BAHD acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *Plant J.* 58, 246–259.

Grothe, T., Lenz, R., and Kutchan, T. M. (2001). Molecular characterization of the salutaridinol 7-O-acetyltransferase involved in morphine biosynthesis in opium poppy Papaver somniferum. *J. Biol. Chem.* 276, 30717–30723. doi: 10.1074/jbc.M102688200

Hampel, D., Mau, C. J., and Croteau, R. B. (2009). Taxol biosynthesis: Identification and characterization of two acetyl CoA:taxoid-O-acetyl transferases that divert pathway flux away from Taxol production. *Arch. Biochem. Biophys.* 487, 91–97. doi: 10.1016/j.abb.2009.05.018

Haslam, T. M., Gerelle, W. K., Graham, S. W., and Kunst, L. (2017). The unique role of the ECERIFERUM2-LIKE clade of the BAHD acyltransferase superfamily in cuticular wax metabolism. *Plants* 6, 23. doi: 10.3390/plants6020023

Haslam, T. M., Haslam, R., Thoraval, D., Pascal, S., Delude, C., Domergue, F., et al. (2015). ECERIFERUM2-LIKE proteins have unique biochemical and physiological functions in very-long-chain fatty acid elongation. *Plant Physiol.* 167, 682–692. doi: 10.1104/pp.114.253195

Haslam, T. M., and Kunst, L. (2020). Arabidopsis ECERIFERUM2-LIKEs are mediators of condensing enzyme function. *Plant Cell Physiol.* 61, 2126–2138. doi: 10.1093/pcp/pcaa133

Huang, X. Q., Li, R., Fu, J., and Dudareva, N. (2022). A peroxisomal heterodimeric enzyme is involved in benzaldehyde synthesis in plants. *Nat. Commun.* 13, 1352. doi: 10.1038/s41467-022-28978-2

Ji, X. H., Liu, F. Z., Wang, B. L., Liu, P. P., and Wang, H. B. (2022). Genetic variation of alcohol acyltransferase encoding gene in grape. *Sci. Agricult. Sin.* 55 (14), 2797–2811. doi: 10.3864/j.issn.0578-1752.2022.14.01

Kage, U., Karre, S., Kushalappa, A. C., and McCartney, C. (2017). Identification and characterization of a fusarium head blight resistance gene TaACT in wheat QTL-2DL. *Plant Biotechnol. J.* 15, 447–457. doi: 10.1111/pbi.12641

Kriegshauser, L., Knosp, S., Grienenberger, E., Tatsumi, K., Gütle, D. D., Sørensen, I., et al. (2021). Function of the HYDROXYCINNAMOYL-CoA:SHIKIMATE HYDROXYCINNAMOYL TRANSFERASE is evolutionarily conserved in embryophytes. *Plant Cell.* 33, 1472–1491. doi: 10.1093/plcell/koab044

Kruse, L. H., Weigle, A. T., Irfan, M., Martínez-Gómez, J., Chobirko, J. D., Schaffer, J. E., et al. (2022). Orthology-based analysis helps map evolutionary diversification and predict substrate class use of BAHD acyltransferases. *Plant J.* 111, 1453–1468. doi: 10.1111/tpj.15902

Kumar, A. (2020). Seasonal and developmental dynamics of the essential oil content versus catalytic activity of alcohol acyltransferase in foliage of citronella. *Russ. J. Plant Physiol.* 67, 369–377. doi: 10.1134/S1021443720020090

Kusano, H., Li, H., Minami, H., Kato, Y., Tabata, H., and Yazaki, K. (2019). Evolutionary developments in plant specialized metabolism, exemplified by two transferase families. *Front. Plant Sci.* 10, 794. doi: 10.3389/fpls.2019.00794

Lallemand, L. A., Zubieta, C., Lee, S. G., Wang, Y., Acajjaoui, S., Timmins, J., et al. (2012). A structural basis for the biosynthesis of the major chlorogenic acids found in coffee. *Plant Physiol.* 160, 249–260. doi: 10.1104/pp.112.202051

Levsh, O., Chiang, Y. C., Tung, C. F., Noel, J. P., Wang, Y., and Weng, J. K. (2016). Dynamic conformational states dictate selectivity toward the native substrate in a substrate-permissive acyltransferase. *Biochemistry* 55, 6314–6326. doi: 10.1021/acs.biochem.6b00887

Li, G., Jones, K. C., Eudes, A., Pidatala, V. R., Sun, J., Xu, F., et al. (2018). Overexpression of a rice BAHD acyltransferase gene in switchgrass (Panicum virgatum L.) enhances saccharification. *BMC Biotechnol.* 18, 54. doi: 10.1186/s12896-018-0464-8

Li, D., Xu, Y., Xu, G., Gu, L., Li, D., and Shu, H. (2006). Molecular cloning and expression of a gene encoding alcohol acyltransferase (MdAAT2) from apple (cv. Golden Delicious). *Phytochemistry* 67, 658–667. doi: 10.1016/j.phytochem.2006.01.027

Liu, X., Dai, S., Zhou, Y., Liu, J., Li, D., Zhang, J., et al. (2022b). Crystal structure of the plant feruloyl–coenzyme a monolignol transferase provides insights into the formation of monolignol ferulate conjugates. *Biochem. Biophys. Res. Commun.* 594, 8–14. doi: 10.1016/j.bbrc.2022.01.037

Liu, S., Jiang, J., Ma, Z., Xiao, M., Yang, L., Tian, B., et al. (2022a). The Role of hydroxycinnamic acid amide pathway in plant immunity. *Front. Plant Sci.* 13, 922119. doi: 10.3389/fpls.2022.922119

Liu, Q., Liu, Y., Xu, Y., Yao, L., Liu, Z., Cheng, H., et al. (2018). Overexpression of and RNA interference with hydroxycinnamoyl-CoA quinate hydroxycinnamoyl transferase affect the chlorogenic acid metabolic pathway and enhance salt tolerance in Taraxacum antungense Kitag. *Phytochem. Lett.* 28, 116–123. doi: 10.1016/j.phytol.2018.10.003

Liu, Y. Y., Mo, T., Wang, X. H., Shi, S. P., Liu, X., and Tu, P. F. (2016a). Research progress of plant BAHD acyltransferase family. *China J. Chin. Mater. Med.* 41, 2175–2182. doi: 10.4268/cjcmm20161201

Liu, C., Qiao, X., Li, Q., Zeng, W., Wei, S., Wang, X., et al. (2020). Genome-wide comparative analysis of the BAHD superfamily in seven Rosaceae species and expression analysis in pear (Pyrus bretschneideri). *BMC Plant Biol.* 20, 14. doi: 10.1186/s12870-019-2230-z

Liu, X., Zeng, X., Zheng, R., Luo, Q., and Wang, C. (2016b). Cloning and expression of the alcohol acyltransferase gene fromOsmanthus fragrans flowers. *J. Huazhong Agric. University* 35 (01), 36–42. doi: 10.13300/j.cnki.hnlkxb.2016.01.006

Luo, J., Nishiyama, Y., Fuell, C., Taguchi, G., Elliott, K., Hill, L., et al. (2007). Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from Arabidopsis thaliana. *Plant J.* 50, 678–695. doi: 10.1111/j.1365-313X.2007.03079.x

Ma, X., Koepke, J., Panjikar, S., Fritzsch, G., and Stöckigt, J. (2005). Crystal structure of vinorine synthase, the first representative of the BAHD superfamily*. *J. Biol. Chem.* 280, 13576–13583. doi: 10.1074/jbc.M414508200

Manjasetty, B. A., Yu, X. H., Panjikar, S., Taguchi, G., Chance, M. R., and Liu, C. J. (2012). Structural basis for modification of flavonol and naphthol glucoconjugates by Nicotiana tabacum malonyltransferase (NtMaT1). *Planta* 236, 781–793. doi: 10.1007/s00425-012-1660-8

McElroy, C., and Jennewein, S. (2018). Taxol® biosynthesis and production: from forests to fermenters. In: Schwab, W., Lange, B., and Wüst, M. (eds) *Biotechnology of Natural Products*. (Cham: Springer) 145–185. doi: 10.1007/978-3-319-67903-7_7

Milde, R., Schnabel, A., Ditfe, T., Hoehenwarter, W., Proksch, C., Westermann, B., et al. (2022). Chemical synthesis of trans 8-methyl-6-nonenoyl-coA and functional expression unravel capsaicin synthase activity encoded by the pun1 locus. *Molecules* 27, 6878. doi: 10.3390/molecules27206878

Moghe, G., Kruse, L. H., Petersen, M., Scossa, F., Fernie, A. R., Gaquerel, E., et al. (2022). BAHD company: the ever-expanding roles of the BAHD acyltransferase gene family in plants. *Annu. Rev. Plant Biol.* 74, 165–194. doi: 10.1146/annurev-arplant-062922-050122

Moghe, G. D., Leong, B. J., Hurney, S. M., Daniel Jones, A., and Last, R. L. (2017). Evolutionary routes to biochemical innovation revealed by integrative analysis of a plant-defense related specialized metabolic pathway. *Elife* 6, e28468. doi: 10.7554/eLife.28468.057

Moglia, A., Acquadro, A., Eljounaidi, K., Milani, A. M., Cagliero, C., Rubiolo, P., et al. (2016). Genome-wide identification of BAHD acyltransferases and *in vivo* characterization of HQT-like enzymes involved in caffeoylquinic acid synthesis in globe artichoke. *Front. Plant Sci.* 7, 1424. doi: 10.3389/fpls.2016.01424

Moglia, A., Lanteri, S., Comino, C., Hill, L., Knevitt, D., Cagliero, C., et al. (2014). Dual catalytic activity of hydroxycinnamoyl-coenzyme A quinate transferase from tomato allows it to moonlight in the synthesis of both mono- and dicaffeoylquinic acids. *Plant Physiol.* 166, 1777–1787. doi: 10.1104/pp.114.251371

Molina, I., and Kosma, D. (2015). Role of HXXXD-motif/BAHD acyltransferases in the biosynthesis of extracellular lipids. *Plant Cell Rep.* 34, 587–601. doi: 10.1007/s00299-014-1721-5

Mondal, A., Gandhi, A., Fimognari, C., Atanasov, A. G., and Bishayee, A. (2019). Alkaloids for cancer prevention and therapy: Current progress and future perspectives. *Eur. J. Pharmacol.* 858, 172472. doi: 10.1016/j.ejphar.2019.172472

Morales-Quintana, L., Moya-León, M. A., and Herrera, R. (2015). Computational study enlightens the structural role of the alcohol acyltransferase DFGWG motif. *J. Mol.* 21, 216. doi: 10.1007/s00894-015-2762-6

Morales-Quintana, L., Nuñez-Tobar, M. X., Moya-León, M. A., and Herrera, R. (2013). Molecular dynamics simulation and site-directed mutagenesis of alcohol acyltransferase: a proposed mechanism of catalysis. *J. Chem. Inf. Model.* 53, 2689–2700. doi: 10.1021/ci400409s

Murayama, K., Kato-Murayama, M., Sato, T., Hosaka, T., Ishiguro, K., Mizuno, T., et al. (2021). Anthocyanin 5,3'-aromatic acyltransferase from Gentiana triflora, a structural insight into biosynthesis of a blue anthocyanin. *Phytochemistry* 186, 112727. doi: 10.1016/j.phytochem.2021.112727

Nakayama, T., Suzuki, H., and Nishino, T. (2003). Anthocyanin acyltransferases: specificities, mechanism, phylogenetics, and applications. *J. Mol. Catalysis B.: Enzymatic.* 23, 117–132. doi: 10.1016/S1381-1177(03)00078-X

Navarro-Retamal, C., Gaete-Eastman, C., Herrera, R., Caballero, J., and Alzate-Morales, J. H. (2016). Structural and Affinity Determinants in the Interaction between Alcohol Acyltransferase from F. x ananassa and Several Alcohol Substrates: A Computational Study. *PloS One* 11, e0153057. doi: 10.1371/journal.pone.0153057

Okada, T., Hirai, M. Y., Suzuki, H., Yamazaki, M., and Saito, K. (2005). Molecular characterization of a novel quinolizidine alkaloid O-tigloyltransferase: cDNA cloning, catalytic activity of recombinant protein and expression analysis in lupinus plants. *Plant Cell Physiol.* 46, 233–244. doi: 10.1093/pcp/pci021

Onkokesung, N., Gaquerel, E., Kotkar, H., Kaur, H., Baldwin, I. T., and Galis, I. (2012). MYB8 controls inducible phenolamide levels by activating three novel hydroxycinnamoyl-coenzyme A:polyamine transferases in Nicotiana attenuata. *Plant Physiol.* 158, 389–407. doi: 10.1104/pp.111.187229

Peng, H., Yang, T., Whitaker, B. D., Trouth, F., Shangguan, L., Wen, D., et al. (2016). Characterization of spermidine hydroxycinnamoyl transferases from eggplant (*Solanum melongena* L.) and its wild relative Solanum richardii Dunal. *Hor. Res.* 3, 16062. doi: 10.1038/hortres.2016.62

Rinaldo, A. R., Cavallini, E., Jia, Y., Moss, S. M., McDavid, D. A., Hooper, L. C., et al. (2015). A grapevine anthocyanin acyltransferase, transcriptionally regulated by vvMYBA, can produce most acylated anthocyanins present in grape skins. *Plant Physiol.* 169, 1897–1916. doi: 10.1104/pp.15.01255

Rosa, N., and Neish, A. C. (1968). Formation and occurrence of N-malonylphenylalanine and related compounds in plants. *Can. J. Biochem.* 46, 799–806. doi: 10.1139/o68-121

Roumani, M., Besseau, S., Gagneul, D., Robin, C., and Larbat, R. (2021). Phenolamides in plants: an update on their function, regulation, and origin of their biosynthetic enzymes. *J. Exp. Bot.* 72, 2334–2355. doi: 10.1093/jxb/eraa582

Saigo, T., Wang, T., Watanabe, M., and Tohge, T. (2020). Diversity of anthocyanin and proanthocyanin biosynthesis in land plants. *Curr. Opin. Plant Biol.* 55, 93–99. doi: 10.1016/j.pbi.2020.04.001

Sander, M., and Petersen, M. (2011). Distinct substrate specificities and unusual substrate flexibilities of two hydroxycinnamoyltransferases, rosmarinic acid synthase and hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl-transferase, from Coleus blumei Benth. *Planta* 233, 1157–1171. doi: 10.1007/s00425-011-1367-2

Schilmiller, A. L., Charbonneau, A. L., and Last, R. L. (2012). Identification of a BAHD acetyltransferase that produces protective acyl sugars in tomato trichomes. *Proc. Natl. Acad. Sci. U. S. A.* 109, 16377–16382. doi: 10.1073/pnas.1207906109

Serrani-Yarce, J. C., Escamilla-Trevino, L., Barros, J., Gallego-Giraldo, L., Pu, Y., Ragauskas, A., et al. (2021). Targeting hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase for lignin modification in Brachypodium distachyon. *Biotechnol. Biofuels.* 14, 50. doi: 10.1186/s13068-021-01905-1

Shalit, M., Guterman, I., Volpin, H., Bar, E., Tamari, T., Menda, N., et al. (2003). Volatile ester formation in roses. Identification of an acetyl-coenzyme A. Geraniol/Citronellol acetyltransferase in developing rose petals. *Plant Physiol.* 131, 1868–1876. doi: 10.1104/pp.102.018572

Song, Z. Z., Peng, B., Gu, Z. X., Tang, M. L., Li, B., Liang, M. X., et al. (2021). Site-directed mutagenesis identified the key active site residues of alcohol acyltransferase PpAAT1 responsible for aroma biosynthesis in peach fruits. *Hor. Res.* 8, 32. doi: 10.1038/s41438-021-00461-x

Sullivan, M. L., and Knollenberg, B. J. (2021). Red clover HDT, a BAHD hydroxycinnamoyl-coenzyme A:L-3,4-dihydroxyphenylalanine (L-DOPA) hydroxycinnamoyl transferase that synthesizes clovamide and other N-hydroxycinnamoyl-aromatic amino acid amides. *Front. Plant Sci.* 12, 727461. doi: 10.3389/fpls.2021.727461

Sun, C. H., Yang, C. Y., and Tzen, J. T. C. (2018). Molecular identification and characterization of hydroxycinnamoyl transferase in tea plants (Camellia sinensis L.). *Int. J. Mol. Sci.* 19, 3938. doi: 10.3390/ijms19123938

Suzuki, H., Nakayama, T., Yamaguchi, M.-a., and Nishino, T. (2004a). cDNA cloning and characterization of two Dendranthema×morifolium anthocyanin malonyltransferases with different functional activities. *Plant Sci.* 166, 89–96. doi: 10.1016/j.plantsci.2003.08.010

Suzuki, H., Sawada, S., Watanabe, K., Nagae, S., Yamaguchi, M. A., Nakayama, T., et al. (2004b). Identification and characterization of a novel anthocyanin malonyltransferase from scarlet sage (Salvia splendens) flowers: an enzyme that is phylogenetically separated from other anthocyanin acyltransferases. *Plant J.* 38, 994–1003. doi: 10.1111/j.1365-313X.2004.02101.x

Taguchi, G., Shitchi, Y., Shirasawa, S., Yamamoto, H., and Hayashida, N. (2005). Molecular cloning, characterization, and downregulation of an acyltransferase that catalyzes the malonylation of flavonoid and naphthol glucosides in tobacco cells. *Plant J.* 42, 481–491. doi: 10.1111/j.1365-313X.2005.02387.x

Tang, W., and Thompson, W. A. (2022). Role of the rice BAHD acyltransferase gene osAt10 in plant cold stress tolerance. *Plant Mol. Biol. Rep.* 40, 482–499. doi: 10.1007/s11105-021-01328-0

Tian, Y., Lin, C. Y., Park, J. H., Wu, C. Y., Kakumanu, R., Pidatala, V. R., et al. (2021). Overexpression of the rice BAHD acyltransferase AT10 increases xylan-bound *p*-coumarate and reduces lignin in Sorghum bicolor. *Bio Biofuels.* 14, 217. doi: 10.1186/s13068-021-02068-9

Tohge, T., Wendenburg, R., Ishihara, H., Nakabayashi, R., Watanabe, M., Sulpice, R., et al. (2016). Characterization of a recently evolved flavonol-phenylacyltransferase gene provides signatures of natural light selection in Brassicaceae. *Nat. Commun.* 7, 12399. doi: 10.1038/ncomms12399

Tuominen, L. K., Johnson, V. E., and Tsai, C.-J. (2011). Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among Populus paralogues. *BMC Genom.* 12, 236. doi: 10.1186/1471-2164-12-236

Ube, N., Nishizaka, M., Ichiyanagi, T., Ueno, K., Taketa, S., and Ishihara, A. (2017). Evolutionary changes in defensive specialized metabolism in the genus Hordeum. *Phytochemistry* 141, 1–10. doi: 10.1016/j.phytochem.2017.05.004

Unno, H., Ichimaida, F., Suzuki, H., Takahashi, S., Tanaka, Y., Saito, A., et al. (2007). Structural and mutational studies of anthocyanin malonyltransferases establish the features of BAHD enzyme catalysis. *J. Biol. Chem.* 282, 15812–15822. doi: 10.1074/jbc.M700638200

Vanholme, R., De Meester, B., Ralph, J., and Boerjan, W. (2019). Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* 56, 230–239. doi: 10.1016/j.copbio.2019.02.018

Walker, A. M., Hayes, R. P., Youn, B., Vermerris, W., Sattler, S. E., and Kang, C. (2013). Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme a-dependent transferases and synthases. *Plant Physiol.* 162, 640–651. doi: 10.1104/pp.113.217836

Wang, L., Chen, K., Zhang, M., Ye, M., and Qiao, X. (2022). Catalytic function, mechanism, and application of plant acyltransferases. *Crit. Rev. Biotechnol.* 42, 125–144. doi: 10.1080/07388551.2021.1931015

Wang, T., Li, L., Zhuang, W., Zhang, F., Shu, X., Wang, N., et al. (2021a). Recent research progress in taxol biosynthetic pathway and acylation reactions mediated by taxus acyltransferases. *Molecules* 26, 2855. doi: 10.3390/molecules26102855

Wang, P. P., Liu, H., Gao, S., and Cheng, A. X. (2017). Functional characterization of a hydroxyacid/alcohol hydroxycinnamoyl transferase produced by the liverwort marchantia emarginata. *Molecules* 22, 1854. doi: 10.3390/molecules22111854

Wang, Y., Zhang, H., Wan, C., He, X., Huang, J., Lyu, M., et al. (2021b). Characterization of two BAHD acetyltransferases highly expressed in the flowers of jasminum sambac (L.) aiton. *Plants* 11, 13. doi: 10.3390/plants11010013

Wilkerson, C. G., Mansfield, S. D., Lu, F., Withers, S., Park, J.-Y., Karlen, S. D., et al. (2014). Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science* 344, 90–93. doi: 10.1126/science.1250161

Xiong, X., Gou, J., Liao, Q., Li, Y., Zhou, Q., Bi, G., et al. (2021). The Taxus genome provides insights into paclitaxel biosynthesis. *Nat. Plants.* 7, 1026–1036. doi: 10.1038/s41477-021-00963-5

Yamane, M., Takenoya, M., Yajima, S., and Sue, M. (2020). Crystal structure of barley agmatine coumaroyltransferase, an N-acyltransferase from the BAHD superfamily. *Acta Crystallogr. D.* 76, 590–596. doi: 10.1107/S2053230X20014880

Yamane, M., Takenoya, M., Yajima, S., and Sue, M. (2021). Molecular and structural characterization of agmatine coumaroyltransferase in Triticeae, the key regulator of hydroxycinnamic acid amide accumulation. *Phytochemistry* 189, 112825. doi: 10.1016/j.phytochem.2021.112825

Yan, X., Qin, X., Li, W., Liang, D., Qiao, J., and Li, Y. (2020). Functional characterization and catalytic activity improvement of BAHD acyltransferase from Celastrus angulatus Maxim. *Planta* 252, 6. doi: 10.1007/s00425-020-03413-2

Yu, X. H., Chen, M. H., and Liu, C. J. (2008). Nucleocytoplasmic-localized acyltransferases catalyze the malonylation of 7-O-glycosidic (iso)flavones in Medicago truncatula. *Plant J.* 55, 382–396. doi: 10.1111/j.1365-313X.2008.03509.x

Yu, X. H., Gou, J. Y., and Liu, C. J. (2009). BAHD superfamily of acyl-CoA dependent acyltransferases in Populus and Arabidopsis: bioinformatics and gene expression. *Mol. Plant Biol.* 70, 421–442. doi: 10.1007/s11103-009-9482-1

Yuan, Z., Yang, H., Pan, L., Zhao, W., Liang, L., Gatera, A., et al. (2022). Systematic identification and expression profiles of the BAHD superfamily acyltransferases in barley (Hordeum vulgare). *Sci. Rep.* 12, 5063. doi: 10.1038/s41598-022-08983-7

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

Zhang, Z., and Xu, L. (2018). Arabidopsis BRASSINOSTEROID INACTIVATOR2 is a typical BAHD acyltransferase involved in brassinosteroid homeostasis. *J. Exp. Bot.* 69, 1925–1941. doi: 10.1093/jxb/ery057

Zhao, D., and Tao, J. (2015). Recent advances on the development and regulation of flower color in ornamental plants. *Front. Plant Sci.* 261. doi: 10.3389/fpls.2015.00261

Zhao, Y., Yu, X., Lam, P. Y., Zhang, K., Tobimatsu, Y., and Liu, C. J. (2021). Monolignol acyltransferase for lignin *p*-hydroxybenzoylation in Populus. *Nat. Plants.* 7, 1288–1300. doi: 10.1038/s41477-021-00975-1

Zhou, W., Kong, W., Yang, C., Feng, R., and Xi, W. (2021). Alcohol acyltransferase is involved in the biosynthesis of C6 esters in apricot (Prunus Armeniaca L.) fruit. *Front. Plant Sci.* 12, 763139. doi: 10.3389/fpls.2021.763139

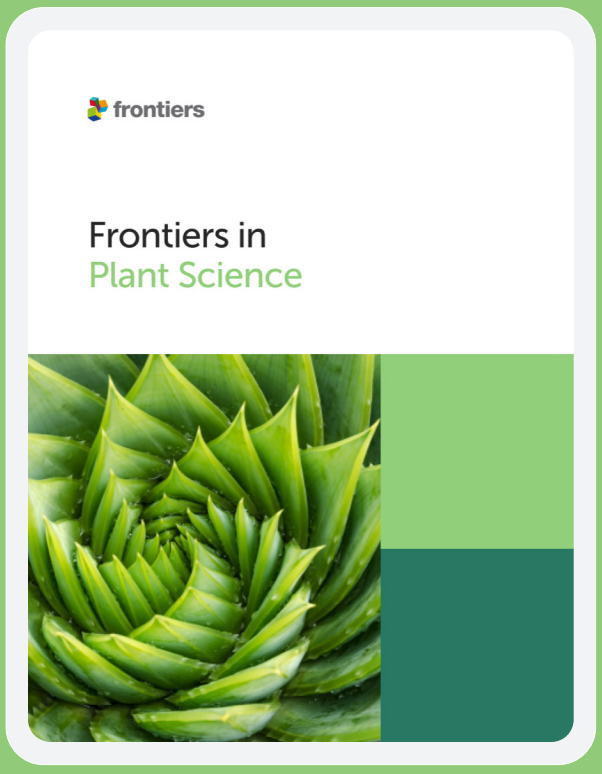
Frontiers in Plant Science

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