FLAVONOIDS: FROM BIOSYNTHESIS AND METABOLISM TO HEALTH BENEFITS

EDITED BY: M. Carmen González-Mas, M. Amparo Blazquez, María Pilar López-Gresa, Pedro Mena and Cristina Garcia-Viguera

PUBLISHED IN: Frontiers in Plant Science and Frontiers in Pharmacology

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88971-670-8
DOI 10.3389/978-2-88971-670-8

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

FLAVONOIDS: FROM BIOSYNTHESIS AND METABOLISM TO HEALTH BENEFITS

Topic Editors:
M. Carmen González-Mas, University of Valencia, Spain
M. Amparo Blazquez, University of Valencia, Spain
María Pilar López-Gresa, Universitat Politècnica de València, Spain
Pedro Mena, University of Parma, Italy
Cristina Garcia-Viguera, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Citation: González-Mas, M. C., Blazquez, M. A., López-Gresa, M. P., Mena, P., Garcia-Viguera, C., eds. (2021). Flavonoids: From Biosynthesis and Metabolism to Health Benefits. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-670-8

Table of Contents

frontiers in Plant Science

EDITORIAL
published: 28 September 2021
doi: 10.3389/fpls.2021.727043

Editorial: Flavonoids: From Biosynthesis and Metabolism to Health Benefits

M. Carmen González-Mas[1†], M. Amparo Blázquez[1], M. Pilar López-Gresa[2], Pedro Mena[3] and Cristina García-Viguera[4]*

[1] Department of Pharmacology, Faculty of Pharmacy, University of Valencia, Valencia, Spain, [2] Institute for Plant Molecular and Cellular Biology, Spanish National Research Council (CSIC)-Polytechnic University of Valencia, Valencia, Spain, [3] Human Nutrition Unit, Department of Food & Drug, University of Parma, Parma, Italy, [4] Phytochemistry and Healthy Foods Laboratory, Department of Food Science and Technology, Centro de Edafología y Biología Aplicada del Segura-Consejo Superior de Investigaciones Científicas (CEBAS–CSIC), Murcia, Spain

Keywords: anthocyanin content, isoflavones, catechin galloylation, pinostrobin pharmacokinetics, xanthohumol, hydroxysafflor Yellow A, vitexin, formononetin

Editorial on the Research Topic

Flavonoids: From Biosynthesis and Metabolism to Health Benefits

Flavonoids are a wide class of polyphenols with multiple biological activities. Unfortunately, many mechanisms of action that explain these activities are still unknown or there are insufficient data on their pharmacokinetics (**Figure 1**). Moreover, there are even undiscovered aspects about flavonoid biosynthesis. Expanding the knowledge on these topics will allow to regulate their biosynthesis, discover efficient preventive/therapeutic flavonoids for patients with certain diseases or design better clinical trials that help confirm their preventive/therapeutic usefulness. This Research Topic shows recent studies advancing along these lines.

Regarding flavonoid biosynthesis, this Research Topic presents a study about the catechin galloylation mechanism in tea plants. Serine carboxypeptidase-like acyltransferases (SCPL) are involved in this mechanism, but their roles have not been clarified so far. In this Research Topic, the tea genome-wide analysis of SCPL gene family is shown (Ahmad et al.), grouping several SCPLs into class IA (enzymes with an acylation function). Three SCPL-IA enzymes (CsSCPL11-IA, CsSCPL13-IA, CsSCPL14-IA) with galloylation activity toward catechins were assessed using recombinant enzymes. The expression levels of these SCPL-IA genes coincided with the galloylated catechin accumulation in tea plants, and their recombinant enzymes also displayed β-glucogallin:catechin galloyl acyltransferase activity. For the first time, genes encoding glucogallin:catechin galloyl acyltransferases with an active function in the galloylated catechin biosynthesis in tea plants have been identified. These results will help to understand the biosynthesis of these galloylated metabolites and the physiological roles of (–)-epicatechin-3-gallate, and (–)-epigallocatechin-3-gallate in tea plants.

The Research Topic also gives information about anthocyanin content in fruits and their biosynthesis regulation. Zeng et al. studied differences of sensory properties and proanthocyanidin and anthocyanin contents, among other phenolics, in rabbiteye blueberry "Brightwell" (*Vaccinium ashei* cv. "Brightwell") fruits. This research showed that many blueberry sensory aspects (bitterness, umami, and sweetness) were positively correlated with total flavonoid, total phenol, proanthocyanidin, and anthocyanin contents. Blueberry plants grown at high altitude locations presented a high content of these compounds along with high scores for sweetness. These results suggested cultivating blueberry at high altitude may produce fruit that not only possess pronounced

Edited and reviewed by:
Selena Ahmed, Montana State University, United States

****Correspondence:***
M. Carmen González-Mas
carmen.gonzalez-mas@uv.es

†Present address:
M. Carmen González-Mas, Department of Plant Biology, Faculty of Biological Sciences, University of Valencia, Valencia, Spain

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

Received: *17 June 2021*
Accepted: *07 September 2021*
Published: *28 September 2021*

Citation:
González-Mas MC, Blázquez MA, López-Gresa MP, Mena P and García-Viguera C (2021) Editorial: Flavonoids: From Biosynthesis and Metabolism to Health Benefits. Front. Plant Sci. 12:727043. doi: 10.3389/fpls.2021.727043

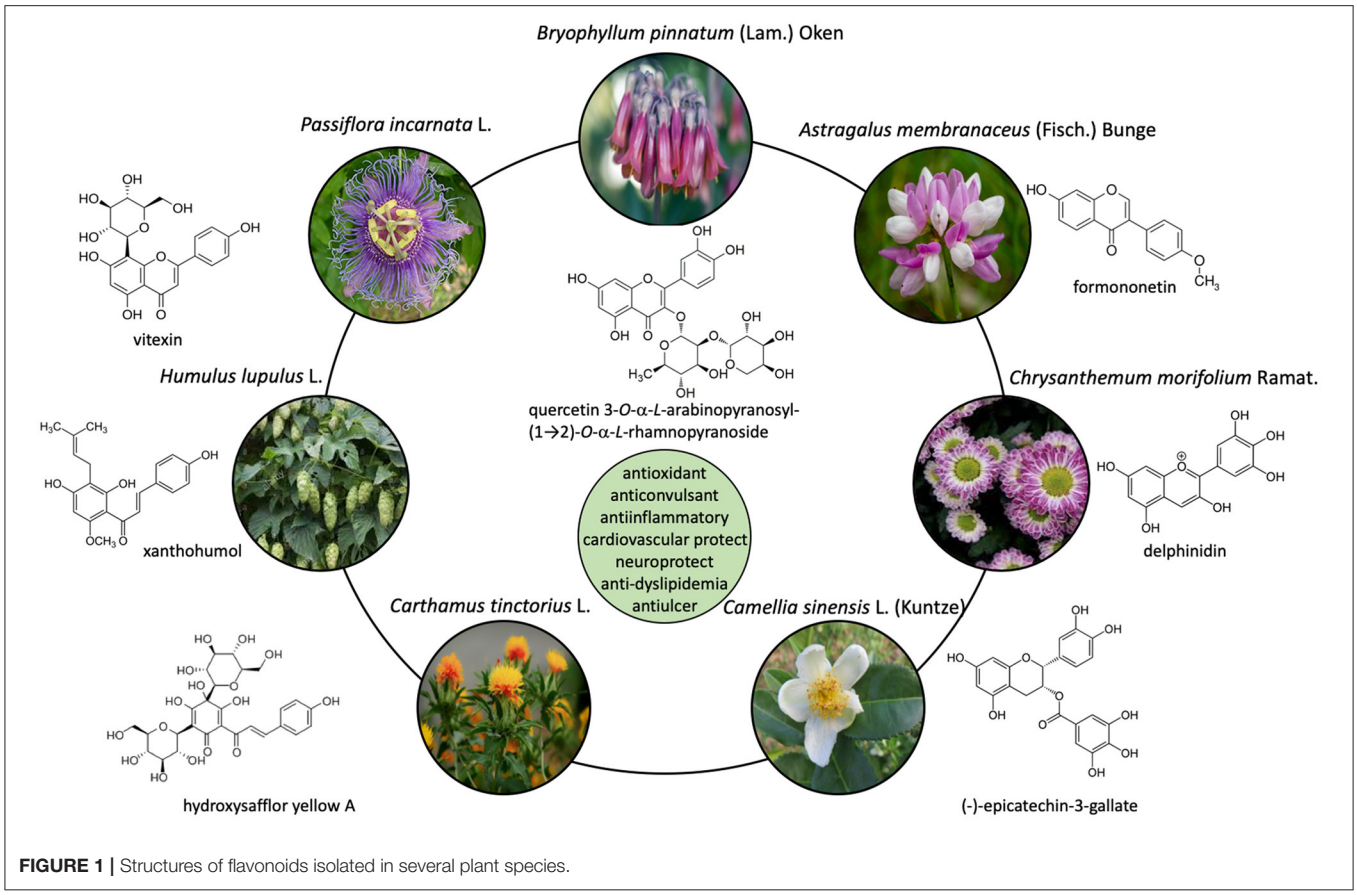

FIGURE 1 | Structures of flavonoids isolated in several plant species.

beneficial compounds but also good taste. In another work, Zhang B. et al. investigated the anthocyanin accumulation regulation in apple peel. This study focused on microRNAs, which participate in plant development by regulating gene expression at the posttranscriptional level. This research evidenced the increase expression level of the microRNA mdm-miR828 in the late apple fruit coloration stage, not in the rapid fruit coloration period, inhibiting anthocyanin synthesis in apple. Thus, mdm-miR828 is involved in a feedback regulatory mechanism associated with anthocyanin biosynthesis. In addition, it was also detected that mdm-miR828 inhibits anthocyanin accumulation in response to high temperature. Also, Li et al. communicated a transcriptome and flavonoid metabolic study in black and white fruits of *Lycium ruthenicum* Murray. Seven transcription factors (tf) associated with anthocyanin biosynthesis were identified. One of them, *LrAN1b*, a basic helix-loop-helix (bHLH) tf, showed the greatest correlations with anthocyanin accumulation, with no expression in white fruits. Moreover, a new activated anthocyanin MYB tf, *LrAN2-like*, was identified. From these results, it was established that *LrAN1b* and *LrAN2-like* can interact on anthocyanin accumulation in these fruits. It explained the color differences of these fruits, speculating that the white fruit phenotype is due to *LrAN1b* abnormal expression. Furthermore, it was studied how light affects anthocyanin accumulation in *Dendrobium candidum* Wall. ex Lindl. stems, used in traditional medicine and in functional foods (Jia et al.). The results determined that strong light increased anthocyanin content and the expression of genes involved in anthocyanin biosynthesis, upregulating *DcTTG1*, a WD40-repeat tf involved in the expression levels of some key anthocyanin biosynthesis genes (*DcCHS2*, *DcCHI*, *DcF3H*, and *DcF3′H*). It was also found that *DcTTG1* restored the *Arabidopsis* ttg1-13 mutant phenotype, which is defect in seed coat anthocyanin pigmentation, suggesting a similar function as *Arabidopsis TTG1*. These results will help improve the understanding of the light-induced anthocyanin synthesis and accumulation mechanisms. Moreover, Lim et al. managed to find out why some white-flowered chrysanthemum (*Chrysanthemum morifolium* Ramat.) cultivars produce red rayflorets under natural cultivation conditions, comparing the expression of anthocyanin biosynthetic and tf genes between white ray florets and those that turned red based on cultivation conditions. Significant differences in the bHLH tf gene *CmbHLH2* expression were detected. *CmbHLH2* generated two alternatively spliced transcripts, $CmbHLH2^{Full}$ and $CmbHLH2^{Short}$. This second transcript encoded a truncated protein $CmbHLH2^{Short}$ that failed to promote anthocyanin biosynthetic genes not being able to interact with CmMYB6, whereas $CmbHLH2^{Full}$ promoted it when simultaneously expressed with CmMYB6. These results represent an important advance to modulate anthocyanin biosynthesis in chrysanthemum flowers.

Beyond anthocyanins, Sohn et al. provided an exhaustive review on the metabolic engineering process of isoflavone

biosynthetic genes in soybean. This study also included a summary about functional roles and health benefits of isoflavones, such as antiosteoporotic properties. In addition, Zhang X. et al. provided remarkable information on the biosynthesis of isoflavone glycosides. Aglycone isoflavones have low water solubility, which limits their therapeutic use; so that, the *O*-glycosylation of isoflavones, catalyzed by *O*-glycosyltransferases (UGTs), would be a way to overcome this problem. So far, studies on isoflavonoid *O*-UGTs were focused on legumes, but this last study is centered on *Iris domestica* (L.) Goldblatt et Mabberley, a non-legume rich in isoflavonoid glycosides. A comparative transcriptome analysis was performed using *I. domestica* seedlings treated with $CuCl_2$. Eight new active BcUGTs with broad substrate spectra were obtained. Real-time quantitative PCR results indicated that the transcriptional levels of BcUGTs were induced by Cu^{2+}. The different expression levels of BcUGTs in this transcriptome analysis seems to denote that BcUGTs play important roles in responses to abiotic stresses.

When it comes to the absorption, distribution, metabolism and excretion of flavonoids, Sun et al. provided pharmacokinetic information on pinostrobin, an aglycone flavanone found in more than 10 families such as Pinaceae. After its oral administration to male rats, concentrations of pinostrobin and its metabolites in plasma, urine, feces, bile, and tissue homogenates were determined by using ultra-high-performance liquid chromatography coupled with mass spectrometry. These results showed that it was mostly distributed in the gastrointestinal tract, which would explain its antiulcer property. Nevertheless, this compound was widely metabolized *in vivo* and it was able to reach all the organs assessed. This study provided a significant basis for further experiments on the potential biological activity of this compounds.

Regarding pharmacological activities, Lourenço et al. identified a mechanism of action of quercetin 3-*O*-α-*L*-arabinopyranosyl-(1→2)-*O*-α-*L*-rhamnopyranoside as anti-inflammatory compound, by a target *in silico* fishing method. This study also explained the *in vitro* anti-inflammatory activity of this flavonol and that of *Bryophyllum pinnatum* (Lam.) Oken, a flavonol-rich plant commonly used in Madagascar to treat inflammatory some processes. Thus, it was established that this flavonol is a potent PDE4B blocker, highly selective to this enzyme.

Concerning other flavonoid-related compounds activities, this Research Topic also shows the role of xanthohumol (XN), a hop prenylated chalcone, on a high-fat diet (Paraiso et al.). XN supplementation resulted in amelioration of hepatic steatosis and decreased bile acid (BA) concentrations in mice with liver farnesoid X receptor (FXR) deficiency, reducing inflammation and tissue damage. This effect was stronger in male mice. XN induces the androstane, pregnane X and glucocorticoid receptors gene expression in the liver of $FXR^{Liver-/-}$ mice. These data indicated sex-dependent relationship between FXR, lipids and BAs, and that XN improves high-fat diet-induced dysfunctional lipid and BA metabolism via FXR-dependent and independent signaling. Also, the Research Topic includes a wide review recapitulating the mechanisms of action and effects of hydroxysafflor Yellow A (HSYA), a quinochalcone *C*-glycoside extracted from *Carthamus tinctorius* L. flowers, on numerous cardio-cerebrovascular diseases (CCVDs) by considering preclinical studies (Bai et al.). Authors emphasized that, in the future, it will be necessary to perform high-quality clinical trials to support the HSYA application for the control of other CCVDs, and not only of acute ischemic stroke with blood stasis syndrome as up to now.

This Research Topic also shows the mechanism of action of the anticonvulsant glycoside flavone vitexin (apigenin-8-*C*-glucoside), found at plants such as *Passiflora incarnata L.* This compound seems to act mainly via modulation of GABAergic neurotransmission, exhibiting selective protection against tonic-clonic seizures triggered by GABA antagonist (Dias de Oliveira et al.). In addition, experiments carried out in a mouse model demonstrated that the aglycone isoflavone formononetin, the main active component of *Astragalus membranaceus*, improves diabetic renal fibrosis. It inhibited hyperglycemia-induced superoxide overproduction by activating the nuclear factor E2-related factor 2/antioxidant response element signaling pathway and increasing sirtuin-1 protein levels in renal tissue (Zhuang et al.).

Finally, Aneklaphakij et al. contributed with a valuable review about the chemical diversity of seed (poly)phenols in majorly consumed nut species coupled to insights into their biological activities. Furthermore, these authors showed how they used omics-based approaches on nut plant species, presenting an example of the annotation of key genes involved in (poly)phenolic biosynthesis in peanut using comparative genomics.

We hope that the studies shown in this Research Topic contribute to improving the knowledge about flavonoid biosynthesis and their roles in plants, as well as enhancing its use for the prevention/treatment of certain diseases.

AUTHOR CONTRIBUTIONS

MG-M participated in the writing of the article and coordinated the work of all the co-authors and designed the figure. MB, CG-V, and PM participated in the manuscript writing. ML-G made the figure and also participated in the writing of the manuscript. All authors contributed to the article and approved the submitted version.

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.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 22 January 2020
doi: 10.3389/fphar.2019.01582

Identification of a Selective PDE4B Inhibitor From *Bryophyllum pinnatum* by Target Fishing Study and *In Vitro* Evaluation of Quercetin 3-*O*-α-*L*-Arabinopyranosyl-(1→2)-*O*-α-*L*-Rhamnopyranoside

Estela M. G. Lourenço[1], *Júlia M. Fernandes*[2], *Vinícius de F. Carvalho*[3], *Raphael Grougnet*[4], *Marco A. Martins*[3], *Alessandro K. Jordão*[1], *Silvana M. Zucolotto*[2] *and Euzébio G. Barbosa*[1]*

[1] Laboratório de Química Farmacêutica Computacional, Departamento de Farmácia, Universidade Federal do Rio Grande do Norte, Natal, Brazil, [2] Laboratório de Produtos Naturais Bioativos, Departamento de Farmácia, Universidade Federal do Rio Grande do Norte, Natal, Brazil, [3] Laboratório de Inflamação, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil, [4] Laboratoire de Pharmacognosie, Faculté de Pharmacie, Université Paris Descartes, Paris, France

OPEN ACCESS

Edited by:
Pedro Mena,
University of Parma, Italy

Reviewed by:
Gerard Pujadas,
Rovira i Virgili University, Spain
Ana Paula Simões-Wüst,
University Hospital Zürich, Switzerland

***Correspondence:**
Euzébio G. Barbosa
euzebiogb@imd.ufrn.br

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: *26 August 2019*
Accepted: *05 December 2019*
Published: *22 January 2020*

Citation:
Lourenço EMG, Fernandes JM, Carvalho VdF, Grougnet R, Martins MA, Jordão AK, Zucolotto SM and Barbosa EG (2020) Identification of a Selective PDE4B Inhibitor From Bryophyllum pinnatum by Target Fishing Study and In Vitro Evaluation of Quercetin 3-O-α-L-Arabinopyranosyl-(1→2)-O-α-L-Rhamnopyranoside. Front. Pharmacol. 10:1582. doi: 10.3389/fphar.2019.01582

Natural products are considered an important source of bioactive compounds especially in biodiversity-rich countries like Brazil. The identification of potential targets is crucial to the development of drugs from natural sources. In this context, *in silico* methodologies, such as inverse virtual screening (target fishing), are interesting tools as they are a rational and direct method that reduces costs and experimental time. Among the species of Brazilian biomes, *Bryophyllum pinnatum* (Lam.) Oken, native to Madagascar, is widely used by the population to treat inflammation conditions. It has a remarkable presence of flavonoids, including quercetin 3-*O*-α-*L*-arabinopyranosyl-(1→2)-*O*-α-*L*-rhamnopyranoside (**1**), considered one of its major compounds. However, until now there were no studies addressing its putative mechanism of action and explaining its pharmacological action. The enzyme PDE4B, known as an antiinflammatory protein, was indicated as a promising target by target fishing methods. This activity was confirmed by *in vitro* enzymatic inhibition, and an expressive selectivity of PDE4B over PDE4A was demonstrated. The interactions were investigated through molecular dynamics simulations. The results were pioneering, representing an advance in the investigation of the antiinflammatory action of *B. pinnatum* and confirm the potential of the flavonoid as a chemical extract marker. Also, the flavonoid was shown to be a promising lead for the design of other selective PDE4B blockers to treat inflammatory diseases.

Keywords: *Bryophyllum pinnatum*, natural products, PDE4, Flavonoid, inverse virtual screening, molecular dynamics

INTRODUCTION

Historically, natural products and their derivatives have a remarkable importance in the process of discovery and development of new drugs, being responsible for several of the recently approved new chemical entities (Patridge et al., 2016) and for about 50% of all drugs approved in clinical use (Horbal et al., 2018). The identification of the mechanism of action for these natural products may be a great tool for modern drug design (Xu et al., 2018). However, the search for targets still follows the traditional methods with the execution of direct biological analyses, contributing to the known challenges that involve the process of drug discovery. In this sense, the utilization of *in silico* techniques has been increasing over the last decades (Freires et al., 2017). These methods are considered rational, reduce costs, experimental time, and number of biological models. It has been important in target identification and discovery of novel potential drugs (Wadood et al., 2013).

Among the *in silico* methodologies, virtual screening (VS) represents one of the biggest advances in drug design (Rodrigues et al., 2012). This method had become a primary component in drug development, often combined with homology modeling, QSAR studies, molecular dynamics simulations (Pérez et al., 2016) and *in vitro* assays (Espargaró et al., 2017). The VS approach, divided into structure and ligand-based, uses large libraries with the objective to identify a compound that is capable of binding to a molecular target, generally a protein or enzyme (Rester, 2008). There are several successful studies that have utilized VS to identify promising ligands and targets, including researches involving secondary metabolites (Xu et al., 2018). Among its advantages, the application of this tool in the target identification of natural products can reduce the number of biological assays and consequently, the amount of isolated compounds used in this process.

The Brazilian biomes are considered a large source of bioactive compounds, hosting 15%–20% of all global biodiversity (Barreiro and Bolzani, 2009). Among the plants of the Brazilian biomes, *Bryophyllum pinnatum* (Lam.) Oken (*B. pinnatum*), native to Madagascar, is important in the therapeutic context since it is widely used to treat inflammation conditions (Panyaphu et al., 2011; Ezuruike and Prieto, 2014; Sreekeesoon and Mahomoodally, 2014). Botanic synonyms as *Bryophyllum calycinum* and *Kalanchoe pinnata* can be found (Fernandes et al., 2019). This species is rich in phenolic compounds, particularly flavonoids (Nascimento et al., 2015). One of its major compounds is a glycosylated quercetin derivative identified as quercetin 3-*O*-α-*L*-arabinopyranosyl-(1→2)-*O*-α-*L*-rhamnopyranoside (**1**) (**Figure 1**) (Muzitano et al., 2006; Fernandes et al., 2016).

This compound, first derivative isolated from *B. pinnatum*, was found in higher proportions than the other secondary metabolites by HPLC analysis of the extract. Interestingly, previous studies had shown that the increase of compound **1,** caused by the different conditions, can be related to the improvement of the effects, as antioxidant properties (Nascimento et al., 2013). Other activities of compound **1** were also been discussed (Muzitano et al., 2006; Sobreira et al., 2017).

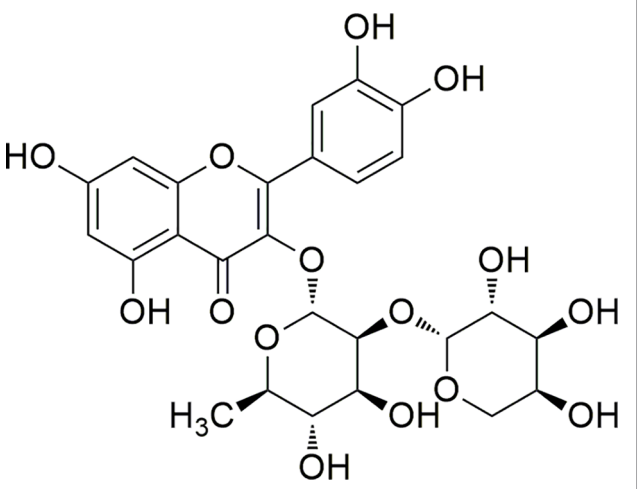

FIGURE 1 | Structure of compound 1.

In addition, the 3-*O* diglycosidic bond and the dimer rhamnose–arabinose are not common structure features and this flavonoid was described for few other species from different botanical families, however not as their major compound (Fernandes et al., 2019). Therefore, the compound **1** may be considered a promising chemical marker of the extract of *B. pinnatum,* making it also important to the conduction of deeper studies aiming to investigate its biological activities (dos Santos Nascimento et al., 2018). However, it is unknown the mechanism of action and/or some probable targets. In addition, this information represents a crucial factor in the discovery and development of new bioactive compounds based on its structure (Xu et al., 2018).

Thus, the aim of this study was to investigate the potential mechanism of action for compound **1** using *in silico* methodologies, known as target fishing (TF), and molecular dynamics simulations, combined with *in vitro* enzymatic inhibition. This investigation can assist in the selection of compound **1** as a possible chemical extract marker and also to the discovery of this flavonoid as a potential lead compound.

MATERIALS AND METHODS

Plant Material

The leaves of *B. pinnatum* were collected in November 2016 from the Escola Agrícola de Jundiaí (EAJ-UFRN), Macaíba city, Rio Grande do Norte state, Brazil, at the following coordinates: 5° 51' 30" S and 5° 51' 30" W. A vouch specimen (UFC. n° 57335) was deposited at the Herbario Prisco Bezerra (EAC) after its botanic identification by Dr. Rúbia Santos Fonseca. The plant material collection was made under the approval of Brazilian Authorization and Biodiversity Information System (SISBIO), process number 35017. Furthermore, the research was registered in the National System of Management of the Genetic Patrimony

and Associated Traditional Knowledge-SISGEN under number A7EA798.

Extraction and Isolation of Compound 1

The fresh leaves without stems (1.6 kg) were subjected to a turbo extraction for 5 min in an industrial blender, using EtOH: H_2O (1:1, v/v) at a plant: solvent proportion of 1:1 (w/v). Subsequently, the hydroethanolic extract (HE) was filtered and the volume was concentrated by rotaevaporation (model v-700, Buchi®). With the HE, a liquid-liquid extraction was made with three solvents of different degrees of polarity. Through this process, the dichloromethane (CH_2Cl_2), ethyl acetate (AcOEt), and *n*-butanol (*n*-BuOH) fractions were obtained.

For the isolation of compound **1**, the *n*-BuOH fraction (2 g) was selected and submitted into a column chromatography (silica gel 60) with gradient elution (30%–0% CH_2Cl_2: 70%–100% AcOEt and 95%–0% AcOEt: 5%–100% MeOH) resulting in nine fractions (G1–G9). The fractions G3 (520.7 mg) and G4 (372.5 mg) were purified separately through a process of size-exclusion column chromatography using Sephadex LH-20 as stationary phase and MeOH as solvent. The process of purification ended with five subfractions (F1–F5) for both primary fractions, G3 and G4. Between them, the subfractions F4 (from G3) and F5 (from G4) showed the same chemical profile by TLC, characterized by the presence of just one yellow band. These subfractions were reunited and submitted to a final purification process through preparative high-performance liquid chromatography (HPLC-UV) (Shimadzu®) using an isocratic elution (80% Milli-Q water and 20% Acetonitrile), a semipreparative column C18 Luna Phenomenex® (250 mm × 10 mm, 5 μm), flow of 4 ml/min and analysis time of 25 min. All fractions obtained during the entire purification process, were monitored by thin layer chromatography (TLC) (Merk silica gel 60) and the bands were visualized by UV (254 and 365 nm). After the purification, the isolation of compound **1** (34.8 mg) was confirmed by UPLC-DAD system used was a Shimadzu Model LC-20AD, with DAD detector SPD-M20A and software LabSolutions. Then, a Phenomenex Kinetex RP-18 column (150 mm × 4.6 mm, 2.6 μm particle size) equipped with a Phenomenex security guard column (4.0 mm × 2.0 mm i.d.) was used. The eluents were: (A) trifluoroacetic acid (TFA) 0.3% and (B) acetonitrile. The following gradient (v/v) was applied: 7%–15% B, 0–3 min; 15%–20% B, 3–12 min; 20%–22% B, 12–30 min; 30 min total analysis time. Flow elution was 0.7 ml/min, and 12 μl of each sample was injected. The UV-DAD detector was programmed to wavelength 200–500 nm and the chromatograms were plotted at 254 and 340 nm. The samples were resuspended in methanol: water, 1:1 (v/v) and the final concentration was 2 mg/ml for the extracts. HPLC-grade acetonitrile and trifluoracetic acid (TFA) were provided from J. T. Baker (Brazil). Water was purified with a Milli-Q system (Millipore). Samples and solvents were filtrated through a membrane (pore-size 0.45 μm) and degassed. The analyses were performed in triplicate. The UPLC-DAD analysis confirmed a purity percentage of approximately 93% (**Figure S1**, **Supplementary Material**). The detailed conditions of all column chromatography are described in **Table S1**, in **Supplementary Material**.

Structural Characterization of Compound 1

Quercetin 3-*O*-α-*L*-arabinopyranosyl-(1→2)-*O*-α-*L*-rhamnopyranoside (1): amorphous yellow solid; ^{1}H NMR (DMSO-d_6, 500 MHz) δ 7.32 (1H, s, H-2'), 7.25 (1H, d, J = 8.30 Hz, H-6'), 6.87 (1H, d, J = 8.25 Hz, H-5'), 6.38 (1H, s, H-8), 6.18 (1H, s, H-6), 5.29 (1H, s, H-1''), 4.08 (1H, d, J = 7.15 Hz, H-1'''), 4.00 (1H, s, H-2''), 3.21 (1H, d, J = 12.0 Hz, H-5'''), 3.12 (1H, t, J = 9.50 Hz, H-2'''), 3.12 (1H, t, J = 9.50 Hz, H-3'''), 0.90 (1H, d, J = 6.10 Hz, H-6''); ^{13}C NMR (DMSO-d_6, 125 MHz) δ 177.92 (C, C-4), 164.48 (C, C-7), 161.40 (C, C-9), 157.14 (C, C-5), 156.56 (C, C-2), 148.78 (C, C-4'), 145.38 (C, C-3'), 134.44 (C, C-3), 121.03 (CH, C-6'), 120.61 (C, C-1'), 115.73 (CH, C-2'), 115.54 (CH, C-5'), 106.5 (CH, C-1'''), 104.2 (C, C-10), 101.05 (CH, C-1''), 98.88 (CH, C-6), 93.84 (CH, C-8), 80.68 (CH, C-2''), 72.56 (CH, C-4''), 71.88 (CH, C-3'''), 71.14 (CH, C-2'''), 70.41 (CH, C-3''), 70.41 (CH, C-5''), 67.90 (CH, C-4'''), 65.93 (CH_2, C-5'''), 17.56 (CH_3, C-6''').

The NMR spectra are described in **Supplementary Material**.

TF Strategies

The 3D structure model of compound **1** was built using the program MarvinSketch 16.9.5 (ChemAxon Ltd.). The program AutoDockVINA (Olson and Trott, 2010) was used to obtain the required files for the inversed VS search. The target library used for structure-based inverse VS was constructed from the RCSB PDB Protein Data Bank (Bernstein et al., 1978) and was comprised of more than 9,000 structures with bound ligands. The TF was performed by molecular docking simulations using AutoDockVINA (Olson and Trott, 2010) automated *ad hoc* by shell scripting. The grid box used for the search was big enough to contain roughly the binding site of the targets and the original ligands were deleted before each simulation. The lowest free energy from the empirical score was used to rank the results. The best 25 results of each compound were selected for manual evaluation.

For the search of targets for compound **1**, it was necessary to screen a simplified moiety (quercetin) as ligand due to the exaggerated number of rotatable bonds present on this flavonoid. The best 50 results were selected for visual evaluation to ensure if the size of the active site of the targets was sufficiently large to accommodate the glycosides bound to compound **1** as a whole. The most promising targets were then selected for the docking with the complete structure of the glycosylated flavonoid.

The compound **1** was also submitted to a ligand-based VS using the web tool SwissSimilarity (Zoete et al., 2016), developed by SIBS Swiss Institute of Bioinformatics, to make a rapid ligand-based VS of small to unprecedented ultralarge libraries of small molecules, including the screening of drugs and bioactive and commercial molecules. With this web tool, the prediction can be carried out using six different approaches, such as 2D molecular fingerprints and super positional and fast nonsuper positional

3D similarity molecules (Zoete et al., 2016). In this study, the combined method was selected for screening and the library of bioactive ligands was used for comparison with the isolated compound constructed from the RCSB PDB Protein Data Bank server (Bernstein et al., 1978). The results were filtered with the aim to consider only similarity coefficients above 0.500 for each ligand of this work.

PDE4 Activity *In Vitro* Evaluation

Human PDE4A and PDE4B activities were measured using an IMAP TR-FRET protocol (kit from Molecular Devices, Sunnyvale, CA, USA) according to the instructions of the manufacturer. Briefly, the enzymatic reactions were carried out at room temperature in a 96-well black plate by co-incubating 100 nM FAM-cAMP (R7513), 10 μM putative inhibitory compounds and 4 ng PDE4A or 10 ng PDE4B isoform dissolved in assay buffer (R7364) for 1 h. The enzymes were obtained from human recombinant sources (MDS PHARMA), whereas the other reagents were purchased from Molecular Devices. Fluorescence polarization intensity was measured at 485 nm excitation and 520 nm emission using a microplate reader, SpectraMax M5 (Molecular Devices, Sunnyvale, CA, USA). Roflumilast was dissolved in dimethyl sulfoxide (DMSO) at a final concentration of 0.1%. The vehicle had no significant effect on PDE4 activity in this condition.

Molecular Dynamics Simulations and Binding-Free Energy Estimations

The simulations were performed using GROMACS simulation version 5 (Van Der Spoel et al., 2005) and CHARMM force field (Vanommeslaeghe et al., 2010). The solvent properties were mimetic using TIP3P water model with a cubic box large enough to allow a minimum of 1.0 nm space from the protein to the box walls. The system charge was neutralized with the addition of ions at the physiological concentration of 0.15 mM. Geometry optimization of the solvated system was performed using the steepest descent algorithm, followed by equilibration simulations with nVT and nPT ensembles keeping the inhibitor and the protein restrained. The temperature was kept at 300°K coupling the system to a V-rescale thermostat (0.1 ps), while the pressure was also kept constant at 1 bar using the Parinello-Rahman coupling algorithm. With the exception of the octahedral configuration structural metals, an unrestrained molecular dynamics simulation was performed until RMSD stabilization. The short range Columbic and Lennard-Jones interaction energies between compound **1** and the surroundings were monitored during the course of the productive simulation step.

To increase the study sample, the ligand binding poses for both isoforms were made based on the previous data found in RCSB PDB Protein Data Bank (Bernstein et al., 1978). The files of all the crystals of PDE4A and PDE4B were used for the alignment between each crystallized ligand and compound **1**. All binding poses resulting from the alignment were filtered to delete the repetitions. To enable the direct comparison between the results of the two isoforms, the exact same poses were used for the molecular dynamics simulations on PDE4A and PDE4B.

After the stabilization of each position, the binding-free energy values were calculated using g_mmpbsa package (Kumari et al., 2014). Between the calculated energy components, the E_{MM} was based on the LJ and Coulomb potential for each complex. The G_{polar} was calculated according to the package and for the $G_{nonpolar}$, the solvent accessible volume (SAV) was chosen as the type of nonpolar solvation model. Considering the results of MM/PBSA, the positions with the most promising binding-free energy were selected to extend the time of simulation, until 14 ns, in order to observe for a longer period the behavior of the complex. The simulations followed the same parameters described above.

RESULTS AND DISCUSSION

Isolation and Structure Elucidation of Compound 1

Compound **1** was obtained as an amorphous yellow solid and its structure was elucidated by comparison of its observed and reported physical data (Muzitano et al., 2006). ^{1}H NMR, ^{13}C NMR, HSQC, and HMBC spectra are available as **Figures S2–S5**, in the **Supplementary Material**. The ^{1}H NMR spectra showed characteristic aglycone signals between 7.32 and 6.18 ppm. The signals in the region of 6.20–6.40 ppm are related with the hydrogens of ring A of the flavonoid. In relation to ring B, the signals are compatible with an ortho coupling. Thus, the pattern of signals and multiplicity observed are compatible with the flavonoid quercetin (3,5,7,3',4'-pentahydroxyflavonol). Also, the spectra had demonstrated that the compound is *O*-glycosylated at the position C-3 of the ring C.

In addition, the data of ^{1}H and ^{13}C NMR spectra suggested the presence of two sugar units. In the ^{1}H NMR, the signal at 0.90 and 5.29 ppm suggested the presence of a unit of rhamnose. This sugar unit was confirmed by ^{13}C NMR spectra and also by the both two-dimensional spectra. The other unit of sugar is evidenced in the ^{1}H NMR spectra at 3.10–4.15 ppm. This signals suggested that the second sugar unit is an arabinose attached to the rhamnose unit. In conclusion, it was possible to confirm the compound **1** as quercetin 3-*O*-α-*L*-arabinopyranosyl-(1→2)-*O*-α-*L*-rhamnopyranoside (**Table 1**).

TF Strategies

The filtered structure-based inverse VS results showed that quercetin appears to be more interactive with proteins of *Homo sapiens* (**Table S2**, **Supplementary Material**). In respect to the therapeutic activities, it was notable the large number of antiinflammatory and antiproliferative targets which can be linked to the activities already reported in the literature for this natural compound (Anand David et al., 2016).

Among the antiinflammatory targets indicated by TF of quercetin, phosphodiesterase 4B (PDE4B) stood out due to its reported action (Jin et al., 2012; Wittmann and Helliwell, 2013). This result can be related to some studies that had evaluated the PDE4B inhibitory capacity of secondary metabolites, including quercetin (Ko et al., 2004; Townsend and Emala, 2013). PDE4 blocker activity was also described with inhibitors derivatives

TABLE 1 | NMR spectroscopic data (500 MHz, DMSO-d_6) for compound 1.

Position	δ_C, type	δ_H, (*J* in Hz)
2	156.56, C	
3	134.44, C	
4	177.92, C	
5	157.14, C	
6	98.88, CH	6.18, s
7	164.48, C	
8	93.84, CH	6.38, s
9	161.40, C	
10	104.02, C	
1'	120.61, C	
2'	115.73, CH	7.32, s
3'	145.38, C	
4'	148.78, C	
5'	115.54, CH	6.87, d (8.25)
6'	121.03, CH	7.25, d (8.30)
1''	101.05, CH	5.29, s
2''	80.68, CH	4.00, s
3''	70.41, CH	*
4''	72.56, CH	*
5''	70.41, CH	*
6''	17.56, CH_3	0.90, d (6.10)
1'''	106.50, CH	4.08, d (7.15)
2'''	71.14, CH	3.12, t (9.50)
3'''	71.88, CH	3.12, t (9.50)
4'''	67.90, CH	*
5'''	65.93, CH_2	3.21, d (12.0)

Signs obscured by the water signal present in the DMSO-d_6

from this flavonol (Chan et al., 2008), which makes PDE4B a promising target for the study of molecules such as compound **1**.

According to this fact, the PDE4B (PDB ID: 4MYQ) (Fox et al., 2014) was highlighted in the ranking that was comprised of the five best results of molecular docking simulations with the complete structure of compound **1** (**Table S3**, **Supplementary Material**). The simulation illustrated a good similarity with the crystallized molecule and a favorable position of the flavonoid at the binding site of the enzyme. The interactions demonstrated a potential stable binding, indicating a hydrophobic interaction between PHE 618 and the ring C of quercetin, an indirect interaction with Mg^{+2} through the water molecules and hydrogen bonds with ASN 567 and TYR 405 (**Figure S6A**, Supplementary Material). In addition, the 2D representation, provided by the program LigPlot (Wallace et al., 1995), also pointed hydrogen bonds between the amino acid residues HIS 406 and GLU 476 (**Figure S6B**, **Supplementary Material**).

Interestingly, another PDE enzyme was indicated as a possible target for the interaction with the compound **1** (PDE10A2), reinforcing the potential capacity of PDE inhibition by this flavonoid (**Table S3**, **Supplementary Material**). The PDE10A2 is a unique gene that encodes the PDE10 family and it is expressed mostly in the brain (Keravis and Lugnier, 2012). Because of its restricted distribution, most tests evaluated the use of its inhibitors for the treatment of neurological diseases (Jørgensen et al., 2013). However, a possible antiinflammatory action has been indicated by a recent study that showed the capacity of small inhibitors to decrease the production of nitrite in LPS-simulated cells (García et al., 2017).

In contrast, the PDE4B inhibitors are widely known for their antiinflammatory action. This enzyme is the major cAMP-metabolizer found in cells of the inflammatory and immune system (Spadaccini et al., 2017) and it is differentiated from the other PDE4 isoforms by its high sensitivity towards the inhibitors (Azam and Tripuraneni, 2014). Among the four isoforms of phosphodiesterase 4 (A-D), PDE4B is believed to play a central role in inflammation, being characterized as the predominant subtype in monocytes and neutrophils (Azam and Tripuraneni, 2014; Yang et al., 2017). Notably, PDE4B has been reported as a pivotal target in the treatment of airway inflammatory changes (Li et al., 2018). This activity is also recognized in the use of the extract of *B. pinnatum* (Cruz et al., 2011), reinforcing the interpretation that the antiinflammatory action of the compound **1** may be a result, at least in part, from the blockade of PDE4B.

The other TF methodology used in this work, the ligand-based VS, also pointed out a target that belongs to the PDE family, increasing the evidences of this possible action by the compound **1**. The cocrystallized ligand within the PDE5A1 enzyme (Ligand PDB ID: 7CA) (Wang et al., 2006) had a similarity score of 0.952, being the only result that showed a similarity score value above 0.5. However, although the targets are both phosphodiesterase enzymes, their therapeutic actions are different and as well as their binding site. The result of protein sequence alignment between the PDE4B and PDE5A1 showed an identity of 25.5% and the difference of their active sites was easily noticed (**Figure S7**, **Supplementary Material**). Therefore, it is probable that compound **1** would not have the same inhibition capacity on PDE4B and PDE5A1.

The choice of the most promising target between the results of both VS techniques was made from a compilation of the lower energies derived from the scoring of docking simulations, the most favorable interaction with the target and the main activity of the extract of *B. pinnatum*. These findings add support to the view that the most promising target for further analyses is indeed the PDE4B enzyme.

PDE4 Activity *In Vitro* Evaluation

With the aim to investigate the PDE4B blocker capacity appropriate for molecular docking simulations, the *in vitro* evaluation was performed. We also tested the effect of compound **1** on PDE4A to determine the selectivity of our compound with the two PDE4 isoforms that are particularly important to the antiinflammatory action regulation (Manning et al., 1999).

The results of compound **1** were considered expressive concerning both isoforms, but the effect upon PDE4B seemed to be superior. It is important to highlight that the inhibition of PDE4B was even better than the standard drug (roflumilast). Furthermore, the results suggest a relative selectiveness between the PDE4A and PDE4B (**Table 2**).

Although the PDE4 is widely studied, the application of phosphodiesterase inhibitors in medicine is still compromised by their side effects, which include nausea and emesis, meaning that their use is limited (Spina, 2008). The exact mechanism that explains their side effects is still unclear, but one of the most accepted explanations involves the nonselective inhibitor effects on the four PDE4 isoforms (Wang et al., 2007). The limitation

TABLE 2 | Results of PDE4 *in vitro* inhibition.

PDE4A	
Compound (10 μM)	**% Inhibition**
Roflumilast	100 ± 9,58
Compound **1**	49,8 ± 7,66
PDE4B	
Compound (10 μM)	**% Inhibition**
Roflumilast	85,7 ± 0,68
Compound **1**	100 ± 0

concerning selectivity is easily justified when the structures of the PDE4 isoforms are analyzed. Their structures differ only in *N*-terminal regions, reflecting the extreme conservation of amino acid residues in their activity sites, particularly between PDE4B and PDE4D (Cheung et al., 2007). In this way, the results showed this flavonoid to be not just a potent PDE4B inhibitor but also a promising compound for the design of other selective analogues.

Molecular Dynamics Simulations and Binding-Free Energy Estimations

The close similarity between the binding sites of the two isoform makes the elucidation of the selectivity a challenging task. Therefore, there are some studies in literature that used experimental tests trying to identify crucial differences between these isoforms (Wang et al., 2007; Feng et al., 2018). However, the reasons behind the PDE4 subfamily selective inhibition are still unclear when it is based essentially on the binding sites of these enzymes. In this context, the use of alternative methods, such as *in silico* studies, can be a promising way to discover different findings and hypotheses related to this issue. Between them, the docking simulations, used in the VS study, have a particular disadvantage due to their lock-key assumption (Mezei, 2003; Andrusier et al., 2008). In contrast, the molecular dynamics simulation (MD) can theoretically predict the move of every atom contained in the system over time, capturing important findings related to the binding process (Hollingsworth and Dror, 2018). Because of this, MD was considered the most promising method to come to a hypothesis that would attempt to explain the selective inhibition demonstrated by compound **1** and, consequently, assist in the design of new selective PDE4B inhibitors.

One of the most widely used parameters in MD simulations is the root-mean-square deviation (RMSD). The RMSD reflects the average displacement of the atoms and its use allows us to determine the structural stabilization in the time scale of the simulation period (Martínez, 2015; Sargsyan et al., 2017). Every one of the five possible positions of the flavonoid was submitted to a lower simulation time before the further analyzes aiming to work with already optimized complexes. The RMSD of each binding pose of the flavonoid on PDE4A (PDB ID: 2QYK) (Wang et al., 2007) and PDE4B (PDB ID: 3O57) (Mitchell et al., 2010) demonstrated a complete optimization of the protein backbone before 5.0 ns, without a significant change of conformation by this time. This fact indicates that the time of the simulations was enough for the refinement of all the structures in the period used. In addition, the high similarity of the RMSD profile of each pose on both isoforms indicates another reflection of their resemblance (**Figure 2**).

To determine the better positions of compound **1** at the activity sites of both isoforms, a Molecular mechanics Poisson-Boltzmann surface area method (MM/PBSA) was carried out. This approach has been used in a lot of studies and can be combined with MD simulations to give important entropic contributions to the total binding energy (Homeyer and Gohlke, 2012; Kumari et al., 2014; Genheden and Ryde, 2015). It is important to emphasize the influence of the choice of nonpolar models in the binding-free energy results. Among the models, the solvent accessible surface area (SASA) is one of the most widely implemented nonpolar types of modeling in the MM/PBSA (Kumari et al., 2014). However, previous works have shown the difficulties of this model in the discrimination of conformational states in explicit solvent simulations, producing bias (Wagoner and Baker, 2006). To overcome this issue, other nonpolar solvent models were tested, such as the SAV and the hybrid model (SASA-SAV) (Wagoner and Baker, 2006; Kumari et al., 2014). In contrast with the SASA-only model, the data results showed that the SAV-only model can reproduce the explicit solvent nonpolar simulation with promising fidelity (Wagoner and Baker, 2006). Therefore, the SAV-only model was chosen for our analysis. Additionally, in MM/PBSA, the

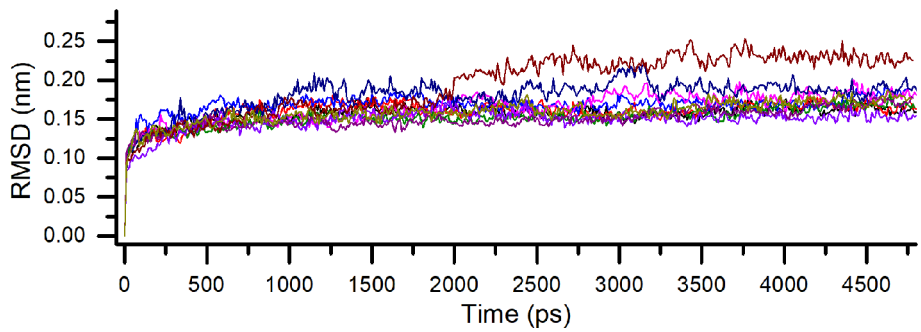

FIGURE 2 | Comparative root-mean-square deviation (RMSD) for all binding poses of compound 1 on PDE4A and PDE4B. Each pose is indicated by a different color.

result of binding-free energy also takes into account the values of electrostatic, Van der Walls and polar energies (Kumari et al., 2014). These values were computed for every one of the five binding poses for **1** on the two isoforms (**Table 3**).

The results demonstrated that, in the vast majority of the poses, the flavonoid had a considerably better binding-free energy value on the PDE4B isoform. These values reinforce the hypothesis of a more favorable interaction in a molecular scenario between the compound **1** and PDE4B and also corroborates with the *in vitro* results. Among the binding poses, the first and fifth ones stood out. The first pose, in particular, had the most promising binding-free energy result on the both PDE4A and PDE4B enzymes (−64.42 and −72.58 kcal mol^{-1}, respectively) which shows it to be an interesting binding position at the two activity sites. The fifth pose had a very similar binding energy on PDE4B to the first one (−72.38 kcal mol^{-1}) and can be considered as a potential occupation mode at the binding site of PDE4B as well.

Based on the results of MM/PBSA, the two most promising positions were selected to extend the time of MD simulations in order to analyze the complex behavior for a longer period. The RMSD of the two different poses, as before, demonstrated a very stable profile of optimization with a complete stabilization before 5.0 ns. Both pose conformations on PDE4A and PDE4B remained stable after the extension of the simulation time (**Figure 3**).

The output results after the extension of the MD time were also analyzed based on the relevant features of the binding site of each isoform. The structure of the activity site and the binding position of cAMP in the PDE4 enzyme are essential to the evaluation of the inhibitor action. The activity of PDE4 enzyme depends on the recognition of the cAMP's nucleotide followed by its hydrolysis. Structurally, the x-ray crystallography of the PDE4 indicated the three most relevant pockets for the activity of this enzyme (Spina, 2008). At the Q pocket, the cAMP is identified by the formation of hydrogen bonds with the nucleotide adenosine and, once it is recognized, the phosphate moiety complex with the metals (Zn^{2+}, Mg^{2+}) of the enzyme's bivalent metal-binding pocket (M pocket) promoting its hydrolysis (Houslay et al., 2005; Spina, 2008). The third pocket (S pocket) has polar amino acid residues and an interesting contact with the external solvent, probably contributing to the stabilization of the inhibitor at the binding site. In this context, the majority of PDE4 inhibitors show a structure capable of occupying specifically the Q and M pockets to prevent the entry and hydrolysis of cAMP (Houslay et al., 2005).

TABLE 3 | Calculated values that contributed to binding-free energy from MM/PBSA[a].

	Pose	ΔE_{vdw}	ΔE_{elec}	ΔG_{polar}	ΔG_{SAV}	ΔG_{bind}
PDE4A	**1**	−51.73 ± 0.34	−8.20 ± 0.21	34.64 ± 0.39	−39.10 ± 0.55	−64.42 ± 0.65
	2	−53.73 ± 0.58	0.14 ± 0.35	61.21 ± 2.0	−38.52 ± 1.03	−30.86 ± 2.29
	3	−45.93 ± 0.53	−11.52 ± 0.43	63.60 ± 3.52	−39.23 ± 1.19	−32.96 ± 2.59
	4	−46.42 ± 0.53	−11.09 ± 0.49	35.08 ± 0.72	−34.59 ± 0.82	−57.07 ± 1.01
	5	−41.88 ± 0.15	−7.20 ± 0.14	60.30 ± 0.86	−31.35 ± 0.22	−20.12 ± 0.86
PDE4B	**1**	−65.80 ± 0.44	−25.98 ± 0.43	73.587 ± 0.49	−54.40 ± 0.52	−72.58 ± 0.77
	2	−48.77 ± 0.63	−8.23 ± 0.34	37.30 ± 0.41	−35.84 ± 0.84	−55.60 ± 1.19
	3	−54.53 ± 0.67	1.18 ± 0.27	51.62 ± 0.70	−40.05 ± 1.02	−41.84 ± 1.41
	4	−54.40 ± 0.43	−6.44 ± 0.35	53.79 ± 0.79	−45.27 ± 1.09	−52.27 ± 1.30
	5	−56.63 ± 0.22	−21.89 ± 0.25	54.07 ± 0.28	−47.94 ± 0.25	−72.38 ± 0.35

a: *All values, standard deviations and averages, are expressed in kcal* mol^{-1}.

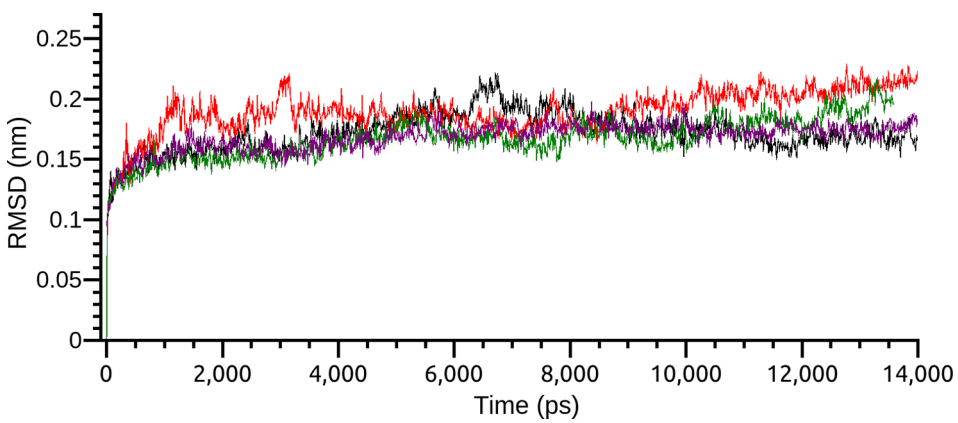

FIGURE 3 | Root-mean-square deviation (RMSD) of the two most interesting binding poses of compound 1 on PDE4A and PDE4B. Each pose is indicated by a different color.

Considering this structural information, the first position of the flavonoid demonstrated considerable intermolecular differences between its occupations at both of the binding sites of the isoforms. On PDE4A, the compound **1** showed a very stable position at the Q pocket with a double parallel π-stacking interaction between the amino acid residues PHE 296 and PHE 265, and also the hydrogen bonds with the residues HIS 201 and TYR 84. However, the compound clearly had an important dislocation at the activity site and had no interactions with the metals of the M pocket (**Figure 4A**). As already mentioned, the interactions with the bivalent metals influence the inhibition of PDE4 enzymes in a direct way (Houslay et al., 2005). Therefore, the behavior of the best pose of the flavonoid indicated by MM/PBSA on PDE4A represents a valid hypothesis for the lower inhibition capacity of compound **1** on this isoform. In contrast, the compound **1** at the activity site of PDE4B also demonstrated the occupation of the Q pocket, showing one parallel π-stacking interaction with the amino acid residue PHE 296. Additionally, there was an increase of hydrogen bonds between the glycosides and the polar amino acid residues and an indirect interaction with the metals through hydrogen bonds with the water molecules (**Figure 4B**). This increase was clearly demonstrated in the 2D representation of the intermolecular interactions of compound **1** at both binding sites. It was detected the presence of 8 valid hydrogen bonds between the natural product and the amino acid residues of PDE4B (**Figure 4D**) and only 2 valid hydrogen bonds with the amino acid residues of PDE4A (**Figure 4C**).

The fifth pose also demonstrated a promising result. This binding position showed a large difference of the binding-free energy of the flavonoid on PDE4A and PDE4B. However, in particular, the results of MM/PBSA on PDE4B pointed to a very similar binding-free energy in comparison with the first pose. Interestingly, the analyses on both the binding sites demonstrated that the most significant variation involved the hydrophobic interactions with the amino acid residues of phenylalanine present in the Q pocket. At the PDE4A, the flavonoid exhibited a parallel π-stacking interaction with PHE 297 (**Figure 5A**). In comparison, the compound **1** had a double parallel π-stacking with PHE 296 and PHE 264 at the PDE4B (**Figure 5B**). However, even with these similar interactions, the results of binding-free energy clearly indicate disadvantages of this pose at the PDE4A activity site. Analyzing the 2D representations, it was possible to confirm the higher number of hydrogen bonds between the compound **1** and the amino acid residues of PDE4B (**Figure 5D**) in comparison with PDE4A

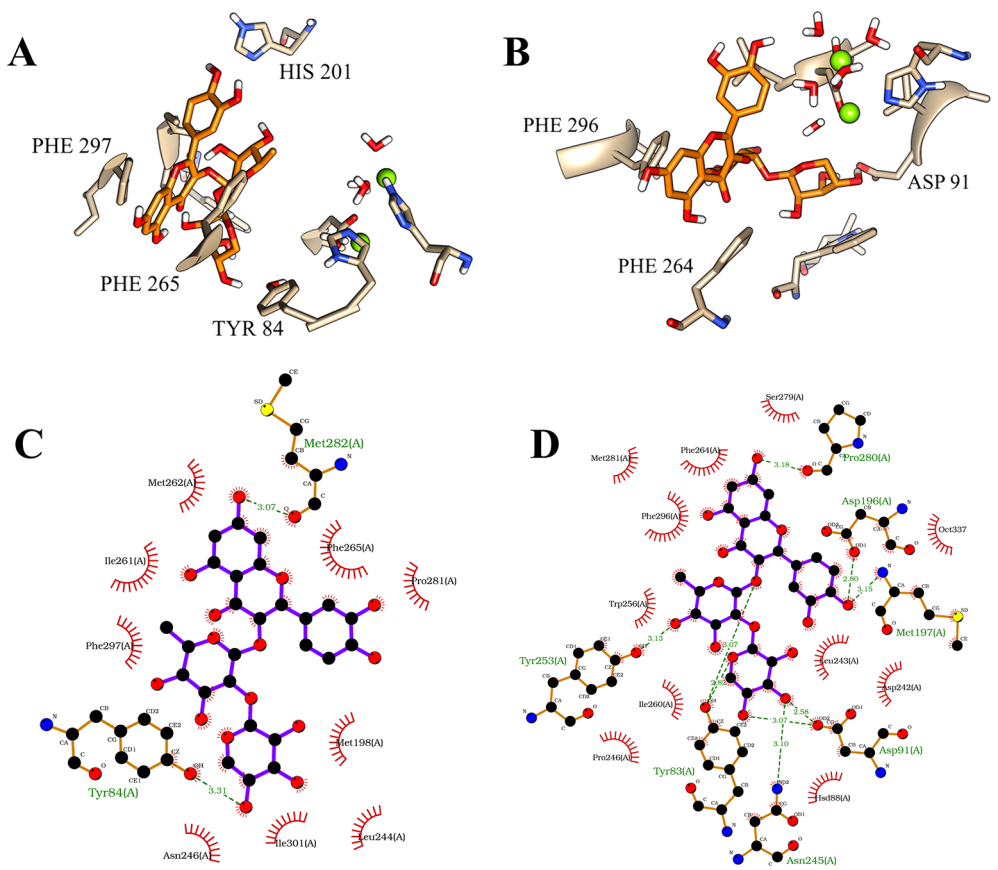

FIGURE 4 | Illustration of interactions between compound 1 (orange) and the amino acid residues of PDE4A **(A)** and PDE4B **(B)** for the first pose. 2D representation of the intermolecular interactions between the compound 1 and PDE4A **(C)** and PDE4B **(D)**. The 2D diagrams was provided by the program LigPlot.

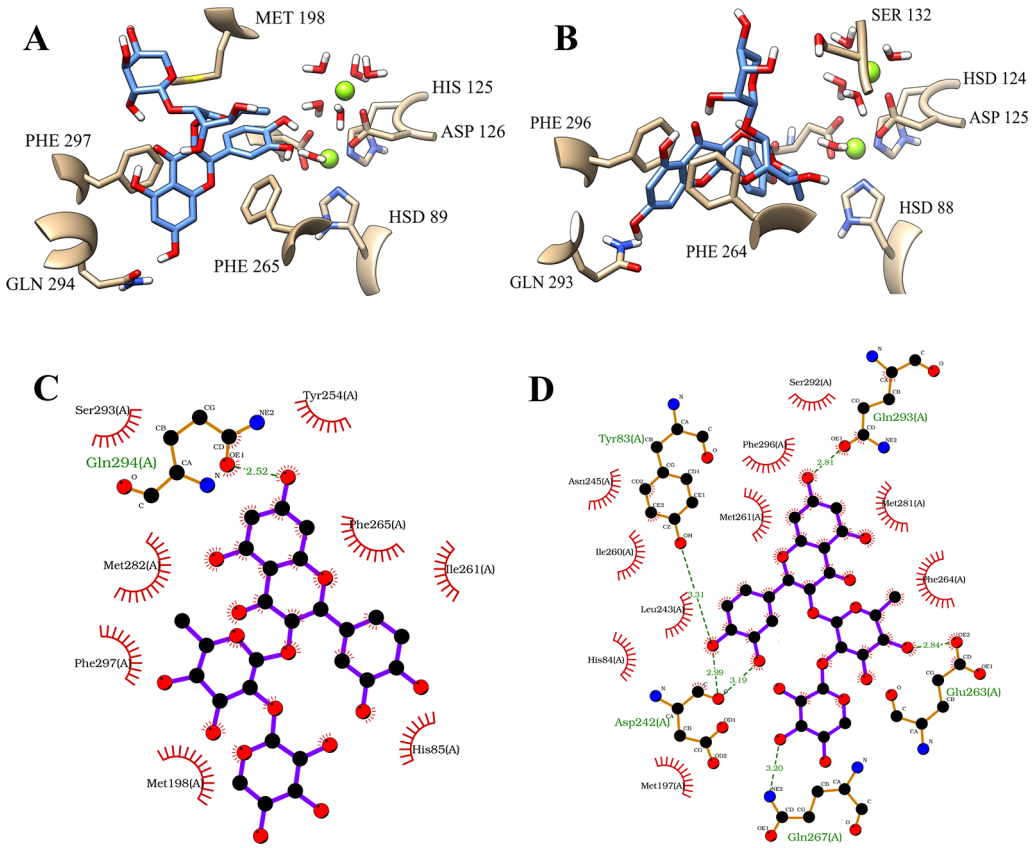

FIGURE 5 | Illustration of interactions between compound 1 (blue) and the amino acid residues of PDE4A **(A)** and PDE4B **(B)** for the fifth pose. 2D representation of the intermolecular interactions between the compound 1 and PDE4A **(C)** and PDE4B **(D)**. The 2D diagrams was provided by the program LigPlot.

(**Figure 5C**). This difference may be related with the better interaction with PDE4B pointed by the favorable binding-free energy value.

Another interesting fact involving these two potential binding poses is the positional reflection between them. However, the behaviors of the poses on PDE4A and PDE4B were substantially different. In regard to the PDE4B isoform, both poses were stable at the activity site. In the first pose the glycosides occupied the polar region between the Q and M pockets and the fifth pose demonstrated that the compound **1** was also exploring polar interactions but at the hydrated S pocket. The same was not observed in PDE4A. The difference in the binding-free energy results of the first and fifth poses (−64.42 and −20.12 kcal mol^{-1}, respectively) was reflected in the interaction of these binding positions in the two isoforms. The compilation results on PDE4A, demonstrated that even the first pose which had the most promising binding-free energy, showed major hindrances in the occupation of the activity site.

The traditional binding pose observed for the known PDE4 inhibitors, including the Roflumilast, had demonstrated that the most important occupation at the active site is with the amino acids belonging to the Q and M pockets (**Figure S8**, Supplementary Material). Especially three types of interaction are crucial to the PDE4 inhibition: interactions with the metal ions through structural water molecules, hydrogen bonds and hydrophobic interactions with the amino acid residues at Q pocket (Card et al., 2004). In this molecular scenario, compound **1** followed the traditional binding pose exhibited by most of the PDE4 inhibitors and had a widespread occupation of the activity site. However, the occupation in PDE4B was demonstrated to be higher after the molecular dynamics simulations, suggesting a formation of a stable system of hydrogen bonds and hydrophobic interactions. Therefore, the small but significative differences between the PDE4 subfamilies found in previous literature on experimental tests (Wang et al., 2007; Feng et al., 2018), seems to be crucial to the whole complex behavior simulated *in silico*.

Although compound **1** showed interesting *in vitro* and *in silico* results, it is important to emphasizing that this compound has disadvantages as a drug candidate. The enzyme lactase phrorizin hydrolase, found in small intestine of mammalians, is capable of hydrolyzing a range of flavonol and isoflavone glycosides which difficult the abortion of these compounds as a whole molecule (Day et al., 2000). Since the molecular dynamics simulations had demonstrated the importance of the glyosidic portion for the PDE4 blocker action and selectivity profile, potential medications consisting only of compound **1** would not be optimal for per oral, the most common type of administration. However, it can be considered as a promising

lead compound and the results of molecular dynamics simulations may be used to design novel potential PDE4B inhibitors based on the structure of compound **1**. Aiming to increase the synthetic accessibility of the designed compounds, the glycosides may be replaced by another polar structure in order to preserve the hydrogen bonds made in S pocket. In contrast, quercetin aglycone appears to be essential for the inhibition of the both isoforms and its replacement should be made by other structure with high similarity.

CONCLUSION

Our studies showed that the employed TF methodologies were efficient approaches to identify the alleged activities of natural products, abruptly reducing the costs and numbers of biological models. Based on the results, this approach proved to be an interesting tool for the chemical and pharmacological investigation of possible natural extract markers such as compound **1**. The flavonoid quercetin 3-*O*-α-*L*-arabinopyranosyl-(1→2)-*O*-α-*L*-rhamnopyranoside, one of the major compounds of *B. pinnatum*, had its antiinflammatory activity explained by the inhibition of PDE4B. This action confirmed by *in vitro* evaluation, indicates that compound **1** is not just a potent PDE4 blocker, but also demonstrated it to be highly selective to PDE4B. In addition, it was possible to explore this inhibiting and selectivity action of compound **1** with molecular dynamics simulations which shed light on the atomic level inhibition properties of this compound. These results show important progress in the investigation of the antiinflammatory properties of *B. pinnatum* and reinforced the possible choice of compound **1** as the extract chemical marker of this plant. Finally, these highly expressive results show this flavonoid to be an interesting prototype for the design of other PDE4B selective inhibitors.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

EL, SZ, and EB conceived and designed the experimental tests. The development of the methodology was made by EL, JF, SZ, VC, MM, and EB. Analysis of the data was made by EL, AJ, VC, MM, and EB. Contributed reagents/materials/analysis: JM, SM, VC, and MM. The paper was written and reviewed by EL, AJ, VC, MM, SZ, and EB.

FUNDING

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

ACKNOWLEDGMENTS

The authors acknowledge the CAPES for the granting of scholarships and the High-Performance Computing Center at UFRN (NPAD/UFRN) for the support offered in this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2019.01582/full#supplementary-material

REFERENCES

Anand David, A., Arulmoli, R., and Parasuraman, S. (2016). Overviews of biological importance of quercetin: a bioactive flavonoid. *Pharmacogn. Rev.* 10 (20), 84–89. doi: 10.4103/0973-7847.194044

Andrusier, N., Mashiach, E., Nussinov, R., and Wolfson, H. J. (2008). Principles of flexible protein-protein docking. *Proteins* 73 (2), 271–289. doi: 10.1002/prot.22170

Azam, M. A., and Tripuraneni, N. S. (2014). Selective phosphodiesterase 4B inhibitors: a review. *Sci. Pharm.* 82 (3), 453–482. doi: 10.3797/scipharm.1404-08

Barreiro, E. J., and Bolzani, V. D. S. (2009). Biodiversidade: fonte potencial para a descoberta de fármacos. *Quim. Nova* 679–688. doi: 10.1590/S0100-40422009000300012

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., et al. (1978). The protein data bank: a computer-based archival file for macromolecular structures. *Arch. Biochem. Biophys.* 80 (2), 584–591. doi: 10.1016/0003-9861(78)90204-7

Card, G. L., England, B. P., Suzuki, Y., Fong, D., Powell, B., Lee, B., et al. (2004). Structural basis for the activity of drugs that inhibit phosphodiesterases. *Structure* 12 (12) 2233–2247. doi: 10.1016/j.str.2004.10.004

Chan, A. L., Huang, H. L., Chien, H. C., Chen, C. M., Lin, C. N., and Ko, W. C. (2008). Inhibitory effects of quercetin derivatives on phosphodiesterase isozymes and high-affinity [3H]-rolipram binding in guinea pig tissues. *Invest. New Drugs* 26 (5), 417–424. doi: 10.1007/s10637-008-9114-7

Cheung, Y.-F., Kan, Z., Garrett-Engele, P., Gall, I., Murdoch, H., Baillie, G. S., et al. (2007). PDE4B5, a novel, super-short, brain-specific cAMP phosphodiesterase-4 variant whose isoform-specifying N-terminal region is identical to that of cAMP phosphodiesterase-4D6 (PDE4D6). *J. Pharmacol. Exp. Ther.* 322 (2), 600–609. doi: 10.1124/jpet.107.122218

Cruz, E. A., Reuter, S., Martin, H., Dehzad, N., Muzitano, M. F., Costa, S. S., et al. (2011). Kalanchoe pinnata inhibits mast cell activation and prevents allergic airway disease. *Phytomedicine* 19 (2), 115–121. doi: 10.1016/j.phymed.2011.06.030

Day, A. J., Cañada, F. J., Díaz, J. C., Kroon, P. A., McLauchlan, R., Faulds, C. B., et al. (2000). Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase. *FEBS Lett.* 468 (2–3), 166–170. doi: 10.1016/S0014-5793(00)01211-4

dos Santos Nascimento, L. B., de Aguiar, P. F., Leal-Costa, M. V., Coutinho, M. A. S., Borsodi, M. P. G., Rossi-Bergmann, B., et al. (2018). Optimization of aqueous extraction from kalanchoe pinnata leaves to obtain the highest content of an anti-inflammatory flavonoid using a response surface model. *Phytochem. Anal.* 29 (3), 308–315. doi: 10.1002/pca.2744

Espargaró, A., Ginex, T., Vadell, M., del, M., Busquets, M. A., Estelrich, J., et al. (2017). Combined *in vitro* Cell-Based/*in silico* screening of naturally occurring flavonoids and phenolic compounds as potential anti-alzheimer drugs. *J. Nat. Prod.* 80 (2), 278–289. doi: 10.1021/acs.jnatprod.6b00643

Ezuruike, U. F., and Prieto, J. M. (2014). The use of plants in the traditional management of diabetes in nigeria: pharmacological and toxicological considerations. *J. Ethnopharmacol.* 155 (2), 857–924. doi: 10.1016/j.jep.2014.05.055

Feng, X., Wang, H., Ye, M., Xu, X. T., Xu, Y., Yang, W., et al. (2018). Identification of a PDE4-Specific Pocket for the Design of Selective Inhibitors. *Biochemistry* 57 (30), 4518–4525. doi: 10.1021/acs.biochem.8b00336

Fernandes, J. M., Félix-Silva, J., da Cunha, L. M., Gomes, J. A., dos, S., Siqueira, E. M. S., et al. (2016). inhibitory effects of hydroethanolic leaf extracts of kalanchoe brasiliensis and kalanchoe pinnata (crassulaceae) against local effects induced by bothrops jararaca snake venom. *PloS One* e0168658. doi: 10.1371/journal.pone.0168658

Fernandes, J. M., Cunha, L. M., Azevedo, E. P., Lourenço, E. M. G., Fernandes-Pedrosa, M. F., and Zucolotto, S. M. (2019). Kalanchoe laciniata and Bryophyllum pinnatum: an updated review about ethnopharmacology, phytochemistry, pharmacology and toxicology. *Rev. Bras. Farmacogn.* 29 (4), 529–558. doi: 10.1016/j.bjp.2019.01.012

Fox, D., Burgin, A. B., and Gurney, M. E. (2014). Structural basis for the design of selective phosphodiesterase 4B inhibitors. *Cell. Signal.* 26 (3), 657–663. doi: 10.1016/j.cellsig.2013.12.003

Freires, I. A., Sardi, J., de, C. O., de Castro, R. D., and Rosalen, P. L. (2017). Alternative animal and non-animal models for drug discovery and development: bonus or burden? *Pharm. Res.* 34, 681–686. doi: 10.1007/s11095-016-2069-z

García, A. M., Brea, J., González-García, A., Pérez, C., Cadavid, M. I., Loza, M. I., et al. (2017). Targeting PDE10A GAF domain with small molecules: A way for allosteric modulation with anti-inflammatory effects. *Molecules* 22 (9), E1472. doi: 10.3390/molecules22091472

Genheden, S., and Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discovery* 10 (5), 449–461. doi: 10.1517/17460441.2015.1032936

Hollingsworth, S. A., and Dror, R. O. (2018). Molecular Dynamics Simulation for All. *Neuron* 99 (6), 1129–1143. doi: 10.1016/j.neuron.2018.08.011

Homeyer, N., and Gohlke, H. (2012). Free energy calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area method. *Mol. Inform.* 31 (2), 114–122. doi: 10.1002/minf.201100135

Horbal, L., Marques, F., Nadmid, S., Mendes, M. V., and Luzhetskyy, A. (2018). Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metab. Eng.* 49, 299–315. doi: 10.1016/j.ymben.2018.09.010

Houslay, M. D., Schafer, P., and Zhang, K. Y. J. (2005). Keynote review: Phosphodiesterase-4 as a therapeutic target. *Drug Discovery Today* 10 (12), 1503–1519. doi: 10.1016/S1359-6446(05)03622-6

Jørgensen, M., Kehler, J., Langgård, M., Svenstrup, N., and Tagmose, L. (2013). Selective inhibitors of PDE2, PDE9, and PDE10: Modulators of activity of the central nervous system. *Annu. Rep. Med. Chem.* 48, 37–55. doi: 10.1016/B978-0-12-417150-3.00004-1

Jin, S. L., Ding, S. L., and Lin, S. C. (2012). Phosphodiesterase 4 and its inhibitors in inflammatory diseases. *Chang Gung Med. J.* 35, 197–210.

Keravis, T., and Lugnier, C. (2012). Cyclic nucleotide phosphodiesterase (PDE) isozymes as targets of the intracellular signalling network: Benefits of PDE inhibitors in various diseases and perspectives for future therapeutic developments. *Br. J. Pharmacol.* 166 (5), 1288–1305. doi: 10.1111/j.1476-5381.2011.01729.x

Ko, W. C., Shih, C. M., Lai, Y. H., Chen, J. H., and Huang, H. L. (2004). Inhibitory effects of flavonoids on phosphodiesterase isozymes from guinea pig and their structure-activity relationships. *Biochem. Pharmacol.* 68 (10), 2087–2094. doi: 10.1016/j.bcp.2004.06.030

Kumari, R., Kumar, R., and Lynn, A. (2014). g_mmpbsa – A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* 54 (7), 1951–1962. doi: 10.1021/ci500020m

Li, H., Zuo, J., and Tang, W. (2018). Phosphodiesterase-4 inhibitors for the treatment of inflammatory diseases. *Front. Pharmacol.* 9, 1048. doi: 10.3389/fphar.2018.01048

Manning, C. D., Burman, M., Christensen, S. B., Cieslinski, L. B., Essayan, D. M., Grous, M., et al. (1999). Suppression of human inflammatory cell function by subtype-selective PDE4 inhibitors correlates with inhibition of PDE4A and PDE4B. *Br. J. Pharmacol.*, 128 (7), 1393–1398. doi: 10.1038/sj.bjp.0702911

Martínez, L. (2015). Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PloS One* e0119264. doi: 10.1371/journal.pone.0119264

Mezei, M. (2003). A new method for mapping macromolecular topography. *J. Mol. Graph. Model.* 21 (15), 463:472. doi: 10.1016/S1093-3263(02)00203-6

Mitchell, C. J., Ballantine, S. P., Coe, D. M., Cook, C. M., Delves, C. J., Dowle, M. D., et al. (2010). Pyrazolopyridines as potent PDE4B inhibitors: 5-Heterocycle SAR. *Bioorg. Med. Chem. Lett.* 20 (19), 5803–5803. doi: 10.1016/j.bmcl.2010.07.136

Muzitano, M. F., Tinoco, L. W., Guette, C., Kaiser, C. R., Rossi-Bergmann, B., and Costa, S. S. (2006). The antileishmanial activity assessment of unusual flavonoids from Kalanchoe pinnata. *Phytochemistry* 67 (18), 2071–2077. doi: 10.1016/j.phytochem.2006.06.027

Nascimento, L. B. S., Leal-Costa, M. V., Coutinho, M. A. S., Moreira, N. D. S., Lage, C. L. S., Barbi, N. D. S., et al. (2013). Increased antioxidant activity and changes in phenolic profile of kalanchoe pinnata (Lamarck) persoon (crassulaceae) specimens grown under supplemental blue light. *Photochem. Photobiol.* 148, 73–81. doi: 10.1111/php.12006

Nascimento, L. B. D. S., Leal-Costa, M. V., Menezes, E. A., Lopes, V. R., Muzitano, M. F., Costa, S. S., et al. (2015). Ultraviolet-B radiation effects on phenolic profile and flavonoid content of Kalanchoe pinnata. *J. Photochem. Photobiol. B Biol.* 73–81. doi: 10.1016/j.jphotobiol.2015.03.011

Olson, A. J., and Trott, O. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* 31 (2), 455–461. doi: 10.1002/jcc.21334

Pérez, G. M., Salomón, L. A., Montero-Cabrera, L. A., de la Vega, J. M. G., and Mascini, M. (2016). Integrating sampling techniques and inverse virtual screening: toward the discovery of artificial peptide-based receptors for ligands. *Mol. Divers.* 421, 438. doi: 10.1007/s11030-015-9648-5

Panyaphu, K., Van On, T., Sirisa-Ard, P., Srisa-Nga, P., Chansakaow, S., and Nathakarnkitkul, S. (2011). Medicinal plants of the Mien (Yao) in Northern Thailand and their potential value in the primary healthcare of postpartum women. *J. Ethnopharmacol.* 226–237. doi: 10.1016/j.jep.2011.03.050

Patridge, E., Gareiss, P., Kinch, M. S., and Hoyer, D. (2016). An analysis of FDA-approved drugs: Natural products and their derivatives. *Drug Discovery Today* 204–207. doi: 10.1016/j.drudis.2015.01.009

Rester, U. (2008). From virtuality to reality - Virtual screening in lead discovery and lead optimization: a medicinal chemistry perspective. *Curr. Opin. Drug Discovery Devel.* 559–568.

Rodrigues, R. P., Mantoani, S. P., de Almeida, J. R., Pinsetta, F. R., Semighini, E. P., da Silva, V. B., et al. (2012). Estratégias de Traigem Virtual no Planejamento de Fármacos. *Rev. Virtual Quim.* 739–776. doi: 10.5935/1984-6835.20120055

Sargsyan, K., Grauffel, C., and Lim, C. (2017). How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* 1518–1524. doi: 10.1021/acs.jctc.7b00028

Sobreira, F., Hernandes, L. S., Vetore-Neto, A., Díaz, I. E. C., de Santana, F. C., Mancini-Filho, J., et al. (2017). Gastroprotective activity of the hydroethanolic extract and ethyl acetate fraction from Kalanchoe pinnata (Lam.) Pers. *Braz. J. Pharm. Sci.* 53 (1), e16027. doi: 10.1590/s2175-97902017000116027

Spadaccini, M., D'Alessio, S., Peyrin-Biroulet, L., and Danese, S. (2017). PDE4 inhibition and inflammatory bowel disease: A novel therapeutic avenue. *Int. J. Mol. Sci.* 18 (6), e1276. doi: 10.3390/ijms18061276

Spina, D. (2008). PDE4 inhibitors: Current status. *Br. J. Pharmacol.* 155 (3), 308–315. doi: 10.1038/bjp.2008.307

Sreekeesoon, D. P., and Mahomoodally, M. F. (2014). Ethnopharmacological analysis of medicinal plants and animals used in the treatment and management of pain in Mauritius. *J. Ethnopharmacol.* 181–200. doi: 10.1016/j.jep.2014.09.030

Townsend, E. A., and Emala, C. W. (2013). Quercetin acutely relaxes airway smooth muscle and potentiates -agonist-induced relaxation via dual phosphodiesterase inhibition of PLC and PDE4. *AJP Lung Cell. Mol. Physiol.* 305 (5), L396–L403. doi: 10.1152/ajplung.00125.2013

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *J. Comput. Chem.* 1701–1718. doi: 10.1002/jcc.20291

Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., et al. (2010). CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* 671–690. doi: 10.1002/jcc.21367

Wadood, A., Ahmed, N., Shah, L., Ahmad, A. H. H., and Shams, S. (2013). *In-silico* drug design: An approach which revolutionarised the drug discovery process. *Open Access Drug Des. Deliv.* 1–3.

Wagoner, J. A., and Baker, N. A. (2006). Assessing implicit models for nonpolar mean solvation forces: the importance of dispersion and volume terms. *Proc. Natl. Acad. Sci.* 103 (22), 8331–8336. doi: 10.1073/pnas.0600118103

Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995). Ligplot: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng. Des. Sel.* 127, 134. doi: 10.1002/jcc.24467

Wang, H., Liu, Y., Huai, Q., Cai, J., Zoraghi, R., Francis, S. H., et al. (2006). Multiple conformations of phosphodiesterase-5: Implications for enzyme function and drug development. *J. Biol. Chem.* 281, 21469–21479. doi: 10.1074/jbc.M512527200

Wang, H., Peng, M.-S., Chen, Y., Geng, J., Robinson, H., Houslay, M. D., et al. (2007). Structures of the four subfamilies of phosphodiesterase-4 provide insight into the selectivity of their inhibitors. *Biochem. J.* 193–201. doi: 10.1042/BJ20070970

Wittmann, M., and Helliwell, P. S. (2013). Phosphodiesterase 4 inhibition in the treatment of psoriasis, psoriatic arthritis and other chronic inflammatory diseases. *Dermatol. Ther. (Heidelb).* 1–15. doi: 10.1007/s13555-013-0023-0

Xu, X., Huang, M., and Zou, X. (2018). Docking-based inverse virtual screening: methods, applications, and challenges. *Biophys. Reports* 1–16. doi: 10.1007/s41048-017-0045-8

Yang, J. X., Hsieh, K. C., Chen, Y. L., Lee, C. K., Conti, M., Chuang, T. H., et al. (2017). Phosphodiesterase 4B negatively regulates endotoxin-activated interleukin-1 receptor antagonist responses in macrophages. *Sci. Rep.* 7 46165. doi: 10.1038/srep46165

Zoete, V., Daina, A., Bovigny, C., and Michielin, O. (2016). SwissSimilarity: a web tool for low to ultra high throughput ligand-based virtual screening. *J. Chem. Inf. Model.* 1399–1404. doi: 10.1021/acs.jcim.6b00174

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 June 2020
doi: 10.3389/fpls.2020.00848

Genome-Wide Analysis of Serine Carboxypeptidase-Like Acyltransferase Gene Family for Evolution and Characterization of Enzymes Involved in the Biosynthesis of Galloylated Catechins in the Tea Plant (*Camellia sinensis*)

*Muhammad Zulfiqar Ahmad[1], Penghui Li[1], Guangbiao She[1], Enhua Xia[1], Vagner A. Benedito[2], Xiao Chun Wan[1] and Jian Zhao[1]**

[1] State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, China, [2] Division of Plant & Soil Sciences, West Virginia University, Morgantown, WV, United States

OPEN ACCESS

Edited by:
M. Carmen González-Mas, University of Valencia, Spain

Reviewed by:
Jing Zhuang, Nanjing Agricultural University, China
Xinchao Wang, Tea Research Institute (CAAS), China
Xinyuan Hao, Tea Research Institute (CAAS), China

***Correspondence:**
Jian Zhao
Jianzhao@ahau.edu.cn

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

Received: *10 March 2020*
Accepted: *26 May 2020*
Published: *25 June 2020*

Citation:
Ahmad MZ, Li P, She G, Xia E, Benedito VA, Wan XC and Zhao J (2020) Genome-Wide Analysis of Serine Carboxypeptidase-Like Acyltransferase Gene Family for Evolution and Characterization of Enzymes Involved in the Biosynthesis of Galloylated Catechins in the Tea Plant (Camellia sinensis). Front. Plant Sci. 11:848. doi: 10.3389/fpls.2020.00848

Tea (*Camellia sinensis* L.) leaves synthesize and concentrate a vast array of galloylated catechins (e.g., EGCG and ECG) and non-galloylated catechins (e.g., EGC, catechin, and epicatechin), together constituting 8%–24% of the dry leaf mass. Galloylated catechins account for a major portion of soluble catechins in tea leaves (up to 75%) and make a major contribution to the astringency and bitter taste of the green tea, and their pharmacological activity for human health. However, the catechin galloylation mechanism in tea plants is largely unknown at molecular levels. Previous studies indicated that glucosyltransferases and serine carboxypeptidase-like acyltransferases (SCPL) might be involved in the process. However, details about the roles of SCPLs in the biosynthesis of galloylated catechins remain to be elucidated. Here, we performed the genome-wide identification of SCPL genes in the tea plant genome. Several SCPLs were grouped into clade IA, which encompasses previously characterized SCPL-IA enzymes with an acylation function. Twenty-eight tea genes in this clade were differentially expressed in young leaves and vegetative buds. We characterized three SCPL-IA enzymes (CsSCPL11-IA, CsSCPL13-IA, CsSCPL14-IA) with galloylation activity toward epicatechins using recombinant enzymes. Not only the expression levels of these SCPLIA genes coincide with the accumulation of galloylated catechins in tea plants, but their recombinant enzymes also displayed β-glucogallin:catechin galloyl acyltransferase activity. These findings provide the first insights into the identities of genes encoding glucogallin:catechin galloyl acyltransferases with an active role in the biosynthesis of galloylated catechins in tea plants.

Keywords: acyltransferase, catechins, enzyme activity, galloylation, health function, specialized metabolism

INTRODUCTION

Many plant metabolites, particularly specialized compounds, are subject to various modifications, such as glycosylation, malonylation, methylation, acylation, and prenylation (Bowles et al., 2005; Kosma et al., 2012; Bontpart et al., 2015; Ahmad et al., 2017). These modifications further diversify plant specialized metabolites and generate modified bioactive molecules with additional physicochemical properties, thereby creating new biological properties that the original metabolites do not have (Wang et al., 2019). Plant genomes evolved to owning a large number of gene families encoding modifying enzymes in order to adapt against adverse environments and survive frequent attacks by pathogens and herbivores (Wilson et al., 2016). Among these modifications, acylation is the most common and important of the modifications on plant metabolites, including phenolics, lipid barriers, sugars, and polyamines. The various modes of acylations enable plants to gain many functions in metabolic and physiological processes, as well as defense responses against abiotic or biotic stresses (Bontpart et al., 2015; Wilson et al., 2016).

Plant phenolics are ubiquitous specialized metabolites that contain a phenolic ring with at least one hydroxyl group (Bontpart et al., 2015). They are involved in growth, development, and defense responses. As the most extensively studied specialized metabolites in plants, their biosynthetic pathways are well known and involve multiple modification types, such as acylation, which culminates with the production of a vast diversity of phenolic products. Moreover, due to the high availability of simple phenolics in most plants, they often are not only the acceptor molecules during acylation but are also modified as energy-rich donors during other types of metabolite modifications in plants (Bontpart et al., 2015; Wilson et al., 2016). Two major acyltransferase families that use phenolic compounds either as acceptor or donor molecules have been described and characterized, each of which uses a distinct type of acyl donor (Bontpart et al., 2015). BAHD acyltransferases, named after the first four enzymes biochemically characterized in the family (BEAT, AHCT, HCBT, and DAT), have been extensively studied in a wide range of plant specialized metabolism and are characterized by using acyl-CoA thioesters as donor molecules (Schilmiller et al., 2012; Bontpart et al., 2015; Wilson et al., 2016; Zhang et al., 2018). On the other hand, serine carboxypeptidase-like (SCPL) acyltransferases use 1-*O*-β-glucose esters as acyl donors to facilitate the transacylation reaction to a large variety of phenolics, acids, saponins, and other compounds were only discovered recently (Mugford and Milkowski, 2012).

SCPLs revealed sequence homology to serine carboxypeptidases (Steffens, 2000; Milkowski and Strack, 2004; Bontpart et al., 2015). SCPLs can produce a variety of phenolic compounds, such as β-1-cinnamoyl-D-glucose, sinapoyl esters, and gallotannins (Wilson et al., 2016). However, not a single gene coding for an SCPL was characterized until the early 2000s (Lehfeldt et al., 2000; Li and Steffens, 2000). SCPLs provide biological relevance to various compounds, such as sinapoylglucose:malate sinapoyltransferase (Lehfeldt et al., 2000), sinapoylglucose:choline sinapoyltransferase (Shirley and Chapple, 2003), sinapoylglucose:sinapoylglucose sinapoyltransferase (Fraser et al., 2007), sinapoylglucose:anthocyanin sinapoyltransferase (Fraser et al., 2007) in Arabidopsis. These enzymes use 1-*O*-sinapoyl-β-D-glucose to form sinapoylmalate, sinapoylcholine (sinapine), 1,2-di-*O*-sinapoyl-β-D-glucose, and sinapoylated anthocyanins, respectively. Subsequently, two *Brassica napus* orthologs BnSCT1 and BnSCT2 were also characterized (Milkowski and Strack, 2004; Weier et al., 2008). Gallic acid is a trihydroxybenzoic acid that occurs in several dicot clades (Karas et al., 2017). Galloylated flavon-3-ols play vital roles in protecting plant cell membranes from oxidative damage (Saffari and Sadrzadeh, 2004). UDP-glucose:sinapic acid glucosyltransferase synthesizes sinapoylglucose in the Brassicaceae (Milkowski and Strack, 2004) whereas three UGTs form hydroxycinnamoyl glucose esters and β-glucogallin in grapevine (Khater et al., 2012), which are used by several SCPLs. Three grapevine (*Vitis vinifera*) glucosyltransferase genes *VvgGT1* to *3* were isolated and characterized (Khater et al., 2012). A purified polypeptide from tea leaves displaying epicatechin:1-*O*-galloyl-β-D-glucose *O*-galloyltransferase activity was previously reported (Liu et al., 2012). Galloyl-glucose esters, such as galloyl 1-*O*-β-D-glucose(β-glucogallin), is required to act as an acyl donor for the galloylation of flavon-3-ols. β-Glucogallin was produced from gallic acid by a UDP-glucosyltransferase in a crude protein extract of tea plant (*Camellia sinensis*) leaves (Liu et al., 2012). However, the gene encoding this enzyme is yet to be reported.

Catechins is the collective term for tea soluble flavan-3-ols, including six major types, (−)-epicatechin (EC), (+)-catechin (C), (−)-epigallocatechin (EGC), (+)-gallocatechin (GC), (−)-epicatechin-3-gallate (ECG), and (−)-epigallocatechin-3-gallate (EGCG) (Wei et al., 2018; Zhao et al., 2020). Among them, galloylated (e.g., EGCG and ECG) and non-galloylated catechins (e.g., EGC, C, and EC) together constitute 8%–24% of the dry leaf mass. As characteristic to teas, galloylated catechins account for up to 75% of the total catechins and make major contributions to the astringency and bitterness tastes of the green tea and numerous pharmacological activities for promoting human health (Zhao et al., 2020). Despite of their importance, the molecular and genetic bases of catechin galloylation remain largely unexplored. In tea plant leaves, several flavonoid 3′-hydroxylases and flavonoid 3′,5′-hydroxylases are two types of enzymes controlling the generation of EC and EGC, respectively, two major non-galloylated catechins (Wei et al., 2015). As precursors of galloylated flavan-3-ols and flavonols, β-glucogallin or polygalloylated glucoses are ubiquitous in the core eudicots (Moore et al., 2005; Liu et al., 2012; Wilson et al., 2016; Ciarkowska et al., 2018). Several enzymatic activities using β-glucogallin to synthesize gallotannins were observed in oak (*Quercus* spp.) (Niemetz and Gross, 2005). Moreover, β-glucogallin serves as a precursor for the biosynthesis of galloylated catechins, and it could be acylated with galloyl moieties in other species, such as tea plant, grapevine, and persimmon (*Diospyros kaki*). β-Glucogallin is formed from glucose and gallic acid in tea plants by the UDP-glucose:galloyl-1-*O*-β-D-glucosyltransferase CsUGT84A22 (Liu et al., 2012;

Cui et al., 2016), which homologs were also identified in pomegranate (*Punica granatum*) and *Eucalyptus camaldulensis* (Ono et al., 2016; Tahara et al., 2018). β-Glucogallin is also used for the biosynthesis of proanthocyanidins via SCPLs VvGAT1 and VvGAT2 in grapevine (Terrier et al., 2009; Carrier et al., 2013; Koyama et al., 2014). Differential expression of the DkSCPL1 and DkSCPL2 genes in persimmon was observed in astringent versus non-astringent fruits (Ikegami et al., 2007; Akagi et al., 2009). In tea leaves, galloylated catechins not only account for most soluble catechins, but are also the prominent contributors to flavor, such as astringency and bitterness, and have higher pharmacological activity for human health than non-galloylated catechins (Ikegami et al., 2007; Akagi et al., 2009; Hayashi et al., 2010; Wilson et al., 2016). β-Glucogallin (β-G) serves not only as an acyl donor in successive transacylation steps to yield hydrolysable tannins (HTs), such as substituted di- through penta-galloylglucose derivatives (Niemetz and Gross, 2005). These HTs may also be used as acyl donors for the generation of galloylated catechins (Bontpart et al., 2015; Wilson et al., 2016; Ciarkowska et al., 2018). Assays with β-G:1,2,3,6-tetra-*O*-galloyl- and β-D-glucose-4-*O*-galloyltransferase activities were also performed in pedunculate oak (*Quercus robur*), although their coding gene is not cloned (Grundhöfer and Gross, 2001).

Similar to HTs, galloylated catechins can be hydrolyzed to non-galloylated catechins and gallic acid by a galloylated catechins hydrolase, which has been widely characterized in bacteria at the catalytic as well as genetic levels (Niemetz and Gross, 2005; de las Rivas et al., 2019). Interestingly, no 1,2,3,4,6-pentagalloylglucose was identified in tea plants, suggesting that the hydrolyase activity toward this compound is very active. More than 30 gallic acid derivatives, including two types of HTs (gallotannins and ellagitannins), as well as several types of polygalloylated glucose (PGG), such as 1,6-digalloylglucose, 1,2,6-trigalloylglucose, 1-galloylglucose, 1,2-digalloylglucose, 6-galloylglucose, 2,6-digalloylglucose, 1,3-digalloylglucose, 1,2,3,4-tetragalloylglucose, and 1,4,6-trigalloylglucose have been identified in tea plant leaves (Yang and Tomás-Barberán, 2018; Wei et al., 2019).

Here, we explored the tea genome and identified novel putative SCPL genes involved in the generation of galloylated catechins in the leaves. We further characterized three SCPL enzymes, which produced *in vitro* galloylated catechins with PGG as an acyl donor and catechins as acceptor substrates. This study provides new insights into the understanding of this important gene family involved in the modification of tea catechins, which are of high relevance in the food chemistry and pharmacological industries.

MATERIALS AND METHODS

Identification of SCPL Family Genes in the *Camellia sinensis* Genome

BLASTP against the *Camellia sinensis* proteome dataset obtained from the genome sequence (Wei et al., 2018; The Tea Plant Information Archive – TPIA public database[1]) was used to search for SCPL-coding genes with characterized protein sequences as queries. The SCP motif (PF00450), was searched through Pfam in each retrieved sequence to confirm it as a putative SCPL protein. Features of SCPL proteins, such as the number of amino acid residues, gene architecture (exon/intron arrangement) and start-to-end position of their respective genes in the genome, were retrieved from the TPIA annotation. Physical parameters of each gene product, such as molecular weight (Mw) and isoelectric point (pI), were calculated using the ExPASy tool Compute pI/Mw[2] using default parameters.

Chromosomal Position, Phylogeny, and Gene Structure Analysis of SCPL Genes

PhenoGram Plot[3] used to create the image of scaffold locations of SCPL genes using the information available at the TPIA database. A phylogenetic analysis with functionally characterized proteins from different species and all SCPL proteins encoded in the *C. sinensis* genome was conducted to explore the evolutionary relationships of this gene family. An unrooted phylogenetic tree was constructed following the Neighbor-Joining method involving 1,000 replicates with the bootstrap test in MEGA 6.0 (Tamura et al., 2013). The Gene Structure Display Server v.2.0 tool (Hu et al., 2015[4]) was used to determine the exon/intron structures of the tea SCPL genes.

Calculation of K_a and K_s Ratios and Evolutionary Selection Analysis

Synonymous and non-synonymous substitution rates (K_s and K_a, respectively) of two closely related CsSCPL genes, along with their K_a/K_s ratios, were calculated online[5] to evaluate the selection mode of each CsSCPL paralogous gene pair identified. Assumptions of negative, neutral, and positive evolution were considered for ratios < 1, $= 1$, and > 1, respectively. The evolutionary signals of positive sites were determined with FEL, IFEL, REL and SLAC tests through DATAMONKEY[6]. For the evolutionary selection analysis, the phylogeny was reconstructed through the HyPhy program with CsSCPL cDNA sequences as input and the Nei-Gojobori method in MEGA 6. Maximum-likelihood based on the Kimura-2 parameter model was used to infer the evolutionary gene histories. The detection of episodic diversification of individual coding sites was performed through different approaches. Episodic diversifying selection and pervasive positive selection were identified at the individual branch site level using the Mixed-Effect Model of Evolution (MEME). The Markov Chain Monte Carlo (MCMC) method was used in the Fast, Unconstrained Bayesian Approximation analysis to ensure the

[1]http://teaplant.org/

[2]https://web.expasy.org/protparam/

[3]http://visualization.ritchielab.psu.edu/phenograms/plot

[4]http://gsds.cbi.pku.edu.cn

[5]http://services.cbu.uib.no/tools/kaks

[6]http://classic.datamonkey.org/

strength over predefined sites via the approximate Bayesian method (Murrell et al., 2012). The parameter $\omega = \beta/\alpha$ was used in MEME to fit the data in the GTR nucleotide model as initial values. The parameters $\beta{:}\beta^- \leq \alpha$ and β^+ were used to measure the selection pressure, whereas β^-, β^+ and α were used to estimate the variability of site-to-site substitution rate. The Likelihood Ratio Test (LRT) based on χ^2 asymptotic distribution was considered significant at p-value < 0.05.

Identification of Conserved Motifs and Promoter Region Analysis

The analysis of conserved regions within CsSCPL family genes was executed through MEME[7], and their genomic assemblies were screened through the Pfam database[8]. Regulatory elements in the 1.5-Kb promoter region upstream of the start codon for each CsSCPL gene were identified using the Plant *cis*-acting Regulatory DNA Elements (PlantCARE) program[9].

In silico Gene Expression Analysis

Public gene expression data of identified CsSCPL genes in seven tissues (apical vegetative bud; young, middle and old leaves; root; stem; and fruit: **Supplementary Dataset S3**) were retrieved from the microarray-based transcriptome tea data for the cultivar Shuchazao (Wei et al., 2018).

Plant Growth Conditions and Treatment Experiments

Tea (*Camellia sinensis* var. *sinensis* cv. Shuchazao) seeds germinated in vermiculite in a growth chamber at 22 ± 2°C. For the aluminum (Al^{3+}) treatment, one-year cuttings were hydroponically grown in Shigeki Konishi solution (**Supplementary Table S1**) under acidic conditions (pH 5.0, 0.4 mM Al^{3+}: control) or Al^{3+} exposure (pH 4.0, 2.5 mM Al^{3+}) in a growth chamber set with light period 16 h/8 h (light/dark) and temperature at 23°C. Root systems were collected at 0 h, 12 h, 24 h, and 48 h (Wang et al., 2017) for gene expression analysis. MeJA treatment experiments were followed as described previously (Shi et al., 2015): detached branches with tender tea shoots were sprayed with 100 μM Methyl Jasmonate (MeJA) solution, while the controls were treated with distilled water. Tea leaves were respectively collected at 0 h, 12 h, 24 h, and 48 h. The polyethylene glycol (PEG) and NaCl exposure experiments were performed as previously (Zhang et al., 2017): 25% PEG or 200 mM NaCl were used respectively to simulate plant drought- and salt-stress conditions for 0 h, 24 h, 48 h, and 72 h. For the cold treatments, tea plant leaves were collected during the cold accumulation (CA) process. Control: 25°C; CA1.6 h: 10°C for 6 h; CA1.7days: from 10°C to 4°C for 6 days; CA2.7days: from 4°C to 0°C for 7 days; DA.7d: recover under 25°C to 20°C for 7 days. For the shading treatment, tea plants cultivated in the 20-year-old tea garden in Anhui, where plants grow under natural condition (full sun exposure) or 70%–90% shading provided by a net cover. Buds were collected at 0 h, 4 h, 8 h, 2 days, 4 days, 8 days, and 14 days as previously reported (Liu et al., 2018).

RNA-Seq Data Analysis, and Gene Expression Analysis, and Validation With qRT-PCR

Tea plant tissues were sampled, rinsed immediately, snap-frozen in liquid N_2, and stored at −80°C for RNA extraction. Total RNA was performed with RNA extraction kits (BIOTECH, Beijing, China). The libraries were synthesized using Illumina HiSeq2500 library prep kits according to the manufacture's protocol. RNA-Seq (PE150bp) sequencing was performed on an Illumina HiSeq2500 platform. A total clean dataset of about 6 Gb (Q30 ≥ 80%) was obtained for each biological sample for analysis according to the tea plant genome sequence and annotation (Wei et al., 2018). Reads Per Kilobase of transcript per Million mapped reads (RPKM) and read counts were calculated using eXpress. The analysis of differential gene expression followed the method previously published (Ahmad et al., 2020).

Transcriptome data for the cv. Shuchazao were retrieved from the tea information archive[10]. Fragments per kilobase of exon per million fragments (FPKM) values were used to estimate gene expression in eight tea plant tissues (root, stem, old leaf, mature leaf, young leaf, apical bud, flower and fruit) and various treatments. The expression levels of *CsSCPL1A-AT* genes were calculated using Log_{10}(FPKM). The data were utilized to quantify the expression of *CsSCPL* genes in the roots of hydroponically growing tea cuttage seedlings in response to Al^{3+} and MeJA stresses (**Supplementary Dataset S3**). Mev4.9.0[11] was used to display the expression data in heatmaps. Expression and metabolite association analyses for *CsSCPL1A* genes potentially associated with galloylated catechins biosynthesis were carried out as in previous studies (Shi et al., 2015; Liu et al., 2018; Wei et al., 2018).

Pearson correlation analysis on *CsSCPL-I* gene expression and the contents of catechins from multiple tissue experimental datasets. R package also was used to evaluate the correlation between gene expression and metabolism and global gene expression. The resulting heatmap was structured by pHeatmap R package. Dark red means positive correlation values, and light red means negative correlation. Transcriptome and metabolic profiling data sets were from different tissues of tea plants (Wei et al., 2018).

Based on their expression dynamics, five tea CsSCPL-I genes [i.e., *CsSCPL2-I* (TEA009664, NCBI Genbank accession MK843824), *CsSCPL5-I* (TEA034028; MK843825), *CsSCPL11-I* (TEA023451; MK843826), *CsSCPL13-I* (TEA034055; MK843827), and *CsSCPL14-I* (TEA027270; MK843828)]

[7]http://meme-suite.org/

[8]http://pfam.xfam.org/

[9]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

[10]http://tpia.teaplant.org/

[11]https://sourceforge.net/projects/mev-tm4/

were selected for validation of their stress responses in different tissues via qRT-PCR. The gene-specific primers used listed in **Supplementary Dataset S4**. An iQ5 Real-Time PCR System (Bio-Rad) was used with 96-well plates and 20-μL reaction volume. Each reaction consisted of 2.5 μL 2X Power SYBR Master Mix (Applied Biosystems), 1 μL primer mix (0.4 μL of 10 mM each primer + 0.2 μL H_2O), and 2 μL cDNA reaction diluted 1:30. The gene expression was normalized using the housekeeping *CsACTIN* (TEA002341) gene.

Cloning and Expression of CsSCPL Recombinant Proteins and Enzyme Activity Assay

The open reading frames (ORF) of the five CsSCPL-I genes listed above were amplified with gene-specific primers. Amplification was performed on a Mastercycler PCR equipment (Eppendorf) using 0.5 μL ExTaq DNA polymerase (Takara) and 2.0 μL cDNA dilutions (1:30) as templates. The gel-purified PCR products inserted into the pGEM-T easy vector (Promega), cloned into *E. coli*, and confirmed by sequencing. For expression of functional SCPL proteins, the truncated CsSCPLs, encoding by the ORFs of CsSCPL11-IA and CsSCPL14-IA with a deletion of the first 150 bp and an ATG codon insertion and full length CsSCPL13-IA (**Supplementary Figure S2B**), were cloned into pDONR221 (Invitrogen) followed by recombination into pDEST17 with LR Recombinase (Invitrogen). The constructs transferred into *E. coli* strain Rosetta for heterologous protein induction. The single colony of confirmed construct taken and overnight grown in 10 mL LB media with ampicillin (50 μg mL^{-1}) at 37°C. Bacterial cultures grown overnight were diluted in 300 mL LB media (1:100) with ampicillin (50 μg mL^{-1}) and kept in a shaker at 37°C. When the OD_{600} reached 0.8 to 1.0, 0.2 mM isopropyl 1-β-D thiogalactoside (IPTG) was added to the culture and kept at 25°C. After a 10-h incubation period, the cultures were collected, centrifuged at 12,000 rpm for 30 min and re-suspended in 20 mL lysis buffer [200 mM Tris–HCl (pH 8.0), 0.1% Triton X-100, and 5 mM β-mercaptoethanol] and kept on ice for 1h. The cells were then broken through ultrasonication and centrifuged at 12,000 rpm for 30 min at 4°C. Supernatants were collected and purified using a nickel-resin purification kit (Promega). Protein extracts were examined through sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS–PAGE).

For the enzyme activity assays, 50-μl total reaction mixtures containing 50 mM potassium phosphate buffer (pH 6.0), 1.0 mM catechin (C), epicatechin (EC), gallocatechin (GC), epigallocatechin (EGC), 0.4 mM 1,2,3,4,6-penta-*O*-galloyl-β-D-glucose (PGG), and 1.0 μg purified recombinant enzyme were used to verify the gallate activity of each selected CsSCPL. The reactions were carried out at 30°C for 20 mins and stopped by adding an equal amount of 100% methanol. The samples were analyzed by high-performance liquid chromatography (HPLC) and liquid chromatography-mass spectrometry (LC-MS), as described previously (Wei et al., 2018).

Analyses of Enzyme Kinetics

The K_M and V_{Max} kinetic parameters were obtained through Lineweaver-Burk plots for *CsSCPL11-IA*, *CsSCPL13-IA*, and *CsSCPL14-IA* with various concentrations of EC and EGC as acceptors (0.5 to 50 mM) and PGG (0.4 mM) as the donor substrate, in a total reaction volume of 50 μL. The PGG kinetic parameters for these recombinant proteins were also measured with the PGG (0.1 to 10.0 mM) as the acyl donor and EGC (fixed concentration at 1.0 mM) as the acceptor substrate. The reactions were analyzed on HPLC.

Statistical Analyses

At least 3 biological replicates were used in this study with at least 2 repetitions to obtain the data for analysis. Differences between paired data from the enzymes and controls under various conditions were analyzed via ANOVA followed by the Student's *t*-test ($n = 3$).

RESULTS

SCPL Gene Mining in the *Camellia sinensis* Genome

A total of 47 SCPL-related genes were identified in the *C. sinensis* genome and annotated according to their distribution in different phylogenetic classes (**Supplementary Dataset S1**). Their coding regions ranged from 717 bp (TEA034030) to 1,758 bp (TEA034034) with an average of 1,369 bp. The molecular weight of the CsSCPL gene products ranged between 27.49 kD (TEA034030) and 73.8 kD (TEA034056) and averaged 51.41 kD. The CsSCPL pI values ranged from 4.67 (TEA010715 and TEA016469) to 9.97 (TEA034027), with 77% (36/47) having acidic pI values. The subcellular localization of a protein influences its function by controlling the ability to obtain and use all types of molecular interacting partners. So, protein localization is an important piece of information in creating hypotheses about cellular functions of newly discovered proteins (Scott et al., 2005). About 61% of the CsSCPL-I proteins are predicted to localize to the plasma membrane, whereas 65% of the CsSCPL-II proteins are estimated to localize to the lysosome. Meanwhile, the remaining members of CsSCPL-I and CsSCPL-II clades are estimated to be found in other organelles, such as the lysosome, nucleus and cytoplasm, as well as extracellularly (**Supplementary Dataset S1**). Details on the products of each CsSCPL gene have shown in **Supplementary Dataset S1**.

Phylogeny and Motif Analysis of CsSCPLs

A phylogenetic tree was constructed from tea protein sequences as well as functionally characterized SCPL sequences from other plant species and the SCPLs from rice and Arabidopsis genomes in order to evaluate the evolutionary relationships among them (**Figure 1**). According to the classification and structural features of SCPL proteins from previous studies in rice and Arabidopsis models, CsSCPL genes are divided into

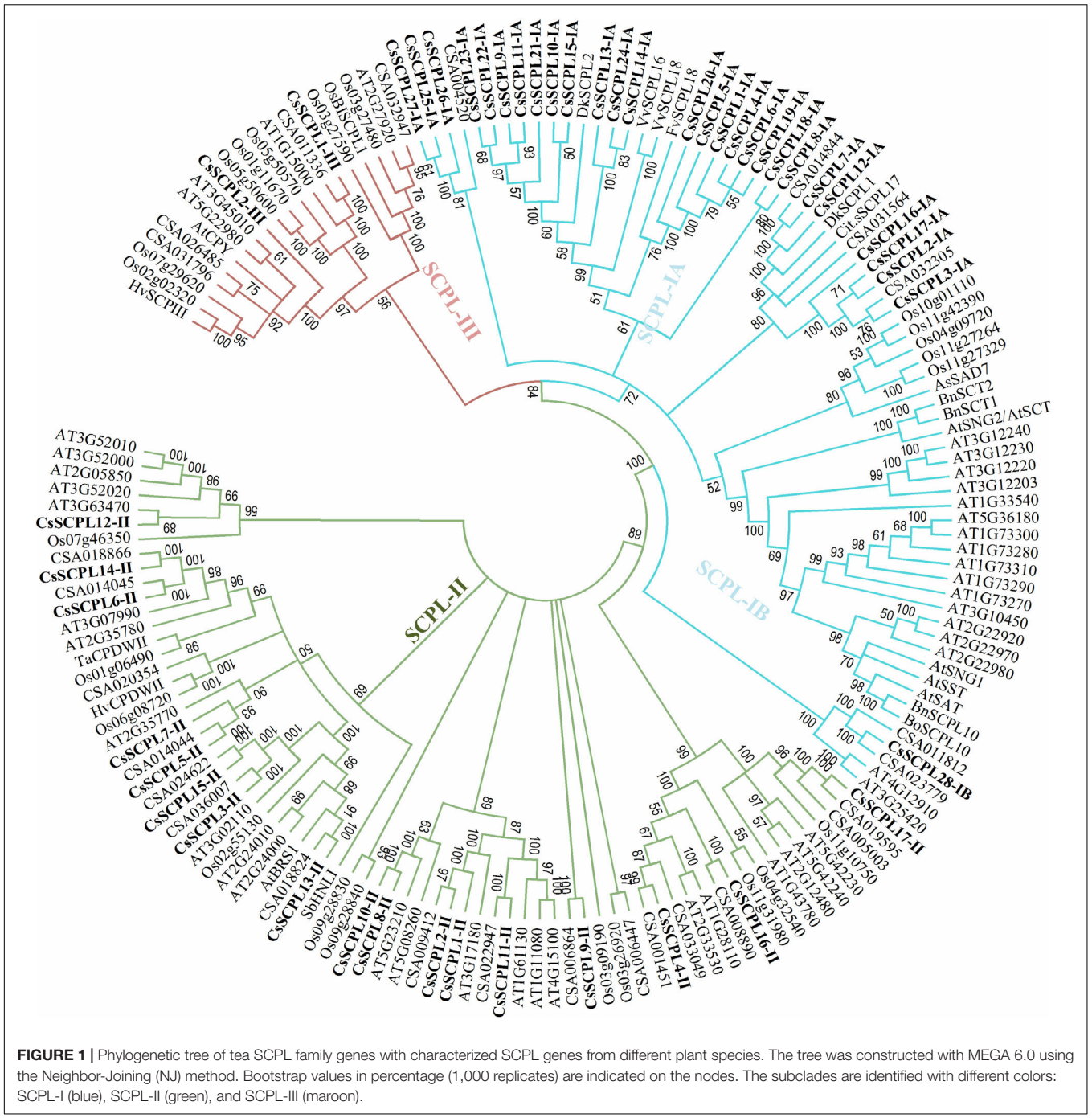

FIGURE 1 | Phylogenetic tree of tea SCPL family genes with characterized SCPL genes from different plant species. The tree was constructed with MEGA 6.0 using the Neighbor-Joining (NJ) method. Bootstrap values in percentage (1,000 replicates) are indicated on the nodes. The subclades are identified with different colors: SCPL-I (blue), SCPL-II (green), and SCPL-III (maroon).

three classes: CsSCPL-I to CsSCPL-III. CsSCPL-I is further split into two subclasses: CsSCPL-IA and CsSCPL-IB (**Figure 1**). The annotation of all 47 tea SCPLs was carried out according to their phylogenetic position: 27 fell into the CsSCPL-IA clade (CsSCPL1-IA to 27-IA) whereas only one member fell into CsSCPL-IB (CsSCPL28-IB). The CsSCPL-II clade contains 17 tea proteins (CsSCPL1-II to 17-II), and CsSCPL-III is limited to only two members (CsSCPL1-III and 2-III) (**Figure 2A**). Tea CsSCPL-I members intermixed with many SCPLs from other species but diverged from most Arabidopsis SCPLs. The conserved domains of CsSCPL proteins have determined by MEME (**Figure 2B**). In total, ten motifs have found, and their annotations confirmed through the Pfam and SMART databases (Letunic et al., 2012). The first four motifs are specific to serine carboxypeptidase enzymes and found in virtually all CsSCPLs (**Supplementary Table S2**; **Figure 2B**). All CsSCPLs motifs were generally well conserved but especially similar within the same phylogenetic class. In addition, each clade contained unique motifs – for example, motifs 5 and 10 were only found in the CsSCPL-I clade, which suggests

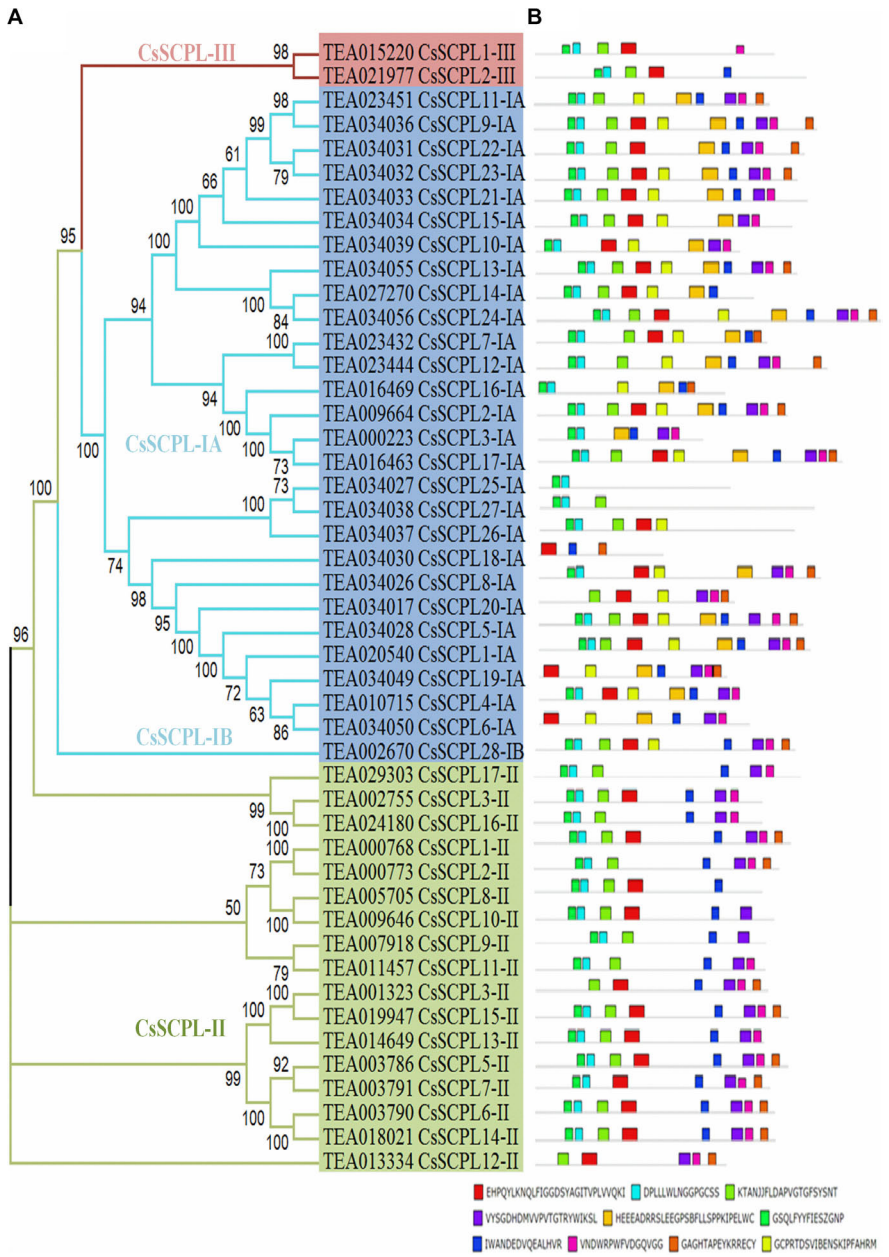

FIGURE 2 | Phylogenetic analyses and conserved motif analysis of tea CsSCPL genes. **(A)** The phylogenetic tree of CsSCPLs genes was constructed through MEGA 6.0 using the Neighbor-Joining (NJ) method. Bootstrap values in percentage (1,000 replicates) are indicated on the nodes. The subclades are identified with different colors: SCPL-I (blue), SCPL-II (green), and SCPL-III (maroon). **(B)** The identified motifs of ScSCPL proteins are shown as colored boxes. The CsSCPL proteins are listed according to their phylogenetic relationships.

that they have distinct features and potential specific functions related to each class.

Exon/Intron Organization Within CsSCPL-I Genes

The exon/intron organization analysis of CsSCPL-I genes was performed to understand their structural diversity (**Figure 3A**). Extensive variation was found in exon numbers within the CsSCPL-I class, ranging from 6 (*CsSCPL18-IA* and *20-IA*) to 14 exons (*CsSCPL13-IA* and *23-IA*), with an average of 10 exons per gene (**Supplementary Dataset S1**; **Figure 3A**). About 11% (3/28) of the CsSCPL-I genes has 8 exons, 21% (6/28) contains 9 exons, 14% (4/28) has 10 exons, another 21% (6/28) shows 11 exons, while the remaining (40%, 11/28) contains 12 to 14 exons (**Supplementary Dataset S1**; **Figure 3A**). The comparison of gene architectures between each of the five CsSCPL paralogous pairs identified revealed that only one pair (CsSCPL27-IA/25-IA) contained the same number of exons.

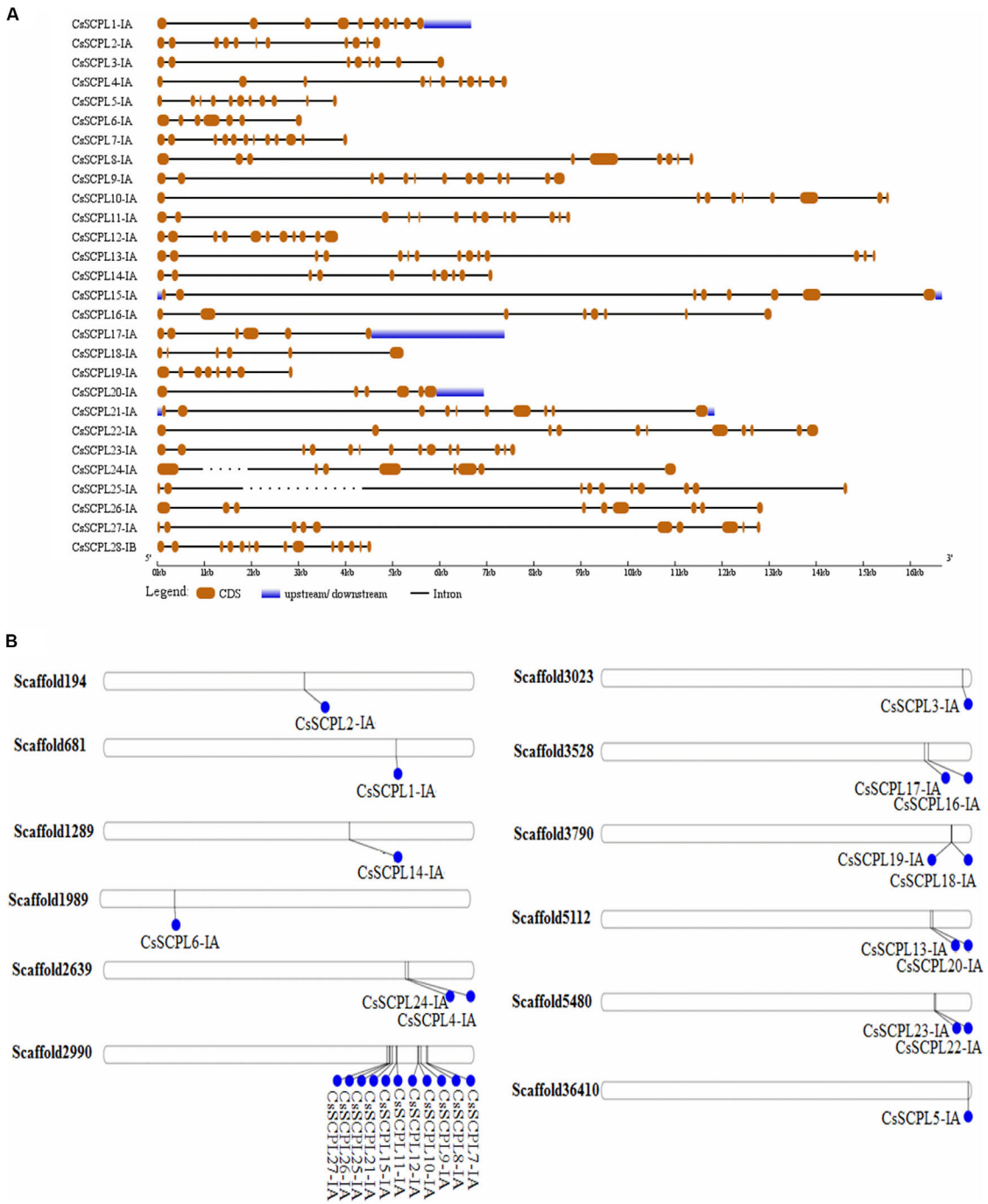

FIGURE 3 | Gene architecture of tea CsSCPL genes. **(A)** The gene structures of CsSCPLI were plotted using red boxes representing exons, black lines representing introns, and blue boxes indicating UTR sequences. The scale in the bottom is in the unit of kilobase (kb). **(B)** Scaffold locations and duplications of CsSCPL-I genes. Scaffold numbers are indicated on leftside of each respective scaffold. The different genes located according to their presence in bps.

Chromosomal Distribution and Evolutionary Analysis of CsSCPL-I Coding Sequences

The 27 CsSCPL-IA genes were unevenly distributed on twelve different scaffolds of the tea genome assembly (**Figure 3B**). Interestingly, while most scaffolds contained one or two CsSCPL-I genes, scaffold2990 contained 11, with the majority clustering together in the phylogenetic tree (**Figure 1**). The expansion of gene families and their neo/sub-functionalization is a common phenomenon in plants and occurs through both, segmental as well as tandem gene duplication (Zhang and He,

2005; Kong et al., 2007). From the phylogenetic analysis, we conclude that the primary method for gene duplication of the CsSCPL-IA clade was segmental duplication while a single locus holds tandem duplications. We also found tandem duplications in two paralogous pairs (*SCPL9-IA/21-IA* and *12-IA/-IA*) along with segmental duplications in five pairs (*SCPL16-IA/20-IA, 13-IA/24-IA, 22-IA/10-IA, 6-IA/14-IA*, and *19-IA/1-IA*) (**Figure 3B**).

The selection history of coding sequences can be assessing through the K_a/K_s ratio. In order to investigate the divergence of duplicated CsSCPL-IA members, K_a and K_s values, as well as the K_a/K_s ratios, were determined for each CsSCPL-IA paralogous pair. K_a/K_s ratios < 1, > 1, or $= 1$ indicate that the paralogous pair are respectively under purifying (negative), positive, or neutral selection (Juretic et al., 2005). The K_a/K_s ratios of the seven CsSCPL-I paralogous pairs were between 0.88 and 1.67 (**Supplementary Table S3**), suggesting that distinct selection forces are acting on each pair. The K_a/K_s ratios of tandemly duplicated gene pairs were 0.88 and 1.31 (**Supplementary Table S3**), indicating that these pairs are also undergoing different selection pressures. On the other hand, all segmentally duplicated paralogous pairs are under strong positive selection pressure (K_a/K_s between 1.18 and 1.67), except for *SCPL6-IA/14-IA* and *19-IA/1-IA*, which are under purifying selection pressures (Ka/Ks 0.89 and 0.88, respectively) (**Supplementary Table S3**).

In order to maintain protein structure and function, conservation of amino acid positions during evolution is essential. Exploring this feature may clarify the selection pressures at work. The variation in substitution rates estimated by the MEC model adaptive selection test revealed that several SCPL protein regions are under positive selection (**Supplementary Figure S1A**). Over 5% (26 out of 476) CsSCPL-I amino acid residues are under positive selection, whereas the remaining are under purification (**Supplementary Figure S1A**). For CsSCPL-II proteins, all residues are under purifying selection (**Supplementary Figure S1B**).

For further corroboration of the selection pressures, we tested additional models. CsSCPL-I and CsSCPL-II members show different average ω ratios across their coding sites. The ω ratio output was CsSCPL-I > 1, and CsSCPL-II $= 1$ (**Supplementary Table S4**). This result suggests that most of the CsSCPL-I coding sites were positively selected and contain conserved amino acid residues exposed to purifying selection. Comparison of the M7 and M8 models was performed to refine the selection test. M8 fit the data more significantly than M7. Both classes, CsSCPL-I and CsSCPL-II, displayed positively selected sites in M8 for 7% of the sites under a ω value of 6.2 for CsSCPL-I, and a very minute proportion of 0.001% of CsSCPL-II sites was under a ω value of 1.0 (**Supplementary Table S4**). Therefore, the number of positively selected sites were significantly higher in the M8 model for CsSCPL-I than CsSCPL-II. Different likelihood-based approaches disclosed that 19 coding sites of CsSCPL-I genes but only 3 sites of CsSCPL-II genes were under positive selection (**Supplementary Table S4**). Various tests (FEL, IFEL, REL and SLAC) were used to identify evolutionary signs of positive selection through the computation of ω values. Each analysis detected a different number of coding sites under positive selection: for CsSCPL-I genes, FEL, IFEL, REL and SLAC revealed six, ten, fourteen, and seven coding sites respectively. Meanwhile, the respective numbers were seven, eleven, one, and four sites for CsSCPL-II genes (**Supplementary Table S4**). All of these tests were significant at a P-value < 0.05, except for REL, which detected sites under positive selection through a Bayes factor > 20. Therefore, reliable indication of positive selection of many CsSCPL-I gene sites and only a few of CsSCPL-II was consistently found throughout these analyses.

Positive Selection Based on Amino Acid Positions

The identification of evolutionarily conserved amino acid positions has a vital role in understanding protein structure and function. The Bayes method was used to reveal functionally relevant sites through posterior probabilities. Amino acid positions with $\omega \geq 1$ indicate positive selection. The average length of 1,377 amino acid residues of CsSCPL-I was analyzed in BEB, which revealed 19 sites under positive selection. Meanwhile, only three positions were under positive selection for the CsSCPL-II average length of 1,370 amino acid residues (**Supplementary Table S5**). This result suggests that the identification of coding sites may explain the selection pressures and points out to functionally relevant amino acid residues in the SCPL proteins.

Cis-Element Analysis of *CsSCPL-I* Gene Promoter Sequences

The exact functions played by CsSCPL-I proteins in plants remain ambiguous. Characterizing the expression of a given gene in time and space is particularly useful to define its function. The dynamics of gene transcription in each cell as responses to environmental stimuli as well as internal cues are ultimately controlled by a modular composition of *cis*-elements present in gene promoter regions. Promoter sequences of *CsSCPL-I* genes were extracted and submitted to PlantCARE for *cis*-regulatory element identification (Lescot et al., 2002). Numerous elements were found in the 1.5-Kb upstream region of CsSCPL-I genes (**Figures 4A–D**; **Supplementary Dataset S2**). Twenty-six out of the 28 *CsSCPL-I* genes contained promoters with elements responsive to light (about 40% of the genes), hormones (15%), environmental stresses (25%), and plant growth (20%) (**Figure 4A**). Moreover, we identified binding sites for transcription factors that regulate responses to several hormones (**Figure 4B**). Among the hormone-response elements, these sensitive to GA were the most abundant in *CsSCPL-I* promoters, followed by MeJA. In addition to light as the most abundant factors may affecting *CsSCPL-I* genes, other abiotic and biotic stresses also impact on their expression significantly (**Figure 4C**). The maximum number of *cis*-elements related to plant growth found in the analysis was associated with the endosperm (**Figure 4D**).

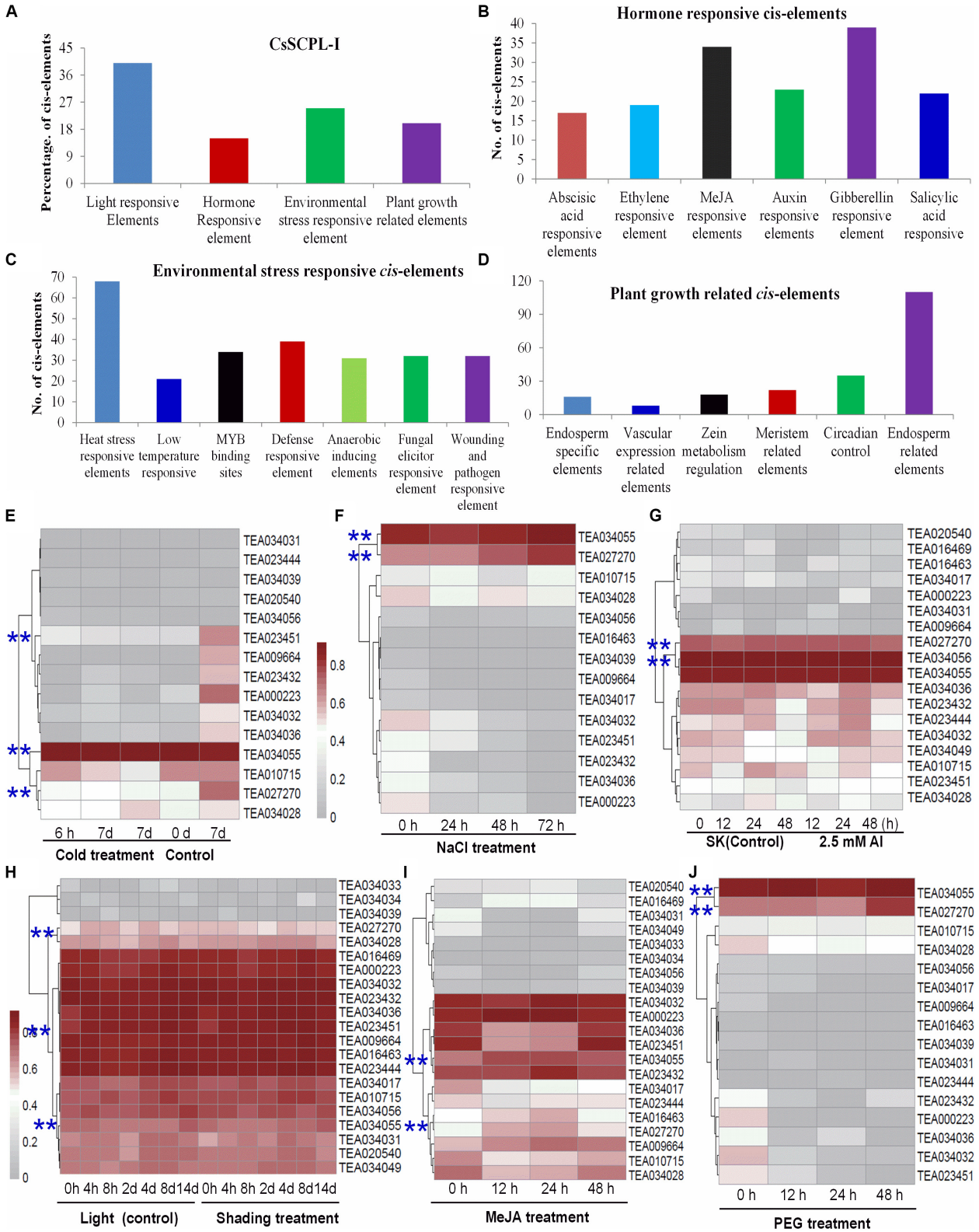

FIGURE 4 | Environmental and hormonal regulation of CsSCPL genes in tea plants. PlantCare were used to analyze the 1,500 bp upstream region of each *CsSCPLI* gene. **(A)** The percentage of light responsive elements, hormone responsive elements, environmental stress related elements, and plant growth responsive elements

(Continued)

FIGURE 4 | Continued
in all *CsSCPLI* family members. **(B)** Different hormone (ABA, ethylene, MeJA, auxin, gibberellinlic acid) responsive elements in *CsSCPLI* genes *cis*-element regions. **(C)** Different environmental stress (heat, cold and dehydration, drought, defense, anaerobic, wound and pathogen) related elements in *CsSCPLI* genes *cis*-element regions. **(D)** Different plant growth related elements in *CsSCPLI* genes *cis*-element regions. **(E–J)** Heatmaps displaying expression patterns of various CsSCPL1A-AT genes under various conditions. Transcriptome data from experiments with tea cv. Shuchazao were retrieved from the tea plant information archive (http://tpia.teaplant.org/index.html). The expression levels of CsSCPL1A genes were normalized as fragments per-kilobase of exon per million fragments (FPKM) in eight tea plant tissues (root; stem; old, mature, and young leaves; apical bud; flower; and fruit) and displayed as Log_{10}(FPKM) in heatmaps using Mev4.9.0 (https://sourceforge.net/projects/mev-tm4/).

The distribution pattern of *cis*-elements differed among *CsSCPL-I* members. For example, *CsSCPL7-IA* and*12-IA* promoters contained most ERE-related elements (five), whereas 62% of the *CsSCPL-I* gene promoters contained none (**Supplementary Dataset S2**). Remarkably, the single motif responsive to ABA was found in the *CsSCPL24-IA* promoter. The greatest number of heat-shock elements (eight) was found in *CsSCPL9-IA* and*22-IA*, followed by six in *CsSCPL11-IA*. A total of six defense-related elements was found in the *CsSCPL13-IA* promoter region (**Supplementary Dataset S2**). *CsSCPL11-IA, 13-IA,* and *14-IA* were differentially expressed in leaves or roots under the various treatments (**Supplementary Dataset S3**). Shading leads to a significant decreased in several galloylated catechins, such as EGCG and GCG (Liu et al., 2018). *CsSCPL13-IA*, *24-IA*, and *14-IA* expression levels were noticeably repressed in shaded leaves compared to plant under full light (**Supplementary Dataset S3**). However, most of the other SCPL genes were unchanged or even induced, such as *SCPL11-IA* (**Figure 4**; **Supplementary Dataset S3**). This result indicates that *CsSCPL11, 13, 14*genes are most likely responsible for the decreased levels of galloylated catechins under shading (Liu et al., 2018). Moreover, both MeJA and NaCl treatments promoted EGCG accumulation in the leaves or roots (Shi et al., 2015; Zhang et al., 2017). Our experiments showed that both of these conditions substantially induced the expression of *CsSCPL13-IA* and *14-IA* (**Figure 4**; **Supplementary Dataset S3**). These genes were also activated by PEG and Al exposure in roots (**Figure 4**), which is consistent with increased galloylated catechins previously reported (Zhang et al., 2017).

CsSCPL Gene Expression Profiling

We first examined the expression pattern of the 47 CsSCPL genes identified by analyzing publicly available RNA-Seq data from eight *C. sinensis* tissues (apical bud; flower; fruit; young, mature, and old leaves; root; and stem) (**Figure 5A**, **Supplementary Dataset S3**). Most genes showed expression in almost all tissues, with a few exceptions. About 25% (12/47) genes showed no or very low expression in all tissues. These genes may be induced by particular conditions or undergoing pseudofunctionalization. Additionally, *CsSCPL2-IA, 3-IA, 22-IA*, and *23-IA* were only expressed in the apical bud, young leaf, and stem. On the other hand, *CsSCPL5-IA, 11-IA, 13-IA, 14-IA,* and *24-IA* were highly expressed in all tissues analyzed (**Figure 5A**). The comparative expression analysis of paralog pairs showed conflicting results. Whereas three of the identified pairs showed very similar expression patterns, each member of the pairs Cs*SCPL13-IA/24-IA, 22-IA/10-IA*, and *19-IA/1-IA* showed distinct expression patterns (**Figure 5A**, **Supplementary Dataset S3**). This result is evidence that each paralogous member underwent or still are undergoing functional divergence.

Pearson correlation analysis between *CsSCPL-I* gene expression levels and the contents of catechins from multiple tissue experiments (Wei et al., 2018), showed that the experession of most *CsSCPL-I* genes was correlated with at least one of these catechins, indicating the redundant functions of *CsSCPL-I* genes on catechins biosyntesis (**Figure 5B**). Based on tissue-specific expression patterns and presence of promoter *cis*-elements responsive to different stresses, five CsSCPL-I genes were chosen for further analysis via qRT-PCR in seven tissues: *CsSCPL2-IA*, *5-IA*, *11-IA*, *13-IA*, and *14-IA* (**Supplementary Dataset S3**). *CsSCPL2-IA* and *11-IA* expressed relatively high in the apical bud and the young leaf (**Figure 5C**). *CsSCPL13-IA* expression was relatively constant across all tissues, whereas *CsSCPL14-IA* expressed very low in all tissues tested. Since our promoter analysis revealed several *cis*-elements associated with MeJA and stress, we examined gene expression under aluminum and MeJA exposures (**Figures 5D–E**). *CsSCPL13-IA* and *14-IA* were markedly induced by Al^{3+} treatment, whereas only low expression levels of *CsSCPL2-IA, 5-IA*, and *11-IA* were detected (**Figure 5D**). *CsSCPL2-IA* expression increased gradually after 12, 24, and 48 h under MeJA treatment whereas *CsSCPL11-IA, 13-IA*, and *14-IA* increased up to 24 h and decreased thereafter as compared to the control (**Figure 5E**). On the other hand, *CsSCPL5-IA* expression was remarkably reduced upon exposure to MeJA compared to the control. The *CsSCPL11-IA* expression maximum was observed under MeJA exposure.

Cloning and Characterization of *CsSCPL-I* Genes

The five CsSCPL-I genes chosen above for expression analysis were cloned in order to produce recombinant proteins in *E. coli*. However, only three of them (*CsSCPL11-IA, CsSCPL13-IA*, and *CsSCPL14-IA*) expressed successfully in *E. coli*, and therefore chosen for further functional characterization. CsSCPL11-IA, 13-IA and 14-IA proteins contain respectively 450, 498 and 413 amino acid residues, and their calculated molecular weights are 51, 56, and 46 kD (**Supplementary Dataset S1**). They contain the conserved serine carboxypeptidase motif (**Figure 6A**) and are generally highly similar to other plant SCPLs. As a typical example

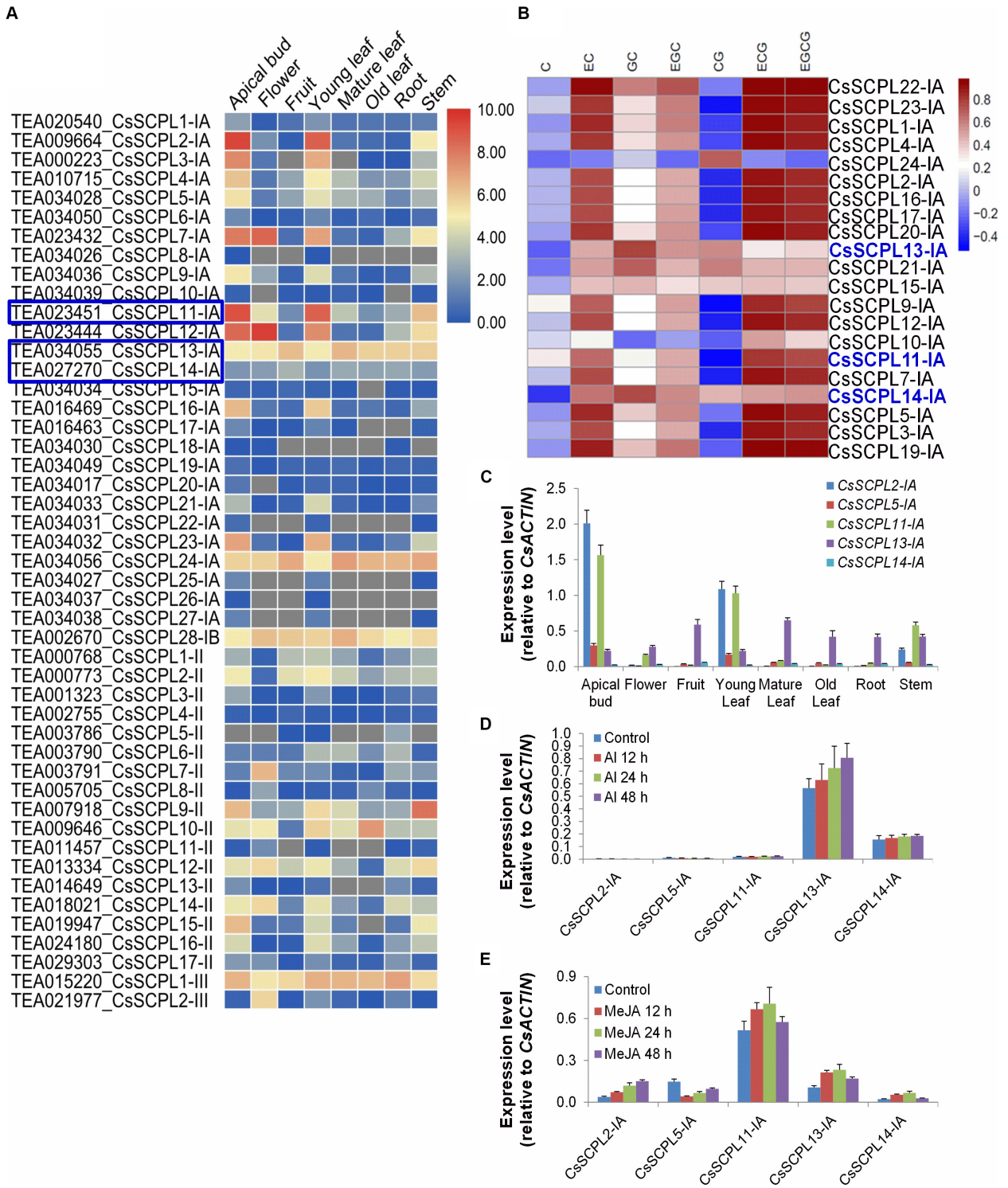

FIGURE 5 | Expression patterns of CsSCPL genes in different Tea tissues and qRT-PCR confirmation. **(A)** The expression levels of CsSCPL genes in different tea tissues retrieved from a public database (Tea Plant Information Archive, http://teaplant.org/) and represented by constructing the heat map using TBTools program. **(B)** Correlation analysis of the contents of catechins of non-galloylated or gallolated in tea plant tissues with the expression levels of CsSCPL-IA genes. *CsSCPL-I* expression data and catechins contents from multiple tissue experiments (Wei et al., 2018). R package was used to evaluate the correlation. Color bar shows Pearson coefficiency values. **(C–E)** Expression of selected CsSCPL-I genes via RT-qPCR in different tea tissues **(C)**, and in response to Al^{3+} **(D)** and MeJA **(E)** exposures. Analytical data were obtained from atleast three biological replicates and are expressed as means ± SD.

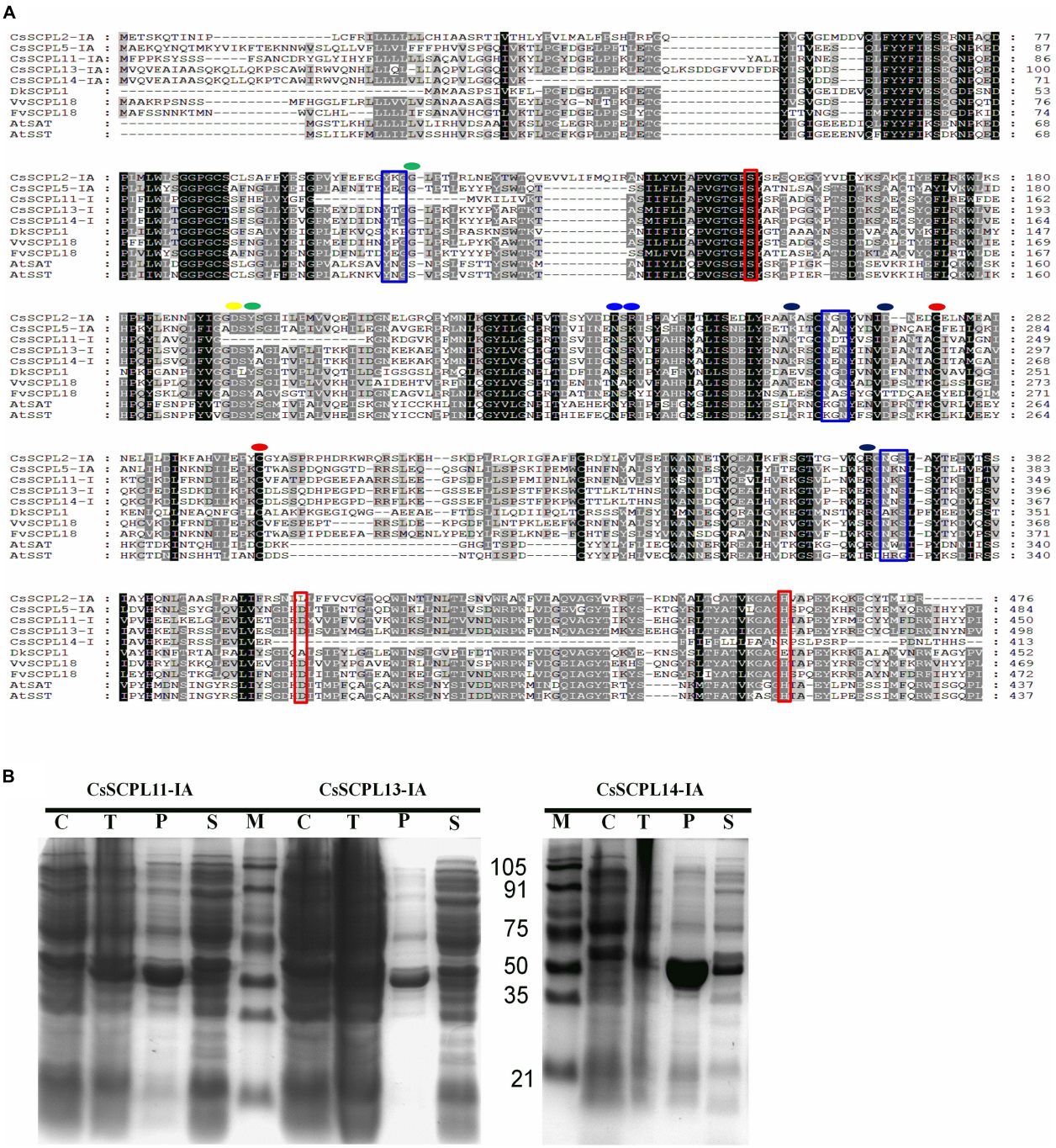

FIGURE 6 | (A) Aligment of CsSCPL-I with characterized SCPLs from other species. Conserved amino acid residues participating in key roles are indicated: Cys residues that are likely to form the intersubunit disulphide bond (red circles), oxyanion hole (green circles), recognition of the sinapoyl moiety of the donor molecule (blue circles), the hydrogen bond network (yellow circle), the primary L-malate recognition (black circle), amino acid tracts corresponding to N-glycosylation sites (blue squares), and the amino acid residues forming the catalytic triad of serine carboxypeptidases (S, D, H) (red squares) (cf. Milkowski and Strack, 2004; Chiu et al., 2016). **(B)** The purified His-tagged CsSCPL11-IA (truncated), 13-IA (complete) and 14-IA (truncated) protein fusion recombinants expressed in *E. coli* and partially purified with nickel resin. Proteins were resolved on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), followed by Coomassie Blue R-250 staining.

in clade IA, the *CsSCPL17-IA* coding region shows > 50% identity with *AtSAT, AtSST, DkSCPL, VvSCPL18,* and *FvSCPL18*. On the other hand, the *CsSCPL14-IA* protein sequence showed < 50% similarity with *AtSAT, AtSST*, and *CsSCPL17-IA* (**Figure 6A**, **Supplementary Dataset S5**). Hydropathy analysis revealed that CsSCPL proteins possess a strong hydropathic

region near their N-termini, which indicates the presence of a signal peptide.

CsSCPL-I Proteins Convert Epicatechin (EC) and Epigallocatechin (EGC) to Their Gallate Forms

Among the aforementioned five genes selected and expressed in His-tagged fusion proteins in *E. coli* strain BL21 (DE3), only CsSCPL13-IA and the truncated CsSCPL11-IA and 14-IA with their *N*-terminal 50 amino acids deleted were successfully expressed and purified with nickel-resin column. The reason for this deletion is that CsSCPL11-IA and 14-IA contain an *N*-terminal transmembrane domain (**Figure 6B**, **Supplementary Figure S2A**). All recombinant proteins successfully expressed showed maximum activity at pH 6.0 with 1,2,3,4,6-pentagalloylglucose (PGG) as an acyl-donor and catechin (C), epicatechin (EC), or gallocatechin (GC) as acceptor substrates. We also used PGG as a donor and C, GC, EC, or EGC as acceptor substrates to assay their activities with partially purified recombinant enzymes. CsSCPL11-IA, 13-IA, and 14-IA preferred PGG as the galloyl donor to convert EC or EGC to ECG or EGCG, respectively (**Figure 7**, **Supplementary Figures S3**, **S4**). High-pressure liquid chromatography (HPLC) coupled with tandem mass spectrometry was used to analyze the reaction and confirm the reaction product with authentic standards (**Figure 7A**, **Supplementary Figures S3**, **S4**).

The kinetic parameters of these three enzymes were calculated through the Lineweaver–Burk plotting method. All of them showed the same trends toward substrate specificity, although the values of kinetic parameters were quite different. CsSCPL13-IA showed higher K_M (160.41 μM) and V_{Max} (37.04 μmol mg^{-1} min^{-1}) for EC than the other enzymes (**Table 1**). Meanwhile, CsSCPL14-IA showed maximum K_M (103.38 μM) and V_{Max} (62.50 μmol mg^{-1} min^{-1}) for EGC (**Table 1**) at the saturated substrate concentration. The maximum specificity constant (K_{cat}/K_M) was observed for CsSCPL14-IA with EC as the substrate (5.19 s^{-1} μM^{-1}) followed by CsSCPL13-IA (3.85 s^{-1} μM^{-1}) and CsSCPL11-IA (3.14 s^{-1} μM^{-1}), meaning that CsSCPL14-IA has a higher binding affinity for EC. Similar trend was observed for EGC, which showed maximum K_{cat}/K_M (10.08 s^{-1} μM^{-1}) for CsSCPL14-IA, followed by CsSCPL13-IA (6.84 s^{-1} μM^{-1}) and CsSCPL11-IA (5.39 s^{-1} μM^{-1}). However, the K_M of CsSCPL13-IA for PGG was lower than the other two enzymes. The maximum K_M (20.81 μM) and V_{Max} (10.64 μmol mg^{-1} min^{-1}) values for PGG were observed for CsSCPL14-IA when EGC was used as acceptor (**Table 1** and **Figure 8**). The K_{cat}/K_M of CsSCPL14-IA was higher for PGG, suggesting that EGC is the preferred substrate for this enzyme.

DISCUSSION

A large eudicot family of serine carboxypeptidase-like (SCPL) acyltransferases was originally recruited from a more ancient serine carboxypeptidase family and adapted to take over acyl transfering functions. SCPL acyltransferases have a catalytic triad formed by a nucleophile, an acid and a histidine residue acting as a charge relay system for the nucleophilic attack on amide or ester bonds (Milkowski and Strack, 2004). Although BAHD enzymes have been known for a long time to acylate anthocyanins and flavan-3-ols (Zhao, 2015), the molecular identities of the enzymes catalyzing transacylations from 1-*O*-β-glucose esters or their involvement in the biosynthesis of varied phenolic compounds were not described until the 1990s (Fujiwara et al., 1998; Niemetz and Gross, 2005). Serine carboxypeptidases (S) catalyze the C-terminal peptide bond in proteins and make the Ser-Asp-His catalytic triad. Many studies showed that SCPL proteins share high sequence similarity with the α/β hydrolase (SCP) family, but they did not have the same hydrolase function as SCPs (Milkowski et al., 2004). Instead, SCPLs show acyltransferase and lyase activities (Li and Steffens, 2000; Shirley et al., 2001). For example, the acyltransferase activity of SCPLs catalyzes the formation of the acylsugar 2,3,4-isobutyryl-glucose (Lehfeldt et al., 2000), while sinapate ester is synthesized in *A. thaliana* by the activity of sinapoyl-glucose:malate sinapoyltransferase and sinapoylglucose:choline sinapoyltransferase (Fraser et al., 2005). The present study is the first analysis that provides comprehensive details of SCPLs at the genomic, evolutionary, and catalytic levels in *C. sinensis*.

Genome-Wide Analysis of SCPL Gene Family in Tea Plant Genome

SCPL acyltransferases play critical roles in many crops, and several of them have been identified and characterized, such as in barley (Baulcombe et al., 1987), rice (Washio and Ishikawa, 1994), pea (Jones et al., 1996), *Arabidopsis* (Li et al., 2001), tomato (Moura et al., 2001), persimmon (Ikegami et al., 2007), and poplar (Zhu et al., 2018). However, so far, not a single study has reported a systematic analysis of SCPLs in *C. sinensis*. Herein, we report a genome-wide identification of SCPL family members, gene expression analysis, and enzymatic assays of select members to explore their substrate specificities and provide clues about their potential functions. We found 47 CsSCPLs distributed into three main phylogenetic classes. Considering genome sizes and ancestral genome duplication events, a ratio of 2:4 SCPL genes is expected between the poplar and *Arabidopsis* genomes (Tang et al., 2008). However, SCPL genes were distributed in these species at the ratio 2:1 in each of the phylogenetic classes (Zhu et al., 2018). This result suggests that SCPL genes were lost in the Arabidopsis genome. Similarly, in our study, the ratios of three SCPL classes in *C. sinensis* and *Arabidopsis* were 1:0.7, 1:1.3, and 1:2.5. The high ratio displayed in *C. sinensis* for the CsSCPL-I class is significant and indicative of genome duplication events. The paralogous pairs, *CsSCPL16-IA/20-IA*, and *12-IA/8-IA* show K_s values of 0.21 and 0.36, respectively (**Supplementary Table S3**), which is similar to the with values of the salicoid lineage duplication event (0.27) that occurred 35 MYA (Guo et al., 2014; Wei et al., 2018). Meanwhile, the other paralogous pairs have lower K_s values and might have originated from recent tandem replication events (Wei et al., 2018). We did not find the evidence of CsSCPL-I paralogous pairs deriving from the ancient γ genome triplication event due to the absence of

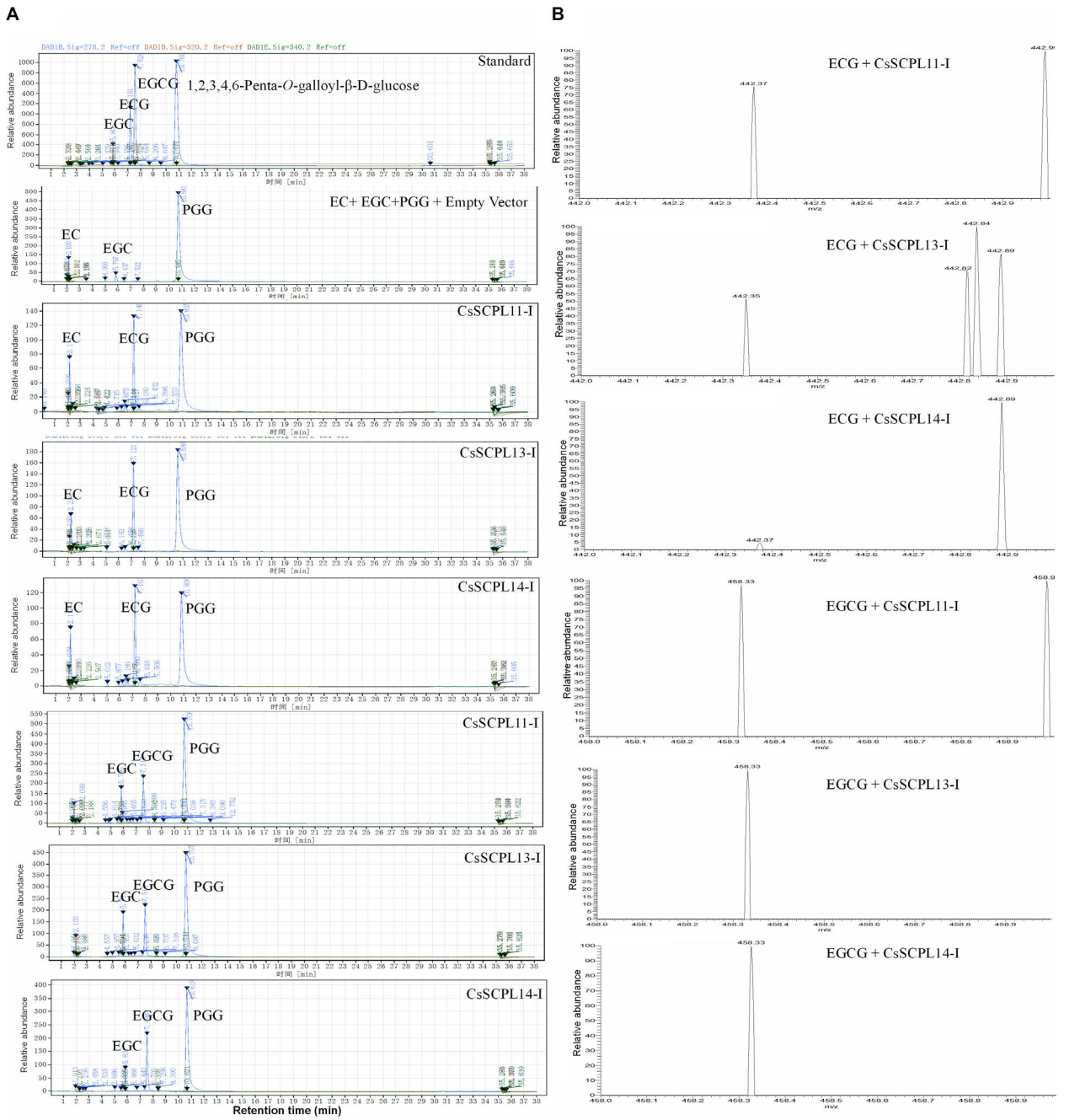

FIGURE 7 | Activity assay of recombinant CsSCPL11-IA, 13-IA, and 14-IA. **(A)** HPLC verification of activity assay with epicatechine (EC), epigallocatechine (EGC) and 1,2,3,4,6-Penta-O-galloyl-β-D-glucose (PGG). **(B)** LC-MS spectra of CsSCPL11-IA, L13-IA and 14-IA activity assay products. ECG, Epicatechin gallate; EGCG, Epigallocatechin gallate. Analytical data were obtained from atleast three biological replicates.

paralogous with K_s values close to that expected for this event (Guo et al., 2014).

Our selection pressure analyses using different models revealed that several codon sites in each branch of the SCPL phylogenetic tree were under positive selection. About 5% (26 out of 476 codons) were under positive or purifying selection in CsSCPL-I genes. Moreover, our analysis revealed an interesting evolutionary dynamic in the CsSCPL-I clade. Grounded on K_a/K_s values, many of these genes are under purifying selection, indicating strong selection pressures after the gene duplication event (**Supplementary Table S3**). This observation showed that the variation in expression patterns of duplicated genes may be evolving toward novel functions after duplication. It is believed that galloylated catechins contents in modern tea

TABLE 1 | Acyltransferase kinetics of the recombinant SCPL11-IA, 13-IA, and 14-IA with different substrates.

Substrate	Enzyme	K_M (μM)	Vmax (μmol $mg^{-1}min^{-1}$)	Kcat (s^{-1})	Kcat/Km (s^{-1} μM^{-1})
EC	*CsSCPL11 – IA*	85.39 ± 10.62	16.13 ± 2.26	268.82 ± 22.76	3.14 ± 0.64
	CsSCPL13 – IA	160.41 ± 21.2	37.04 ± 5.25	617.28 ± 62.48	3.85 ± 0.42
	CsSCPL14 – IA	64.28 ± 0.81	20.00 ± 4.65	333.33 ± 35.25	5.19 ± 0.48
EGC	*CsSCPL11 – IA*	40.67 ± 5.56	13.16 ± 1.64	219.30 ± 31.74	5.39 ± 0.81
	CsSCPL13 – IA	49.73 ± 8.36	20.41 ± 3.95	340.14 ± 45.43	6.84 ± 0.94
	CsSCPL14 – IA	103.38 ± 15.63	62.50 ± 5.62	1041.67 ± 146.8	10.08 ± 1.64
PGG	*CsSCPL11 – IA*	13.34 ± 1.78	4.05 ± 0.62	67.48 ± 7.34	5.06 ± 0.97
	CsSCPL13 – IA	10.43 ± 1.25	3.67 ± 0.52	61.05 ± 8.82	5.85 ± 0.84
	CsSCPL14 – IA	20.81 ± 3.51	10.64 ± 1.44	177.31 ± 22.76	8.52 ± 1.38

Analytical data were obtained from atleast three biological replicates and are expressed as means ± SD. EGC was used as an acceptor when kinetic parameters were measured for PGG. EC, epicatechin; EGC, epigallocatechin; PGG, 1,2,3,4,6-penta-O-galloyl-β-D-glucose.

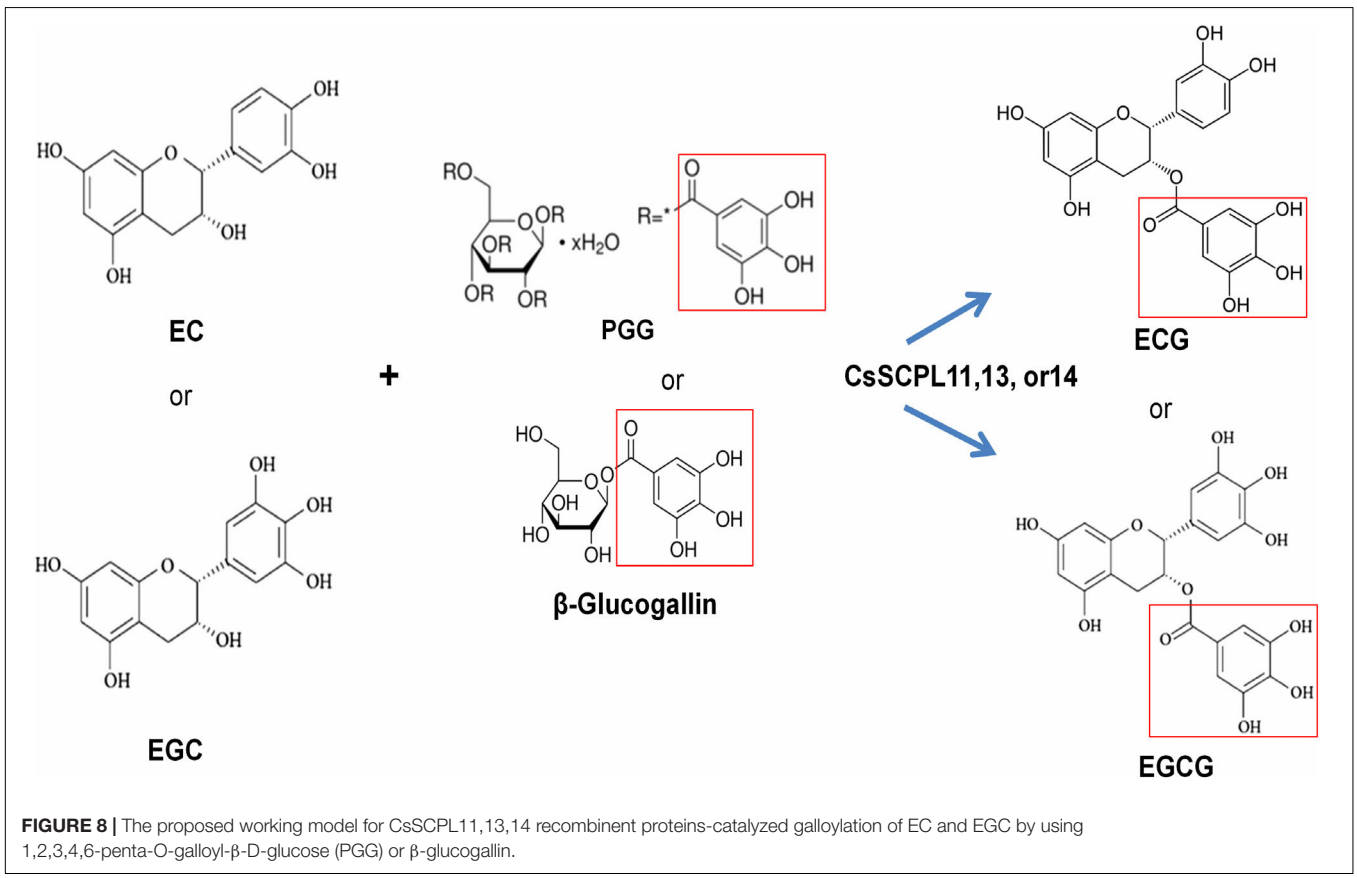

FIGURE 8 | The proposed working model for CsSCPL11,13,14 recombinent proteins-catalyzed galloylation of EC and EGC by using 1,2,3,4,6-penta-O-galloyl-β-D-glucose (PGG) or β-glucogallin.

plant cultivars are one of domestication traits (Wei et al., 2018). Consistent with high farmer selection pressures on higher catechins contents, primarily galloylated catechins, in modern tea plant cultivars during thousand years of cultivation history, domestications and evolution, SCPL-I, but not SCPL-II genes were also subject to positive selection presures during the evolution (**Supplementary Figure S1**).

In our phylogenetic analysis, the number of SCPL genes found in the genomes of different plant species varied in each clade. This finding suggests that the different species studied exhibit conserved evolution and that the gene family evolved in multiple directions in these lineages. The gene clustering events observed in two blocks of scaffold2990 suggest the CsSCPL-I clade evolved more rapidly compared to the other two branches. The exon/intron arrangement was similar within the same clade (Du et al., 2012) and the modal length of the first two exons is well conserved, as observed in Vitis and Arabidopsis (Matus et al., 2008). The phylogenetic distribution of genes in our analysis corroborated previous reports. However, a few members showed distinct exon/intron arrangements, indicating that these SCPL genes may play different functions. CsSCPL genes contain no more than 14 exons, which is comparable to the exon numbers in *O. sativa* and *A. thaliana* SCPL genes. Therefore, we conclude that the SCPL family members in *C. sinensis* maintained their

exon/intron structures over the course the evolution, unlike what happened in poplar (Zhu et al., 2018).

Interestingly, some *CsSCPL-I* paralogous pairs were dissimilar in their exon/intron structure, indicating divergence at their gene architecture. Moreover, the conserved motifs found in *CsSCPL* genes shared similar characteristics that indicate a close evolutionary relationship in the family, especially within the same class. The same motif structures found in paralogous pairs reveal that CsSCPL proteins are potentially functionally redundant, whereas, on the other hand, differences in motif number imply functional variation and divergence. Genome duplication events can potentially lead to the modification of gene properties and significantly increase functional variation (Liu et al., 2014; Ahmad et al., 2019). Two paralogous pairs showed K_a/K_s ratios significantly < 1, indicating that they are under strong purifying process and contained highly conserved amino acid residues whereas the three remaining paralogous pairs had K_a/K_s ratios significantly > 1, suggesting that tight evolutionary constraints are under effect to sustain their stability.

Identification and Characterization of Genes Putatively Involved in Catechins Galloylation

Serine carboxypeptidase-like proteins are involved in the biosynthesis of secondary metabolites conferring tolerance to biotic and abiotic stresses (Wilson et al., 2016). Our study clearly shows that most CsSCPL1As share high similarity in enzyme structure, gene expression pattern, regulatory mechanism, and likely enzymatic function. As a result of whole-genome and recent tandem gene duplication events, the expanded CsSCPL1A genes should play key roles in the formation of tea characteristic secondary metabolites, under both environmental and farmer selection pressures for more galloylated catechins contents in shoot tips. While it is difficult to distinguish clear or subtle differences between these CsSCPL genes, that still undergo convergent or divergent evolution under natural or artificial selection, we further combined gene expression and catechins metabolite correlation analyses to deciphor CsSCPL1As for putative functional differences. Both developmental and environmental factors drastically regulate catechins biosyntesis and accumulation, thus, gene expression patterns in tea plant tissues under normal or stressful environments were analyzed concerning catechins accumulation (**Figures 4**, **5**). These data betrayed their generally similar but individually differential roles and putatively diverse functions.

Our sequence profiling revealed conserved regulatory *cis*-acting elements present in CsSCPL-I gene promoters. Their modular compositions varied among the CsSCPL-I members, potentially coordinating responses to complex stimuli, such as light, hormones, and biotic and abiotic stresses. The analysis of transcriptional patterns can offer clues to explore gene functions. Expression of CsSCPL genes was assessed in different tissues and suggested functional differences that *CsSCPL* genes play *in planta*. In addition, about 61% (17/28) of the *CsSCPL* genes showed high expression in young leaves while 57% (16/28) was highly expressed in the apical bud, suggesting they play roles in these organs. Our qRT-PCR analysis of selected CsSCPL genes was consistent with the public RNA-Seq data analyzed. CsSCPL expression generally increased in response to Al^{3+} and MeJA exposures, except for *CsSCPL11-IA*, which expression was repressed in response to MeJA (**Figure 4**). CsSCPL transcriptional activity was induced in response to heat, but it decreased in response to cold, high salinity, and drought stresses (Chiu et al., 2016). Studies in rice SCPLs suggested that OsBISCPL1 plays a role in defense against multiple stresses (Liu et al., 2008).

It has been speculated that Cs*SCPL-I* genes are involved in the synthesis of galloylated catechins in young leaves and the apical bud for defense against insects or pathogens in the tea plant (Wei et al., 2018). Each member of the paralogous pairs *CsSCPL9-I/11-IA, CsSCPL15-I/10-IA*, and *CsSCPL19-IA/1-IA* shows distinct expression patterns and gene architectures, implying that they may be undergoing functional divergence. On the other hand, *CsSCPL27-IA/24-IA* and *CsSCPL14-IA/24-IA* showed, in addition to sequence similarity, very similar expression patterns and gene structures, strongly indicating functional redundancy. Cs*SCPL2-IA* and CsSCPL*11-IA* were highly induced under MeJA treatment. Therefore, these genes may be interesting candidates for playing roles in disease resistance mechanisms.

The biosynthesis of galloylated catechins is based on galloyl transacylation reactions at the 3-position in the C ring of non-galloylated catechins. The transacylation reactions are accomplished via activated donor molecules in plants. Coenzyme A thioesters in the BAHD family and 1-*O*-glucose esters in SCPLs serve as activated donors due to their high free energy of hydrolysis (Fraser et al., 2007). Among 1-*O*-glucose gallic esters, PGG or its hydrolyzing products, such as di-, tri-, tetra-, and mono-galloylglucoses, such as β-glucogallin, can act as galloyl donors used to modify EC or EGC to form ECG or EGCG, respectively. Tea plant leaves produce many polygalloylated glucose derivatives (Yang and Tomás-Barberán, 2018; Wei et al., 2019). β-Glucogallin is generated from UDP-glucose and gallic acid in tea plants by UDP-glucose: galloyl-1-O-β-D-glucosyltransferase CsUGT84A22 (Liu et al., 2012), whose homologs were also reported in several other plants (Ono et al., 2016; Tahara et al., 2018; Zhao et al., 2020). β-Glucogallin is further used as the galloyl donor in the sequential galloylation of 1,6-di-O-galloyl-β-D-glucose to generate 1,2,6- tri-, 1,2,3,6-tetra-, and 1,2,3,4,6-penta-O-galloyl-β-D-glucose (Niemetz and Gross, 2005). In the biosynthesis of gallotannins, 1,6-di-O-galloyl-glucose: 1,6-di-O-galloylglucose 2-O-galloyltransferase activity was also detected, suggesting that 1,6-di-O-galloyl-glucose can also act as an acyl donor (Niemetz and Gross, 2005).

On the other hand, hydrolyzable tannins, including PGG (pentagalloylglucose) and EGCG, can be readily hydrolyzed enzymatically to generate series of galloylated glucose derivatives, such as di-, tri-, tetra-, and mono-galloylglucoses, as well as gallic acid and catechins, respectively, by many organisms, such as plants and microbes (Jana et al., 2014; Dai et al., 2020; Zhao et al., 2020). Indeed, our assay showed that empty vector control displayed significant hydrolysis of PGG (**Figure 7**). That explains why EGCG in tea plant leaves is not stable, so did PGG, and perhaps, why PGG can be used as acyl donors. Tea plant

leaves accumulated high levels of di-, tri-, tetra-, and mono-galloylglucoses[12] (Yang and Tomás-Barberán, 2018; Wei et al., 2019).

Microorganisms, such as bacteria and phytopathogens, have particularly active enzyme tannases or hydrolyzable tannin hydrolases, to overcome the plant chemical defense and utilize energy provided by plants (Jana et al., 2014). Therefore, the partially purified enzyme extracts in this study might contain some bacterial tannase-activity proteins, could trigger the hydrolysis of PGG into various galloyl glucoses that may act as proper acyl donors for the catechins galloylations, which could also occur in tea plants. Nevertheless, our work is the first study to display galloylated catechin synthesis through an enzymatic assay catalyzed by CsSCPLs from in tea plants. Previously, two reaction steps of an enzyme from plant extracts was carried out, involving UGGT and another involving ECGT in *C. sinensis* and used β-glucogallin as a donor molecule for the transacylation reactions (Liu et al., 2012). But the gene coding the enzyme is not identified yet.

CsSCPL enzymes are hydrophobic glycoproteins, and more than one serine residue, rather than metal ions, is involved in their active sites. Only a few studies e.g., (Ikegami et al., 2007; Terrier et al., 2009, in persimmon and grapevine, respectively) have investigated at the genetic level SCPL acyltransferases, which are most likely involved in the biosynthesis of galloylated catechins. However, their enzymatic functions have not been verified, and the correlations with galloylation of catechins were only hypothetical. Five candidate genes clustered into the CsSCPL-I clade capable of catalyzing the formation of glucose esters were selected from the *C. sinensis* genome. Their functional analyses were assessed through enzyme activity assays. Among them, three genes were successfully expressed in *E. coli* and showed higher binding affinity to EGC than EC. The discrepancies between the catalytic properties observed in our study and that reported by Liu et al. (2012) may be due to the substrate used for the reactions or the enzyme sources since that study used plant extracts whereas we used purified recombinant enzymes.

Three SCPL-IA Enzymes Are Involved in the Biosynthesis of Galloylated Catechins

The availability of the tea genome sequence enabled the identification of the IA subclade of the SCPL acyltransferase gene family, which is likely involved in synthesizing major portions of monomeric galloylated catechins in the leaves (Wei et al., 2018). Unlike the insoluble polymerized proanthocyanidins (PAs, or condensed tannins) that are primarily present in the majority of other plant groups, tea leaves mainly accumulate soluble monomeric catechins (e.g., C, GC, EC, EGC) and their galloylation derivatives (CG, ECG, GCG, EGCG) instead (Wei et al., 2015). These molecules account for up to 75% of the total catechins in leaves and have a major impact on tea quality. However, the genetic basis for the biosynthesis of galloylated catechins is not fully understood, which currently one of the fundamental biological questions in tea biology (Zhao et al., 2020). Biochemical studies demonstrated that the biosynthesis of flavan-3-ols via 1-*O*-glucose ester-dependent reactions are catalyzed by galloyl-1-*O*-β-D-glucosyltransferase (UGGT) and that epicatechin:1-*O*-galloyl-β-D-glucose *O*-galloyltransferase (ECGT) (Liu et al., 2012). Similar reactions have been found in other plant species, such as grapevine and persimmon (Ikegami et al., 2007; Terrier et al., 2009; Wilson et al., 2016). So far, among at least 22 CsSCPL-I genes redundantly present in the tea genome, which have been regarded as the most likely candidates responsible for the galloylation of catechins, no molecular evidence of any gene involved in galloylated catechins biosyntesis in tea plants has been published so far (Wei et al., 2018). Our comprehensive analyses imply that these CsSCPL-I may have overlapping but differential functions. Due to the presence of diverse and complex hydrolyzable tannins, including various *O*-glucose gallic esters and galloylated catechins, CsSCPL-I acyltransferases may have similar but differential enzymatic functions.

We identified and characterized three SCPL-IA genes which expression levels were highly correlated with the accumulation of EGCG and ECG, given their high transcriptional activities in apical buds and young leaves, where most galloylated flavan-3-ols accumulate (**Figure 5B**). We further confirmed enzymatically that the recombinant enzymes were able to catalyze the production of galloylated catechins *in vitro* (**Table 1** and **Figure 8**). A recent study showed that the galloylated catechins, such as EGCGs, are primarily localized to the central vacuole for storage (Xu et al., 2016). Consistently, our analysis showed that CsSCPL11 and CsSCPL13 are predicted to be localized to the lososomes, including vacuoles (**Supplementary Dataset S1**). The measured enzymatic activities of these SCPLs were relatively low *in vitro*, considering the high levels of galloylated catechins that accumulate as major forms in tea leaves. Several reasons could explain this result. Firstly, PGG, rather than β-glucogallin, was used as an acyl donor (**Figure 8**). Although it was shown that PGG could be efficiently hydrolyzed into different galloylglucoses, including β-glucogallin, the enzyme efficiency was low. Secondly, protein modifications (e.g., glycosylation, phosphorylation), or molecular interactions (e.g., homo- or heteromerization, allosteric control) may be required *in celula* to enhance enzyme activity. This is the first report on CsSCPL-I genes responsible for galloylated catechins biosyntesis. Notwithstanding, our results prove the hypothesis that SCPLs are involved in the galloylation of EC or EGC, given that all the three enzymes studied preferred to modify EC or EGC into their galloylated forms. Further studies on these aspects are essential to further our understanding of the roles SCPL enzymes play in plant physiology.

CONCLUSION

We demonstrated that both convergent and divergent evolution of CsSCPL1A genes in tea plant genome and their generally similar but differential gene expression patterns in various tea

[12]http://pcsb.ahau.edu.cn:8080/TCDB/f

plant tissues as effects of developmental and environmental factors, and that CsSCPL11-IA, 13-IA and 14-IA are SCPL acyltransferases that share similar enzymatic kinetics in the galloylation of EC or EGC. Our genome-wide analysis of the SCPL gene family in tea, and biochemical characterization of three recombinant CsSCPLs, such as substrate specificity and enzymatic kinetic parameters, revealed important results to understand the physiological roles these compounds play in tea. The three CsSCPL genes functionally characterized in the study display distinct expression patterns in different tissues of the plant and response to abiotic stress and hormones. The insights provided by this study will not only help to understand the biosynthesis timing and location of these galloylated metabolites but also which physiological roles that ECG or EGCG play in tea plants.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The Tea Plant Information Archive – TPIA public database: http://teaplant.org/; http://tpia.teaplant.org/index.html. CsSCPL2-I (TEA009664, NCBI Genbank accession MK843824), CsSCPL5-I (TEA034028; MK843825), CsSCPL11-I (TEA023451; MK843826), CsSCPL13-I (TEA034055; MK843827), and CsSCPL14-I (TEA027270; MK843828)].

AUTHOR CONTRIBUTIONS

JZ planned and designed the research. MA, PL, GS, and EX performed experiments and analyzed data. MA, JZ, VB, and XW wrote and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2018YFD1000601), the Key Research and Development (R&D) Program of Anhui Province (18030701155), Anhui Agricultural University, and the State Key Laboratory of Tea Plant Biology and Utilization.

ACKNOWLEDGMENTS

The authors thank Prof. Zhao's lab members for all assistance with experiments, discussion, and data analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00848/full#supplementary-material

FIGURE S1 | Selection pressures among CsSCPLI and II gene sequences using mechanistic empirical combination (MEC) model.

FIGURE S2 | Topological analysis of CsSCPLs and designed expression of truncated proteins.

FIGURE S3 | HPLC analyses of CsSCPLs-enzymatic reactions.

FIGURE S4 | LC-MS/MS analysis of CsSCPLs-enzymatic products in negative mode.

TABLE S1 | Composition of tea plant standard culture solution of Shigeki Konishi.

TABLE S2 | Conserved domain analysis of CsSCPL proteins.

TABLE S3 | Synteny analysis of CsSCPL gene paralogues.

TABLE S4 | Log-likelihood values and statistics for PAML site models of positively selection.

TABLE S5 | Positively selected sites under different PAML site models using Bayes empirical analysis.

DATASET S1 | Physical parameters of tea proteins coded by CsSCPL genes.

DATASET S2 | *Cis*-acting analysis of CsSCPL gene promoter regions.

DATASET S3 | Expression level of CsSCPLs in different Tea tissues, response to cold, NaCl, $Al3^{+}/H^{+}$ stresses, Shedding, MejA and PEG treatment.

DATASET S4 | List of primers of CsSCPL genes used for qRT-PCR validation.

DATASET S5 | Identity and similarities at amino-acid level among tea and different characterized SCPLs.

REFERENCES

Ahmad, M. Z., Li, P., Wang, J., Rehman, N. U., and Zhao, J. (2017). Isoflavone malonyltransferases *GmIMaT1* and *GmIMaT3* differently modify isoflavone glucosides in soybean (*Glycine max*) under various stresses. *Front. Plant Sci.* 8:735. doi: 10.3389/fpls.2017.00735

Ahmad, M. Z., Rehman, N. U., Yu, S., Zhou, Y., Haq, B. U., Wang, J., et al. (2020). GmMAX2–D14 and–KAI interaction-mediated SL and KAR signaling play essential roles in soybean root nodulation. *Plant J.* 101, 334–351. doi: 10.1111/tpj.14545

Ahmad, M. Z., Sana, A., Jamil, A., Nasir, J. A., Ahmed, S., Hameed, M. U., et al. (2019). A genome-wide approach to the comprehensive analysis of GASA gene family in *Glycine max*. *Plant Mol. Biol.* 100, 607–620. doi: 10.1007/s11103-019-00883-1

Akagi, T., Ikegami, A., Suzuki, Y., Yoshida, J., Yamada, M., Sato, A., et al. (2009). Expression balances of structural genes in shikimate and flavonoid biosynthesis cause a difference in proanthocyanidin accumulation in persimmon (*Diospyros kaki* Thunb.) fruit. *Planta* 230, 899–915. doi: 10.1007/s00425-009-0991-6

Baulcombe, D. C., Barker, R., and Jarvis, M. (1987). A gibberellin responsive wheat gene has homology to yeast carboxypeptidase Y. *J. Biol. Chem.* 262, 13726–13735.

Bontpart, T., Cheynier, V., Ageorges, A., and Terrier, N. (2015). BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds. *New Phytol.* 208, 695–707. doi: 10.1111/nph.13498

Bowles, D., Isayenkova, J., Lim, E.-K., and Poppenberger, B. (2005). Glycosyltransferases: managers of small molecules. *Curr. Opin. Plant Biol.* 8, 254–263. doi: 10.1016/j.pbi.2005.03.007

Carrier, G., Huang, Y.-F., Le Cunff, L., Fournier-Level, A., Vialet, S., Souquet, J.-M., et al. (2013). Selection of candidate genes for grape proanthocyanidin pathway by an integrative approach. *Plant Physiol. Biochem.* 72, 87–95. doi: 10.1016/j.plaphy.2013.04.014

Chiu, C.-H., Chen, G.-H., Tzen, J. T., and Yang, C.-Y. (2016). Molecular identification and characterization of a serine carboxypeptidase-like gene

associated with abiotic stress in tea plant, *Camellia sinensis* (L.). *Plant Growth Regul.* 79, 345–353. doi: 10.1007/s10725-015-0138-7

Ciarkowska, A., Ostrowski, M., and Jakubowska, A. (2018). A serine carboxypeptidase-like acyltransferase catalyzes synthesis of indole-3-acetic (IAA) ester conjugate in rice (*Oryza sativa*). *Plant Physiol. Biochem.* 125, 126–135. doi: 10.1016/j.plaphy.2018.02.007

Cui, L., Yao, S., Dai, X., Yin, Q., Liu, Y., Jiang, X., et al. (2016). Identification of UDP-glycosyltransferases involved in the biosynthesis of astringent taste compounds in tea (*Camellia sinensis*). *J. Exper. Bot.* 67, 2285–2297. doi: 10.1093/jxb/erw053

Dai, X., Liu, Y., Zhuang, J., Yao, S., Liu, L., Jiang, X., et al. (2020). Discovery and characterization of tannase genes in plants: roles in hydrolysis of tannins. *New Phytol.* 226, 1104–1116. doi: 10.1111/nph.16425

de las Rivas, B., Rodríguez, H., Anguita, J., and Muñoz, R. (2019). Bacterial tannases: classification and biochemical properties. *Appl. Microbiol. Biotechnol.* 103, 603–623. doi: 10.1007/s00253-018-9519-y

Du, H., Yang, S.-S., Liang, Z., Feng, B.-R., Liu, L., Huang, Y.-B., et al. (2012). Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biol.* 12:106. doi: 10.1186/1471-2229-12-106

Fraser, C. M., Rider, L. W., and Chapple, C. (2005). An expression and bioinformatics analysis of the *Arabidopsis serine* carboxypeptidase-like gene family. *Plant Physiol.* 138, 1136–1148. doi: 10.1104/pp.104.057950

Fraser, C. M., Thompson, M. G., Shirley, A. M., Ralph, J., Schoenherr, J. A., Sinlapadech, T., et al. (2007). Related *Arabidopsis* serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. *Plant Physiol.* 144, 1986–1999. doi: 10.1104/pp.107.098970

Fujiwara, H., Tanaka, Y., Yonekura-Sakakibara, K., Fukuchi-Mizutani, M., Nakao, M., Fukui, Y., et al. (1998). cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. *Plant J.* 16, 421–431. doi: 10.1046/j.1365-313x.1998.00312.x

Grundhöfer, P., and Gross, G. G. (2001). Immunocytochemical studies on the origin and deposition sites of hydrolyzable tannins. *Plant Sci.* 160, 987–995. doi: 10.1016/s0168-9452(01)00341-7

Guo, L., Chen, Y., Ye, N., Dai, X., Yang, W., and Yin, T. (2014). Differential retention and expansion of the ancestral genes associated with the paleopolyploidies in modern rosid plants, as revealed by analysis of the extensins super-gene family. *BMC Genom.* 15:612. doi: 10.1186/1471-2229-12-612

Hayashi, N., Chen, R., Hiraoka, M., Ujihara, T., and Ikezaki, H. (2010). β-Cyclodextrin/surface plasmon resonance detection system for sensing bitter-astringent taste intensity of green tea catechins. *J. Agric. Food Chem.* 58, 8351–8356. doi: 10.1021/jf1012693

Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu817

Ikegami, A., Eguchi, S., Kitajima, A., Inoue, K., and Yonemori, K. (2007). Identification of genes involved in proanthocyanidin biosynthesis of persimmon (*Diospyros kaki*) fruit. *Plant Sci.* 172, 1037–1047. doi: 10.1016/j.plantsci.2007.02.010

Jana, A., Halder, S. K., Banerjee, A., Paul, T., Pati, B. R., Mondal, K. C., et al. (2014). Biosynthesis, structural architecture and biotechnological potential of bacterial tannase: a molecular advancement. *Bioresour. Technol.* 157, 327–340. doi: 10.1016/j.biortech.2014.02.017

Jones, C. G., Lycett, G. W., and Tucker, G. A. (1996). Protease inhibitor studies and cloning of a serine carboxypeptidase cDNA from germinating seeds of pea (*Pisum sativum* L.). *Eur. J. Biochem.* 235, 574–578. doi: 10.1111/j.1432-1033.1996.00574.x

Juretic, N., Hoen, D. R., Huynh, M. L., Harrison, P. M., and Bureau, T. E. (2005). The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res.* 15, 1292–1297. doi: 10.1101/gr.4064205

Karas, D., Ulrichová, J., and Valentová, K. (2017). Galloylation of polyphenols alters their biological activity. *Food Chem. Toxicol.* 105, 223–240. doi: 10.1016/j.fct.2017.04.021

Khater, F., Fournand, D., Vialet, S., Meudec, E., Cheynier, V., and Terrier, N. (2012). Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis. *J. Exper. Bot.* 63, 1201–1214. doi: 10.1093/jxb/err340

Kong, H., Landherr, L. L., Frohlich, M. W., Leebens-Mack, J., Ma, H., and DePamphilis, C. W. (2007). Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth. *Plant J.* 50, 873–885. doi: 10.1111/j.1365-313x.2007.03097.x

Kosma, D. K., Molina, I., Ohlrogge, J. B., and Pollard, M. (2012). Identification of an *Arabidopsis* fatty alcohol: caffeoyl-coenzyme A acyltransferase required for the synthesis of alkyl hydroxycinnamates in root waxes. *Plant Physiol.* 160, 237–248. doi: 10.1104/pp.112.201822

Koyama, K., Numata, M., Nakajima, I., Goto-Yamamoto, N., Matsumura, H., and Tanaka, N. (2014). Functional characterization of a new grapevine MYB transcription factor and regulation of proanthocyanidin biosynthesis in grapes. *J. Exper. Bot.* 65, 4433–4449. doi: 10.1093/jxb/eru213

Lehfeldt, C., Shirley, A. M., Meyer, K., Ruegger, M. O., Cusumano, J. C., Viitanen, P. V., et al. (2000). Cloning of the *SNG1* gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *Plant Cell.* 12, 1295–1306. doi: 10.1105/tpc.12.8.1295

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

Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40, D302–D305.

Li, A. X., and Steffens, J. C. (2000). An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6902–6907. doi: 10.1073/pnas.110154197

Li, J., Lease, K. A., Tax, F. E., and Walker, J. C. (2001). BRS1, a serine carboxypeptidase, regulates BRI1 signaling in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 5916–5921. doi: 10.1073/pnas.091065998

Liu, H., Wang, X., Zhang, H., Yang, Y., Ge, X., and Song, F. (2008). A rice serine carboxypeptidase-like gene *OsBISCPL1* is involved in regulation of defense responses against biotic and oxidative stress. *Gene* 420, 57–65. doi: 10.1016/j.gene.2008.05.006

Liu, L., Li, Y., She, G., Zhang, X., Jordan, B., Chen, Q., et al. (2018). Metabolite profiling and transcriptomic analyses reveal an essential role of UVR8-mediated signal transduction pathway in regulating flavonoid biosynthesis in tea plants (*Camellia sinensis*) in response to shading. *BMC Plant Biol.* 18:233. doi: 10.1186/1471-2229-12-233

Liu, Y., Gao, L., Liu, L., Yang, Q., Lu, Z., Nie, Z., et al. (2012). Purification and characterization of a novel galloyltransferase involved in catechin galloylation in the tea plant (*Camellia sinensis*). *J. Biol. Chem.* 287, 44406–44417. doi: 10.1074/jbc.m112.403071

Liu, Z., Zhang, M., Kong, L., Lv, Y., Zou, M., Lu, G., et al. (2014). Genome-wide identification, phylogeny, duplication, and expression analyses of two-component system genes in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *DNA Res.* 21, 379–396. doi: 10.1093/dnares/dsu004

Matus, J. T., Aquea, F., and Arce-Johnson, P. (2008). Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across vitis and *Arabidopsis* genomes. *BMC Plant Biol.* 8:83. doi: 10.1186/1471-2229-12-83

Milkowski, C., Baumert, A., Schmidt, D., Nehlin, L., and Strack, D. (2004). Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. *Plant J.* 38, 80–92. doi: 10.1111/j.1365-313x.2004.02036.x

Milkowski, C., and Strack, D. (2004). Serine carboxypeptidase-like acyltransferases. *Phytochemistry* 65, 517–524. doi: 10.1016/j.phytochem.2003.12.018

Moore, J. P., Westall, K. L., Ravenscroft, N., Farrant, J. M., Lindsey, G. G., and Brandt, W. F. (2005). The predominant polyphenol in the leaves of the resurrection plant *Myrothamnus flabellifolius*, 3, 4, 5 tri-*O*-galloylquinic acid, protects membranes against desiccation and free radical-induced oxidation. *Biochem. J.* 385, 301–308. doi: 10.1042/bj20040499

Moura, D. S., Bergey, D. R., and Ryan, C. A. (2001). Characterization and localization of a wound-inducible type I serine-carboxypeptidase from leaves of tomato plants (*Lycopersicon esculentum* Mill.). *Planta* 212, 222–230. doi: 10.1007/s004250000380

Mugford, S. T., and Milkowski, C. (2012). Serine carboxypeptidase-like acyltransferases from plants. *Methods Enzymol.* 516, 279–297. doi: 10.1016/b978-0-12-394291-3.00006-x

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., and Pond, S. L. K. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* 8:e1002764. doi: 10.1371/journal.pgen.1002764

Niemetz, R., and Gross, G. G. (2005). Enzymology of gallotannin and ellagitannin biosynthesis. *Phytochemistry* 66, 2001–2011. doi: 10.1016/j.phytochem.2005.01.009

Ono, N. N., Qin, X., Wilson, A. E., Li, G., and Tian, L. (206). Two UGT84 family glycosyltransferases catalyze a critical reaction of hydrolyzabl1e tannin biosynthesis in pomegranate (*Punica granatum*). *PLoS One* 11:e0156319. doi: 10.1371/journal.pgen.100156319

Saffari, Y., and Sadrzadeh, S. H. (2004). Green tea metabolite EGCG protects membranes against oxidative damage in vitro. *Life Sci.* 74, 1513–1518. Saffari and Sadrzadeh

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

Scott, M. S., Calafell, S. J., Thomas, D. Y., and Hallett, M. T. (2005). Refining protein subcellular localization. *PLoS Comput. Biol.* 1:e66. doi: 10.1371/journal.pgen.100e66

Shi, J., Ma, C., Qi, D., Lv, H., Yang, T., Peng, Q., et al. (2015). Transcriptional responses and flavor volatiles biosynthesis in methyl jasmonate-treated tea leaves. *BMC Plant Biol.* 15:233. doi: 10.1186/s12870-015-0609-z

Shirley, A. M., and Chapple, C. (2003). Biochemical characterization of sinapoylglucose: choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. *J. Biol. Chem.* 278, 19870–19877. doi: 10.1074/jbc.m302362200

Shirley, A. M., McMichael, C. M., and Chapple, C. (2001). The *sng2* mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose: choline sinapoyltransferase. *Plant J.* 28, 83–94. doi: 10.1046/j.1365-313x.2001.01123.x

Steffens, J. C. (2000). Acyltransferases in protease's clothing. *Plant Cell* 12, 1253–1256.

Tahara, K., Nishiguchi, M., Frolov, A., Mittasch, J., and Milkowski, C. (2018). Identification of UDP glucosyltransferases from the aluminum-resistant tree Eucalyptus camaldulensis forming β-glucogallin, the precursor of hydrolyzable tannins. *Phytochemistry* 152, 154–161. doi: 10.1016/j.phytochem.2018.05.005

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tang, H., Bowers, J. E., Wang, X., Ming, R., Alam, M., and Paterson, A. H. (2008). Synteny and collinearity in plant genomes. *Science* 320, 486–488. doi: 10.1126/science.1153917

Terrier, N., Torregrosa, L., Ageorges, A., Vialet, S., Verries, C., Cheynier, V., et al. (2009). Ectopic expression of *VvMybPA2* promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. *Plant Physiol.* 149, 1028–1041. doi: 10.1104/pp.108.131862

Wang, J., Hou, Q., Li, P., Yang, L., Sun, X., Benedito, V. A., et al. (2017). Diverse functions of multidrug and toxin extrusion (MATE) transporters in citric acid efflux and metal homeostasis in *Medicago truncatula*. *Plant J.* 90, 79–95. doi: 10.1111/tpj.13471

Wang, S., Alseekh, S., Fernie, A. R., and Luo, J. (2019). The structure and function of major plant metabolite modifications. *Mol. Plant* 12, 899–919. doi: 10.1016/j.molp.2019.06.001

Washio, K., and Ishikawa, K. (1994). Cloning and sequencing of the gene for type I carboxypeptidase in rice. *Biochim. Biophys. Acta Gen. Subj.* 1199, 311–314. doi: 10.1016/0304-4165(94)90012-4

Wei, C., Yang, H., Wang, S., Zhao, J., Liu, C., Gao, L., et al. (2018). Draft genome sequence of *Camellia sinensis* var. sinensis provides insights into the evolution of the tea genome and tea quality. *Proc. Natl. Acad. Sci. U.S.A.* 115, E4151–E4158.

Wei, K., He, H., Li, H., Wang, L., Ruan, L., Pang, D., et al. (2019). Gallotannin 1, 2, 6-tri-*O*-galloyl-β-D-glucopyranose: Its availability and changing patterns in tea (*Camellia sinensis*). *Food Chem.* 296, 40–46. doi: 10.1016/j.foodchem.2019.05.144

Wei, K., Wang, L., Zhang, C., Wu, L., Li, H., Zhang, F., et al. (2015). Transcriptome analysis reveals key flavonoid 3′-hydroxylase and flavonoid 3′, 5′-hydroxylase genes in affecting the ratio of dihydroxylated to trihydroxylated catechins in *Camellia sinensis*. *PLoS One* 10:e0137925. doi: 10.1371/journal.pgen.00137925

Weier, D., Mittasch, J., Strack, D., and Milkowski, C. (2008). The genes *BnSCT1* and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression. *Planta* 227, 375–385. doi: 10.1007/s00425-007-0624-x

Wilson, A. E., Matel, H. D., and Tian, L. (2016). Glucose ester enabled acylation in plant specialized metabolism. *Phytochem. Rev.* 15, 1057–1074. doi: 10.1007/s11101-016-9467-z

Xu, H., Wang, Y., Chen, Y., Zhang, P., Zhao, Y., Huang, Y., et al. (2016). Subcellular localization of galloylated catechins in tea plants [*Camellia sinensis* (L.) O. Kuntze] assessed via immunohistochemistry. *Front. Plant Sci.* 7:728. doi: 10.3389/fpls.2017.00728

Yang, X., and Tomás-Barberán, F. A. (2018). Tea is a significant dietary source of ellagitannins and ellagic acid. *J. Agric. Food Chem.* 67, 5394–5404. doi: 10.1021/acs.jafc.8b05010

Zhang, J., and He, X. (2005). Significant impact of protein dispensability on the instantaneous rate of protein evolution. *Mol. Biol. Evol.* 22, 1147–1155. doi: 10.1093/molbev/msi101

Zhang, Q., Cai, M., Yu, X., Wang, L., Guo, C., Ming, R., et al. (2017). Transcriptome dynamics of *Camellia sinensis* in response to continuous salinity and drought stress. *Tree Genet. Genom.* 13:78.

Zhang, Y., Wei, K., Li, H., Wang, L., Ruan, L., Pang, D., and Cheng, H. (2018). Identification of key genes involved in catechin metabolism in tea seedlings based on transcriptomic and HPLC analysis. *Plant Physiol. Biochem.* 133, 107–115.

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

Zhao, J., Li, P., Xia, T., and Wan, X. (2020). Exploring plant metabolic genomics: chemical diversity, metabolic complexity in the biosynthesis and transport of specialized metabolites with the tea plant as a model. *Crit. Rev. Biotechnol.* 22, 1–22. doi: 10.1080/07388551.2020.1752617

Zhu, D., Chu, W., Wang, Y., Yan, H., Chen, Z., and Xiang, Y. (2018). Genome-wide identification, classification and expression analysis of the serine carboxypeptidase-like protein family in poplar. *Physiol. Plant.* 162:33.

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.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 11 August 2020
doi: 10.3389/fphar.2020.01181

Vitexin Possesses Anticonvulsant and Anxiolytic-Like Effects in Murine Animal Models

Denise Dias de Oliveira[1], *Cassio Prinholato da Silva*[1], *Bruno Benincasa Iglesias*[1] *and Renê O. Beleboni*[1,2*]

[1] Department of Biotechnology, University of Ribeirão Preto, Ribeirão Preto, Brazil, [2] School of Medicine, University of Ribeirão Preto, Ribeirão Preto, Brazil

OPEN ACCESS

Edited by:
María Pilar López-Gresa, Universitat Politècnica de València, Spain

Reviewed by:
Alexandra Olimpio Siqueira Cunha, University of São Paulo, Brazil
Donatus Wewura Adongo, University of Health and Allied Sciences, Ghana

***Correspondence:**
Renê O. Beleboni
rbeleboni@unaerp.br;
reneusp@yahoo.com

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: 04 February 2020
Accepted: 20 July 2020
Published: 11 August 2020

Citation:
de Oliveira DD, da Silva CP, Iglesias BB and Beleboni RO (2020) Vitexin Possesses Anticonvulsant and Anxiolytic-Like Effects in Murine Animal Models. Front. Pharmacol. 11:1181. doi: 10.3389/fphar.2020.01181

Different types of epilepsy and forms of pathological anxiety have been described as significant neurological disorders that may exist as comorbidities. Some of those disorders share the association of affected limbic areas/neuropathological triggers as well as the use of drugs for their clinical management. The aim of this work was to investigate the anticonvulsant and anxiolytic properties of the vitexin (apigenin-8-C-glucoside), since this compound is a flavonoid usually found as one of the major constituents in several medicinal plants claimed as anxiolytics and/or anticonvulsants. This investigation was performed by the use of a series of classical murine animal models of chemically induced-seizures and of anxiety-related tests (open-field, elevated plus-maze, and light-dark box tests). Here, we show that the systemic administration of vitexin (1.25; 2.5 and 5 mg/kg; i.p.) exhibited selective protection against chemically-induced seizures. Vitexin did not block seizures evoked by glutamate receptors agonists (NMDA and kainic acid), and it did not interfere with the latencies for these seizures. Conversely, the same treatments protected the animals in a dose-dependent manner against the seizures evoked by the Gabaergic antagonists picrotoxin and PTZ and rise the latency time for the first seizure on non-protected animals. The higher dose of vitexin protected 100% of animals against the tonic-clonic seizures triggered by GABA antagonists. The results from open-field, elevated plus-maze, and light-dark box tests indicated the anxiolytic properties of vitexin at similar range of doses described for the anticonvulsant action screening. Furthermore, these results pointed that vitexin did not cause sedation or locomotor impairment on animals. The selective action of vitexin against picrotoxin and PTZ may reinforce the hypothesis by which this compound acts mainly by the modulation of GABAergic neurotransmission and/or related pathways. This could be useful to explain the dual activity of vitexin as anticonvulsant and anxiolytic, and highlight the pharmacological interest on this promising flavonoid.

Keywords: epilepsy, anxiety, etnopharmacology, flavonoid, γ - aminobutyric acid (GABA), glutamate

INTRODUCTION

Epilepsy is a progressive chronic disorder characterized by an abnormal and synchronous neuronal firing associated with recurrent and unpredictable seizures (England et al., 2012). It affects more than 50 million people worldwide. The current antiepileptic drug arsenal fails to evoke a positive response in about 30% of diagnosed cases of epilepsy (World Health Organization, 2019a). Patients suffering from different types of epilepsy are often affected by psychiatric and behavioral comorbidities such as mood disorders, anxiety, psychoses, motor disorders, cognitive deficits, and social dysfunctions (Salpekar and Mula, 2018). Besides the broad epidemiological relevance of anxiety disorders at different segments of our modern society, anxiety disorders are usually observed in a large number of epileptic patients, which represents additional negative impacts in their already compromised quality of lives (Gandy et al., 2015). It is well known that the alleviation of anxiety symptoms may have a positive effect in the progress of treatment of epileptic patients contributing to ameliorate their general health condition (Fisher and Noble, 2017).

The pathophysiological mechanism related to the onset and maintenance of different types of epilepsy/seizures, includes the imbalance between excitatory (glutamate) and inhibitory (γ-aminobutyric acid (GABA) synaptic activities that are located at different brain areas (Bialer and White, 2010). Some brain networks are commonly involved at prevention or appearance of seizures as well as in the regulation of behavior and mood (Kwon and Park, 2014). Although the huge differences in the molecular and cellular bases involved in onset and development of epilepsy and anxiety, the same imbalance between GABA and Glutamate is also related to anxiety (Masneuf et al., 2014; Gauthier and Nuss, 2015). Of course, several intervenient factors can differentially contribute to the neuropathology underlying both disorders, particularly in terms of magnitude, recruited structures, and brain areas (Martin et al., 2009; Wu et al., 2018). In any case, the screening for new antiepileptic/anxiolytic drugs that are able to modulate the inhibitory and/or excitatory neurotransmission pathways may be worth for the treatment of epilepsy and/or anxiety (Moto et al., 2018).

Vitexin (apigenin-8-C-glucoside) has received great attention by presenting a wide range of pharmacological effects. Vitexin has been found, in some cases as the major constituent, at different medicinal plants potentially useful to treat anxiety and/or epilepsy (Nassiri-Asl et al., 2007; He et al., 2016). More specifically, recent studies have shown the protective effects of vitexin against some neurological and psychiatric diseases (Abbasi et al., 2013; Can et al., 2013; Guimarães et al., 2015). In spite of the interesting pharmacological properties of vitexin; only a few and preliminary efforts have been conducted to prove its potential anxiolytic and anticonvulsant activities. Therefore, the main aim of this work is to investigate these potential effects using a complementary set of experiments at a well-established animal model arrangement.

MATERIALS AND METHODS

Vitexin and Other Drugs

Vitexin (apigenin-8-C-glucoside) was purchased from Cayman Chemical Company, USA (CAS Number 3681-93-4, purity degree ≥ 98%). The proconvulsant drugs N-Methyl-D-aspartate (NMDA) (CAS Number: 6384-92-5; ≥ 98%); kainic acid (CAS Number: 58002-62-3; ≥ 98%); Pentylenetetrazol (PTZ) (CAS Number: 54-95-5) and picrotoxin (CAS Number 204-716-6) were purchased from Sigma Aldrich, St. Louis, MO. Diazepam (DZP) (injectable solution in 0.9% saline) and dimethyl sulfoxyde (DMSO) were acquired from União Química (Brazil) and Synth (Brazil), respectively.

Animals and Experimental Conditions

All experiments were carried out using male adult Wistar rats within 4 to 5 weeks of age weighing about 150 to 180 g. Rats were housed five per cage in a controlled condition of humidity and temperature under 12 h light/dark schedule at 6 am/6 pm. They were allowed to acclimatize to our host facilities for at least 3 days prior to any experimental manipulation. Chow and water were provided *ad libitum*. The procedures were in accordance with the University of Ribeirão Preto Ethic Committee (approval number 11/2015) and with the Guide for the Care and Use of Laboratory Animals (National Research Council, US).

The animals were divided in control and experimental groups (n=06) in all sets of experiments. All the experiments were carried out in an acoustic isolated room between 1:00 to 5:00 p.m. All apparatus or arenas used in experiments were cleaned with 70% ethanol after observation of each animal in all of the procedures. Rats were habituated to the testing room for 30-min before experimental procedures. All experiments were recorded by a digital video camera and copies of files kept in our laboratory as official documents for public consultation.

ASSESSMENT OF ANXIOLYTIC-LIKE EFFECTS OF VITEXIN

The anxiolytic activity of vitexin was evaluated by elevated plus-maze, light-dark box, and open-field tests. The behavior of animals on each experimental apparatus for anxiolytic effects was observed 30 min later after control or experimental treatments. For all experiments, Diazepam 2 mg/kg (dissolved in 0.9% saline) (i.p.) served as positive control while 1% DMSO in 0.9% saline (i.p.) and 0.9% saline solution (i.p.) served as controls. Experimental groups were composed of animals treated with different i.p. doses of vitexin (0.75, 1.25 and 2.5 mg/kg). Doses of vitexin were selected based on pilot scale tests while the dose of DZP were based on previous scientific reports (Kazdoba et al., 2015; Zemdegs et al., 2018)

Elevated Plus-Maze Test

The elevated plus maze (EPM) was conducted in accordance with the method validated by Pellow et al. (1985).The maze comprised of two wood open arms (50 × 10 cm) surrounded by a short (0.5 cm) acrylic edge to avoid falls and two wood enclosed arms (50 × 10 × 40 cm) arranged such that the two open arms were opposite to each other. The arms were connected by a central platform (10 × 10 cm) and the maze was kept 50 cm above the floor. Each rat was placed in the middle compartment (head facing an open arm) and allowed to freely explore the apparatus for 5 min. It was measured (i) the time spent(s) in the

open and closed arms of the EPM, (ii) the percentage of time spent in open arms, (iii) protected and unprotected stretch attend postures, and (iv) protected and unprotected head dipping events. Stretch-attend was defined when the animal stretches forward and retracts without moving its feet. Head dipping was defined as the exploratory movement in which the rodent head protruding over the edge toward the floor on the open arms. Behaviors were defined unprotected when they were exhibited in the open arm region of the EPM and were protected if it occurs in closed arms or central platform of the maze (Walf and Frye, 2007).

Open Field Test

The open field test was performed according to the method previously described (Funchal and Dani, 2014).The open-field used was a circular arena measuring 60 cm in diameter with 50-cm high circular acrylic wall (OP0199, Insight). The apparatus floor was divided into 12 squares (4 squares corresponding a central zone and 8 squares to the peripheral zones).Every animal was placed in the central region of the open field, and the number of lines crossed (all four limbs), rearing, grooming, and the times spent(s) in the central zone of the apparatus were recorded for 5 minutes. Also, for an indirect measurement of locomotor performance of animals, the number of crossings was checked in time blocks along the experiment comparing control and experimental groups, since the animals usually explore more of the apparatus in the first blocks of time and less at the end of the task, as they habituate.

Light-Dark Box Test

The light-dark box consist of a *box* (46 x 27 x 30 cm) divided into a dark and illuminated compartment (EP 158, Insight, fluorescent lamp- 20 W). The chambers were connected by a small opening (7.5 × 7.5 cm) in the middle of the wall separating the two chambers. Rats were placed in the middle of the light chamber and then released to explore for 5 min (Hascoet and Bourin, 1998).The time that rats spent in the light/dark compartments and the number of transitions between them were recorded. Transition was defined as the placement of all four paws in the entry chamber.

ASSESSMENT OF ANTICONVULSANT ACTIVITY OF VITEXIN

The anticonvulsant effect of vitexin was evaluated against chemically-induced seizures by *intraperitoneal (i.p.) injection* of proconvulsant agents: NMDA, 150 mg/kg; kainic acid, 30 mg/kg; picrotoxin, 6 mg/kg; and PTZ 90, mg/kg. NMDA, kainic acid, picrotoxin, and PTZ were dissolved in 0.9% saline while vitexin dissolved in 0.9% saline containing dimethyl sulfoxyde – DMSO 1%. The animals control and experimental groups were divided according to the assessment of anxiolytic-like effects of vitexin, except by the range of vitexin i.p. doses, which were slightly up changed for anticonvulsant screening (1.25, 2.5, and 5 mg/kg). Doses of vitexin were selected based on pilot scale tests while the dose of DZP were selected based on previous scientific reports, as considered above for anxiolytic-related tests. Thirty minutes after the administration of each proconvulsant drug, each animal was maintained individually in a transparent *acrylic arena* (60 cm × 40 cm). The rats were observed for 30, 40, 40, and 120 min respectively for kainic acid, NMDA, picrotoxin, and PTZ assays. It was evaluated the latency time to the first generalized tonic-clonic seizure and the incidence/number of animals exhibiting generalized tonic-clonic seizures (Funchal and Dani, 2014). Animal behavior seizure rating was classified/certified according to the Racine's scale, modified by Pinel and Rovner (Pinel and Rovner, 1978). Generalized tonic-clonic seizures were considered as those rating score 7 or more in accordance to this scale and analyses of our video-recorded files.

Statistical Analysis

Except by the Open-Field test which was performed one and two-way ANOVA test, other experiments statistical analyses were performed using One-way ANOVA, followed by the Tukey post-hoc test (GraphPad Prism; version 7.0, Graphpad Software, USA). Data were expressed as the means ± standard error of mean (SEM) (n=06/group). A level of $p< 0.05$ was accepted as statistically significant.

RESULTS

Anxiolytic-Like Effects of Vitexin

Elevated Plus-Maze

The amount of time (sec) spent in the open and closed arms and percentage of time spent in open arms in the elevated plus maze are shown in **Figure 1**. The *post hoc* test showed that the time spent on the open arms was longer ($p < 0.05$) in the diazepam treated group (positive control). Also, the data shows that all doses of vitexin increased significantly the time spent in the open arms compared to the saline control ($F_{(5,60)} = 103.9$; $p<0.05$). Pretreatment with vitexin produced a significant increase in unprotected head dipping ($F_{(5,30)} = 74.25$; $p<0.05$) and unprotected stretch-attend postures ($F_{(5,30)} = 27.99$; $p<0.05$).

Open Field Test

The number of line crossings during open field test showed that vitexin increased locomotion on apparatus ($F_{(5,26)} = 45.33$; $p<0.05$) (**Figure 2A**). Vitexin significantly improved locomotion mainly in the first 3 minutes of experimentation when compared with the control group ($F_{(20, 140)} = 26,23$ $p<0.05$) (**Figure 2B**). In the open field test, animals treated with vitexin (all doses) spent more time on the center of arena when compared with those from the control group ($F_{(5,30)} = 76.82$; $p<0.05$) (**Figure 2C**). Vertical activity behaviors (rearing and grooming) are shown in **Figure 2D**. The analysis of total number of the rearing showed significantly higher for vitexin treatments than that in the control group (saline) ($F_{(5,27)} = 67.49$; $p<0.05$). The *post hoc* test revealed that vitexin 0.75 and 1.25 mg/kg reduced grooming behaviors (compared with the control group) in a way similar to the diazepam ($F_{(5,27)} = 40.02$; $p<0.05$) (**Figure 2D**).

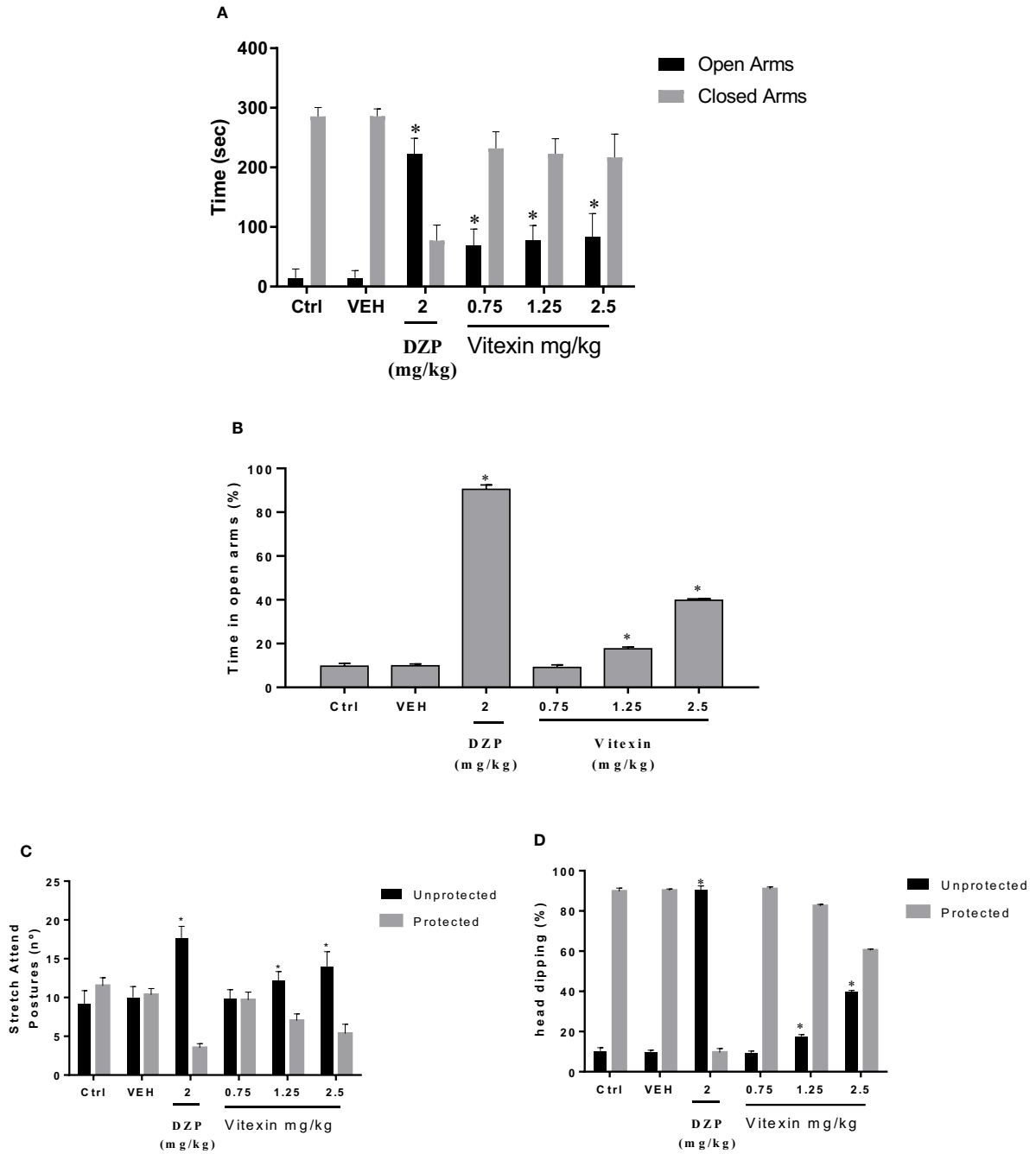

FIGURE 1 | Elevated Plus Maze. **(A)** Total time spent on the open and closed arms, **(B)** percentage of time spent in opens arms, **(C)** protected and unprotected stretch attend postures and **(D)** protected and unprotected head dipping between the groups treated with different doses of vitexin, diazepam (DZP), and vehicle (VEH), compared with control group (Ctrl - Saline Control). Significantly different from saline control group: $^{*}p < 0.05$. One-way ANOVA; Tukey *post hoc* test.

Light-Dark Test

Vitexin at the dose of 0.75, 1.25, and 2.5 mg/kg and diazepam (2 mg/kg) induced a significant increase in the time spent by rats on the illuminated side of the apparatus compared with the saline control group ($F_{(5;50)}$ = 26.87; p<0.05) (**Figure 3A**). Analysis of the light/dark transitions revealed significant greater number of transitions for animals treated with vitexin than in the control group ($F_{(5,28)}$ = 21.33; p<0.05) (**Figure 3B**).

Anticonvulsant Screening

Acute administration of vitexin prior to kainic acid and NMDA injections did not protect animals against seizures. However,

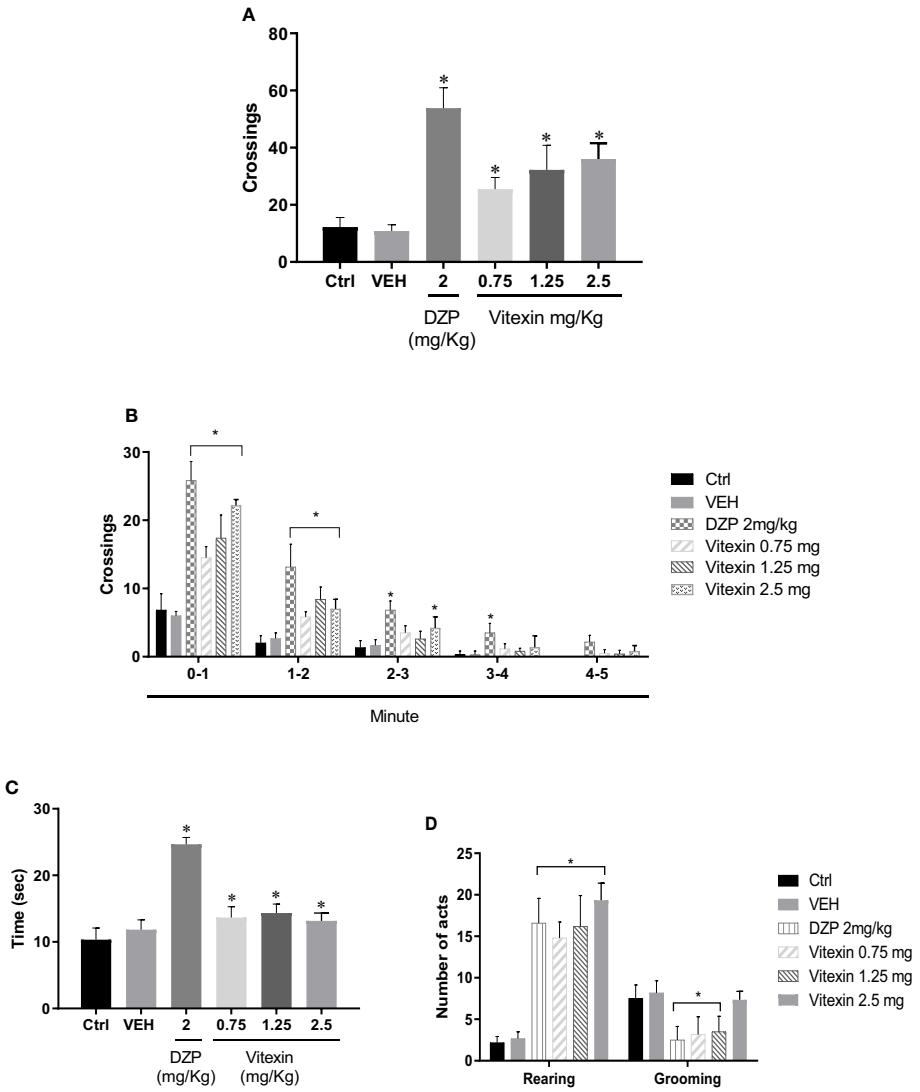

FIGURE 2 | Open Field Test. **(A)** number of crossings made in the open field, **(B)** Crossings in function of the time, **(C)** time spent in the center and **(D)** rearing and grooming behavior of the open field by different doses of vitexin, diazepam (DZP), and vehicle (VEH), compared with control group (Ctrl—saline control). Significantly different from saline control: $^*p < 0.05$. For data in **(A, C** and **D)** was performed one-way ANOVA and in **(B)** was performed two-way ANOVA. For all data were performed Tukey *post hoc* test.

clonic seizures induced by i.p. administrations of PTZ and picrotoxin were inhibited by all doses of vitexin (**Figures 4A, B**). Vitexin significantly protected animals against seizures and showed 16.66, 50 and 100% of protection against picrotoxin-induced seizures (**Figure 4C**) and 16.66%, 33.33%, and 100% against PTZ-induced seizures (**Figure 4D**), respectively, on the doses of 1.25; 2.5, and 5 mg/kg. Therefore, the anticonvulsant activity of vitexin exhibited a dose-dependent profile on PTZ and picrotoxin assays. In addition, the latency to to the first seizure induced by PTZ and picrotoxin was significantly higher for groups treated with vitexin, particularly on the dose of 2.5 mg/kg (**Table 1**).

DISCUSSION

In the present study we demonstrated that vitexin presents both anxiolytic and anticonvulsant properties which are possibly mediated by modulation of GABAergic neurotransmission and/or related synaptic pathways. Since those pharmacological activities were achieved after systemic administration (i.p.) of vitexin, it is plausible to affirm that this compound overcomes the blood-brain barrier, making it more attractive in terms of pharmacokinetic or even pharmaceutical perspectives. Vitexin has been found at different medicinal plants with anxiolytic and/or anti-epileptic properties (Nassiri-Asl et al., 2007; Miroddi

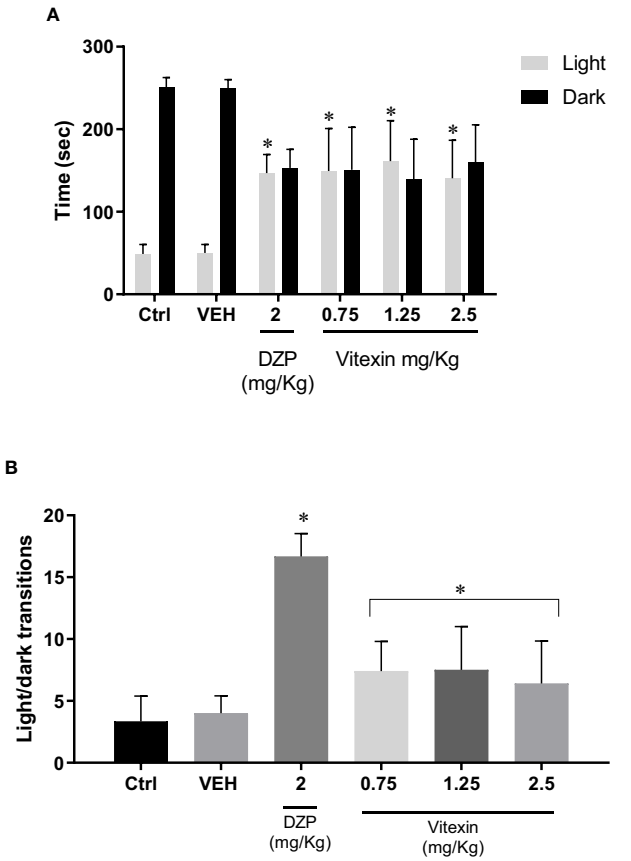

FIGURE 3 | Light dark box test. **(A)** times spent(s) on light/dark area and **(B)** light/dark transitions by different doses of vitexin, diazepam (DZP) and vehicle (VEH) compared with control group (Ctrl—saline control). Significantly different from saline control: $^{*}p < 0.05$. One-way ANOVA; Tukey *post hoc* test.

et al., 2013). Based on this and also on our results, it is also possible to suggest that vitexin is at least partially responsible for those pharmacological activities, in most of cases and more probabilistically, acting in a synergistic way with other(s) active compound(s) present in those plants claimed as anxiolytics and/or anti-epileptics.

In the present study, anxiolytic-like effect of vitexin was confirmed by three different and complementary animal screening tests. This combined arrangement of tests allows a more reliable and complete evaluation of anxiety-related behaviors in rodents (Ramos et al., 2008). Overall results from different tests confirmed the increase of exploratory performance and the decrease of aversive behaviors of animals, indicating an anxiolytic-like effect promoted by the different doses of vitexin. Anxiolytics agents reduce protected risk assessment behaviors and increase the frequency of unprotected ethological behaviors (Benneh et al., 2018). Similar to diazepam, pretreatment with vitexin significantly decreased protected head dips and stretch attend postures and increase unprotected behavior suggesting the anxiolytic property of vitexin. In addition, it is possible to reinforce that vitexin did not caused any type of locomotor impairment or even did not produced signals of sedation on animals at same time that the anticonvulsant activity produced by vitexin is not caused by interference with locomotor capacity of the treated animals. Those conclusions are supported by the increased number of crossings at open field test and the increased number of light/dark transitions exhibited by vitexin treated animals in the light/dark box test when compared with the saline control. Locomotor impairment or severe sedation are constantly claimed as important side effects caused by different pharmacological agents used in the treatment of anxiety and epilepsy and should be avoided in the clinical management of those medical conditions (Canevini et al., 2010; Guina and Merrill, 2018). Therefore, besides the useful dual pharmacological effect (anxiolytic and anticonvulsant) vitexin presents some advantages in terms of its toxicological profile. The dual effect of vitexin at a very similar range of doses is particularly important for a potential combined and future management of both epilepsy and anxiety, which commonly appears as comorbidities and request the use of several combined drugs. However, this dual pharmacological effect is not surprising or exclusive of vitexin, since epilepsy and anxiety are partially overlapped in terms of neurochemical and neurobiological pathways and some drugs depending on the dose (like benzodiazepines) may act against both disorders despite important observed side effects.

The effect of acute intraperitoneal injection of the vitexin for the prospect of its anticonvulsant activity was evaluated using different acute seizure models in which the seizures were chemically induced. Given the possibility of false negatives results on drug screening, potential antiepileptic compounds should not be screened solely against one single model but to be tested across a few different seizure models, exploring at same time different neurotransmission pathways involved on initiation and spraying of induced seizures (Yuen and Troconiz, 2015). These type of animal models are very useful for the rapid and economical screening of new anticonvulsant drugs (Löscher, 2011). Kainic acid and NMDA have significant convulsant effects, fundamentally by activating glutamate receptors; glutamate is the most prominent excitatory neurotransmitter in the mammalian CNS (Bhandage et al., 2017). Moreover, in the current study, we explored the therapeutic potential of vitexin using the $GABA_A$ receptor antagonists PTZ and picrotoxin (Colas et al., 2013). On the basis of the results presented here, acute treatment with different doses of vitexin applied against the kainic acid and NMDA seizures protocol was not effective when compared with the saline control. In contrast, vitexin exhibited a very clear dose-dependent anticonvulsant effect against seizures induced by PTZ and picrotoxin, reaching the maximum protection (100%) at higher dose of 5 mg/kg. In addition, the lower doses (1.25 and 2.5 mg/kg) delayed the onset of seizures for non-completely protected animals against seizures. The combination of these actions indicated an effective anticonvulsant effect of vitexin at least against PTZ and picrotoxin-induced seizures.

The underlying molecular mechanisms involved in the anticonvulsant or anxiolytic actions of vitexin were not the central scope of the present study. However, the selective protection offered by vitexin against GABA receptors antagonists at the expense of Glutamate receptor agonists are

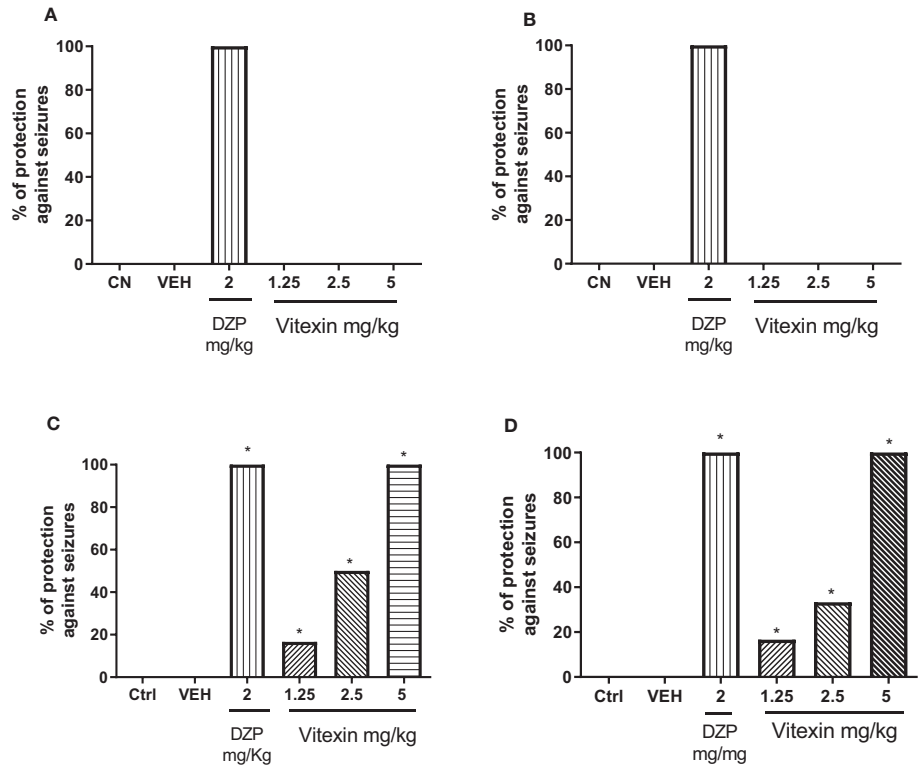

FIGURE 4 | Percentages of animals protected against seizures induced by **(A)** kainic acid 150 mg/kg, **(B)** NMDA 150 mg/kg, **(C)** picrotoxin 6 mg/kg and **(D)** PTZ 90 mg/kg in function of different doses of vitexin, diazepam (DZP), and vehicle (VEH) compared with control group (Ctrl-Saline Control). Significantly different from control: *$p < 0.05$ versus saline control group. One-way ANOVA; Tukey *post hoc* test.

TABLE 1 | Latency to onset of seizure (score 7-8) on vitexin treatments.

Convulsant Agent	Latency to onset of seizure (seconds)					
	Ctrl	VEH	DZP (2 mg/kg)	Vitexin (1.25 mg/kg)	Vitexin (2.5 mg/kg)	Vitexin (5 mg/kg)
NMDA (150 mg/kg)	4409.50 ± 92.48	4326.30 ± 43.49	NS	4403.83 ± 98.63	4514.50 ± 63.59	4548.50 ± 70.52
Kainic acid (30 mg/kg)	4491.33 ± 77.88	4546.33 ± 77.52	NS	4693.0 ± 74.59	4404.83 ± 116.83	4468.72 ± 124.72
PTZ (90 mg/kg)	77.4 ± 14.63	78.00 ± 7.38	NS	53.66 ± 9.07$^{\#}$	141.2 ± 29.08$^{*\#}$	$^{*\#}$ NS
Picrotoxin (6 mg/kg)	1101.0 ± 47.29	1103.66 ± 47.99	NS	745.8 ± 65.35$^{*\#}$	1396.6 ± 43.14$^{*\#}$	$^{*\#}$ NS

$p < 0.05$ versus control group (Ctrl – Saline Control) and $^{\#}p < 0.05$ between different doses of vitexin. One-way ANOVA; Tukey post hoc test. NS, no seizure.

at least intriguing and may suggest that the modulation of GABA neurotransmission or related pathways may be involved on mode of action of vitexin. It is noteworthy to point out that different natural flavonoids and various synthetic derivatives from natural flavonoids are well known to be positive modulators of $GABA_A$ receptors (Hanrahan et al., 2011; Oliveira et al., 2018). Indeed, many studies agree that vitexin may exert its CNS effects *via* $GABA_A$ receptor (benzodiazepine site) modulation (Abbasi et al., 2012; Oliveira et al., 2014; World Health Organization, 2019b). However, further neurochemical investigations are required for a more complete understanding by the way vitexin exerts its pharmacological activities. The GABA neurotransmission is a common link between epilepsy and anxiety. In fact, the down regulation of GABA activity is frequently evoked on the neurobiological explanation for both disorders. Also, the up modulation of GABAergic neurotransmission is assumed on the explanation about the mode of action of some drugs such as benzodiazepines, which are well known as a $GABA_A$ receptors positive allosteric modulators and commonly used for the clinical management for both disorders (Beyenburg et al., 2005).

Finally, epilepsy and anxiety frequently appear as comorbidities. Indeed, one of the main psychiatric conditions observed for epileptic patients is anxiety. Anxiety affects approximately 20% of the people diagnosed with epilepsy and is responsible for a

decreased quality of life and can worsen epilepsy control (Dworetzky, 2017; World Health Organization, 2019b). The actual pharmacological arsenal used in the clinical treatment for both disorders often causes important side effects, such as cognitive/locomotor impairments and sedation, which negatively affects therapeutic adherence (Diniz et al., 2019). Thus, the prospecting/developing of new pharmacological probes is relevant especially when those new potential drugs exhibit potential advantages in terms of pharmacokinetics, toxicological, and pharmacodynamic features.

CONCLUSION

The results of this study showed that vitexin possess both anxiolytic and anticonvulsant properties. Since vitexin has been found in different medicinal plants popularly claimed as anxiolytics and/or anticonvulsants, this work has both ethnopharmacological and pharmaceutical relevance. This work may be useful for the development of new pharmacological probes not only for anxiety and epilepsy patients relieve in the future but also, and indirectly, for the better understanding of pathological events underlying both anxiety and epilepsy themselves.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by University of Ribeirão Preto Ethic Committee (11/2015).

AUTHOR CONTRIBUTIONS

Conceived and Designed: DO and RB. Data Collection: DO, CS and BI. Acquisition, Analysis and Interpretation of Data: DO and CS. Article Writing by: DO and RB. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Brazilian National Council for Scientific and Technological Development (CNPq).

ACKNOWLEDGMENTS

The authors appreciate the financial support from (CAPES and CNPq) and equally thanking the Biotechnology Department of University of Ribeirão Preto (UNAERP) for the infrastructure, technical support and facilities to carry out this work.

REFERENCES

Abbasi, E., Nassiri-Asl, M., Shafeei, M., and Sheikhi, M. (2012). Neuroprotective Effects of Vitexin, a Flavonoid, on Pentylenetetrazole-Induced Seizure in Rats. *Chem. Biol. Drug Des.* 80 (2), 274–278. doi: 10.1111/j.1747-0285.2012.01400.x

Abbasi, E., Nassiri-Asl, M., Sheikhi, M., and Shafiee, M. (2013). Effects of vitexin on scopolamine induced memory impairment in rats. *Chin. J. Physiol.* 56 (3), 184–189. doi: 10.4077/CJP.2013.BAB123

Benneh, C., Biney, R. P., Adongo, D. W., Mante, P. K., Ampadu, F. A., Tandoh, A., et al. (2018). Anxiolytic and Antidepressant Effects of *Maerua angolensis DC.* Stem Bark Extract in Mice. *Depression Res. Treat*, 1–18. doi: 10.1155/2018/1537371

Beyenburg, S., Mitchell, A. J., Schmidt, D., Elger, C. E., and Reuber, M. (2005). Anxiety in patients with epilepsy: Systematic review and suggestions for clinical management. *Epilepsy Behav.* 7 (2), 161–171. doi: 10.1016/j.yebeh.2005.05.014

Bhandage, A. K., Jin, Z., Hellgren, C., Korol, S. V., Nowak, K., Williamsson, L., et al. (2017). AMPA, NMDA and kainate glutamate receptor subunits are expressed in human peripheral blood mononuclear cells (PBMCs) where the expression of GluK4 is altered by pregnancy and GluN2D by depression in pregnant women. *J. Neuroimmunol.* 305, 51–58. doi: 10.1016/j.jneuroim.2017.01.013

Bialer, M., and White, H. S. (2010). Key factors in the discovery and development of new antiepileptic drugs. *Natural Rev. Drug Discovery* 9, 68–82. doi: 10.1038/nrd2997

Can, O. D., Demir Ozkay, U., and Ucel, U. I. (2013). Anti-depressant-like effect of vitexin in BALB/c mice and evidence for the involvement of monoaminergic mechanisms. *Europen J. Pharmacol.* 699, 250–257. doi: 10.1016/j.ejphar.2012.10.017

Canevini, M. P., Sarro, G., Galimberti, C. A., Gatti, G., Liccheta, L., Malerba, A., et al. (2010). Relationship between adverse effects of antiepileptic drugs, number of coprescribed drugs, and drug load in a large cohort of consecutive patients with drug-refractory epilepsy. *Epilepsia* 51 (5), 797–804. doi: 10.1111/j.1528-1167.2010.02520.x

Colas, D., Chuluun, B., Warrier, D., Blank, M., Wetmore, D. Z., Buckmaster, P., et al. (2013). Short-term treatment with the GABAA receptor antagonist pentylenetetrazole produces a sustained pro-cognitive benefit in a mouse model of Down's syndrome. *Br. J. Pharmacol.* 169 (5), 963–973. doi: 10.1111/bph.12169

Diniz, T. C., Júnior, R. G. O., Medeiros, M. A. M. B. , Silva, M. G., Teles, R. B. A., Menezes, P. P., et al. (2019). Anticonvulsant, sedative, anxiolytic and antidepressant activities of the essential oil of Annona vepretorum in mice: Involvement of GABAergic and serotonergic systems, (2019). *Biomed. Pharmacother.* 111, 1074–1087. doi: 10.1016/j.biopha.2018.12.114

Dworetzky, B. A. (2017). Worrying More about Anxiety in Patients with Epilepsy. *Epilepsy Curr.* 17 (6), 353–354. doi: 10.5698/1535-7597.17.6.353

England, M. J., Liverman, C. T., Schultz, A. M., and Strawbridge, L. M. (2012). Epilepsy across 455 the spectrum: Promoting health and understanding. *Epilepsy Behav. E&B* 25 (2), 266–276. doi: 10.1016/j.yebeh.2012.06.016

Fisher, P. L., and Noble, A. J. (2017). Anxiety and depression in people with epilepsy: The contribution of metacognitive beliefs. *Seizure* 50, 153–159. doi: 10.1016/j.seizure.2017.06.012

Funchal, C., and Dani, C. (2014). Neurosciências: Modelos Experimentais Animais. *EdiPUCRS*, 280.

Gandy, M., Sharpe, L., Perry, K. N., Miller, L., Thayer, Z., Boseiro, J., et al. (2015). Anxiety in epilepsy: a neglected disorder. *J. Psychosom. Res.* 78, 149–155. doi: 10.1016/j.jpsychores.2014.12.002

Gauthier, I., and Nuss, P. (2015). Anxiety disorders and GABA neurotransmission: a disturbance of modulation. *Neuropsychiatr. Dis. Treat.* 11, 165–175. doi: 10.2147/NDT.S58841

Guimarães, C. C., Oliveira, D. D., Valdevite, M., Fachin, A. L. S., Pereira, S. I. V., França, S. C., et al. (2015). The glycosylated flavonoids vitexin, isovitexin, and quercetrin isolated from Serjania erecta Radlk (Sapindaceae) leaves protect

PC12 cells against amyloid-beta peptide-induced toxicity. *Food Chem. Toxicol.* 86, 88–94. doi: 10.1016/j.fct.2015.09.002

Guina, J., and Merrill, B. (2018). Benzodiazepines I: Upping the Care on Downers. *J. Clin. Med.* 7 (2), 1–22. doi: 10.3390/jcm7020017

Hanrahan, J. R., Chebib, M., and Johnston, G. A. R. (2011). Flavonoid modulation of GABAA receptors. *Br. J. Pharmacol.* 163 (2), 234–245. doi: 10.1111/j.1476-5381.2011.01228.x

Hascoet, M., and Bourin, M. (1998). A new approach to the light/dark procedure in mice. *Pharmacol. Biochem. Behav.* 60, 645–653. doi: 10.1016/s0091-3057(98)00031-8

He, M., Min, J. W., Kong, W. L., He, X. H., Li, J. X., and Peng, B. W. (2016). A review on the pharmacological effects of vitexin and isovitexin. *Fitoterapia* 115, 74–85. doi: 10.1016/j.fitote.2016.09.011

Kazdoba, T. M., Hagerman, R. J., Zolkowska, D., Rogawski, M. L. A., and Crawley, J. N. (2015). Evaluation of the neuroactive steroid ganaxolone on social and repetitive behaviors in the BTBR mouse model of autism. *Psychopharmacology* 233, 309–323. doi: 10.1007/s00213-015-4115-7

Kwon, O., and Park, S. (2014). Depression and Anxiety in People with Epilepsy. *J. Clin. Neurol.* 10 (3), 175–188. doi: 10.3988/jcn.2014.10.3.175

Löscher, W. (2011). Critical review of current animal models of seizures and epilepsy used in the discovery and development of new antiepileptic drugs. *Seizure* 20 (5), 359–368. doi: 10.1016/j.seizure.2011.01.003

Martin, E.II, Ressler, K. J., Binder, E., and Nemeroff, C. B. (2009). The Neurobiology of Anxiety Disorders: Brain Imaging, Genetics, and pychoneuroendocrinology. *Psychiatr. Clinics North America* 32 (3), 549–575. doi: 10.1016/j.psc.2009.05.00

Masneuf, S., Lowery-Gionta, E., Colacicco, G., Pleil, K. E., Li, C., Crowley, N., et al. (2014). Glutamatergic mechanisms associated with stress-induced amygdala excitability and anxiety-related behavior. *Neuropharmacology* 85, 190–197. doi: 10.1016/j.neuropharm.2014.04.015

Miroddi, M., Calapai, G., Navarra, M., Minciullo, P. L., and Gangemi, S. (2013). Passiflora incarnata L.: Ethnopharmacology, clinical application, safety and evaluation of clinical trials. *J. Ethnopharmacol.* 150 (3), 791–804. doi: 10.1016/j.jep.2013.09.047

Moto, F. C. O., Arsa'a, A., Ngoupaye, G. T., Taiwe, G. S., Njapdounke, J. S. K., Kandeda, A. K., et al. (2018). Anxiolytic and Antiepileptic Properties of the Aqueous Extract of *Cissus quadrangularis* (Vitaceae) in Mice Pilocarpine Model of Epilepsy. *Front. Pharmacol.* 9, 751. doi: 10.3389/fphar.2018.00751

Nassiri-Asl, M., Shariati-Rad, S., and Zamansoltani, F. (2007). Anticonvulsant effects of aerial parts of *Passiflora incarnata* extract in mice: involvement of benzodiazepine and opioid receptors. *BMC Complement. Altern. Med.* 7. doi: 10.1186/1472-6882-7-26

Oliveira, D. R., Zamberlam, C. R., Gaiardo, R. B., Rêgo, G. M., Cerutti, J. M., Cavalheiro, A. J., et al. (2014). Flavones from *Erythrina falcata are* modulators of fear memory. *BMC Complement. Altern. Med.* 14 (1), 1–17. doi: 10.1186/1472-6882-14-288

Oliveira, D. R., Todo, A. H., Rêgo, G. M., Cerutti, J. M., Cavalheiro, A. J., Rando, D. G. G., et al. (2018). Flavones-bound in Benzodiazepine Site on GABA A Receptor: Concomitant Anxiolytic-Like and Cognitive-Enhancing Effects Produced by Isovitexin and 6-C-glycoside-Diosmetin. *Eur. J. Pharmacol.* 15 (831), 77–86. doi: 10.1016/j.ejphar.2018.05.004

Pellow, S., Chopin, P., File, S. E., and Briley, M. (1985). Validation of open: closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *J. Neurosci. Methods* 14, 149–167. doi: 10.1016/0165-0270(85)90031-7

Pinel, J. P., and Rovner, L.II (1978). Electrode pacement and kindling: induced experimental epilepsy. *Exp. Neurol.* 58 (15), 35–346-202. doi: 10.1016/0014-4886(78)90145-0

Ramos, A., Pereira, E., Martins, G. C., Wehrmeister, T. D., and Izidio, G. S. (2008). Integrating the open field, elevated plus maze and light/dark box to assess different types of emotional behaviors in one single trial. *Behav. Brain Res.* 193, 277–288. doi: 10.1016/j.bbr.2008.06.007

Salpekar, J. A., and Mula, M. (2018). Common psychiatric comorbidities in epilepsy: How big of a problem is it? *Epilepsy Behav.* 18, 1–5. doi: 10.1016/j.yebeh.2018.07.023

Walf, A. A., and Frye, A. C. (2007). The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. *Nat. Protoc.* 2, 322–328. doi: 10.1038/nprot.2007.44

World Health Organization (2019a). *Epilepsy key facts*. Available at: https://www.who.int/news-room/fact-sheets/detail/epilepsy (Accessed October 04, 2019).

World Health Organization (2019b). *Epilepsy: a public health imperative*. Available at: https://www.who.int/mental_health/neurology/epilepsy/report_2019/en/ (Accessed October 15, 2019).

Wu, Q., Zhao, C. W., Long, Z., Xiao, B., and Feng, L. (2018). Anatomy Based Networks and Topology Alteration in Seizure-Related Cognitive Outcomes. *Front. Neuroanat.* 12, 25. doi: 10.3389/fnana.2018.00025

Yuen, E. S. M., and Troconiz, I. F. (2015). Can pentylenetetrazole and maximal electroshock rodent seizure models quantitatively predict antiepileptic efficacy in humans? *Seizure* 24, 21–27. doi: 10.1016/j.seizure.2014.11.006

Zemdegs, J., Rainer, Q., Grossmann, C. P., Rousseau-Ralliard, D., Grynberg, A., Ribeiro, E., et al. (2018). Anxiolytic- and Antidepressant-Like Effects of Fish Oil-Enriched Diet in Brain-Derived Neurotrophic Factor Deficient Mice. *Front. Neurosci.* 12:2018.00974. doi: 10.3389/fnins.2018.00974

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 August 2020
doi: 10.3389/fpls.2020.01256

Transcriptome and Flavonoids Metabolomic Analysis Identifies Regulatory Networks and Hub Genes in Black and White Fruits of *Lycium ruthenicum* Murray

Tingting Li[1†], Yunfang Fan[2,3†], Huan Qin[1], Guoli Dai[2,3], Guoxiu Li[1], Yanlong Li[2,3], Jingjin Wang[1], Yue Yin[2,3], Fang Chen[1], Xiaoya Qin[2,3], Youlong Cao[2,3] and Lin Tang[1*]*

[1] Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China, [2] Institute of Wolfberry Engineering Technology, Ningxia Academy of Agriculture and Forestry Sciences, Yinchuan, China, [3] National Wolfberry Engineering Technology Research Center, Yinchuan, China

Lycium ruthenicum Murry. is a highly nutritional cash crop due to its fruit abundant anthocyanins. To understand the complex metabolic networks underlying the color formation in black and white fruits of *L. ruthenicum*, we conducted transcriptome and flavonoid metabolic profiling to identify the candidate genes possibly involved in flavonoid biosynthesis. As a result, 147 flavonoids were identified and there was almost no anthocyanin in white fruits, while luteolin, kaempferol, and quercetin derivatives showed markedly higher abundance. Furthermore, applying weighted gene co-expression network analyses, 3 MYB, 2 bHLH, 1WRKY and 1 NAC transcription factor, associated with anthocyanin biosynthesis were identified. A bHLH transcription factor, *LrAN1b* showed the greatest correlations with anthocyanin accumulation with no expression in white fruits. In addition, gene function analysis and qRT-PCR experiments identified a new activated anthocyanin MYB transcription factor designed as *LrAN2-like*. Yeast two-hybrid and transient tobacco overexpression experiments showed that *LrAN1b* could interact with *LrAN2-like* and *LrAN11* to form MBW complex to activate the anthocyanin pathway. The yeast one-hybrid experiment indicated that *LrAN2-like* bonded anthocyanin structural gene *LrDFR* and *LrANS* promoters. Heterologous expression of *LrAN1b* in tobacco can significantly increase the anthocyanin content of tobacco florals and capsules, and activate anthocyanin synthesis related genes. Taken together, an anthocyanin regulatory network model in *L. ruthenicum* fruit was proposed firstly and we speculate that the white fruit phenotype was due to abnormal expression of *LrAN1b*. The findings provide new insight into the underlying mechanism of flavonoids, laying the foundation for future functional and molecular biological research in *L. ruthenicum*.

Keywords: anthocyanin, fruits, flavonoid, *Lycium ruthenicum*, metabolic profiling, transcriptome

OPEN ACCESS

Edited by:
Cristina Garcia-Viguera, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
Shouchuang Wang, Hainan University, China
Atsushi Fukushima, RIKEN, Japan

***Correspondence:**
Youlong Cao
youlongchk@163.com
Lin Tang
tangl666@126.com

[†]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 May 2020*
Accepted: *30 July 2020*
Published: *14 August 2020*

Citation:
Li T, Fan Y, Qin H, Dai G, Li G, Li Y, Wang J, Yin Y, Chen F, Qin X, Cao Y and Tang L (2020) Transcriptome and Flavonoids Metabolomic Analysis Identifies Regulatory Networks and Hub Genes in Black and White Fruits of Lycium ruthenicum Murray. Front. Plant Sci. 11:1256. doi: 10.3389/fpls.2020.01256

INTRODUCTION

Anthocyanins are an important subclass of water-soluble flavonoids pigments that can give plants a bright red to blue color. Anthocyanin has many beneficial functions for human health, such as antioxidant activity (Jain et al., 2008) inhibition of tumor cells and prevention of chronic human diseases (Butelli et al., 2008). Anthocyanin is a downstream product of the flavonoid pathway. It takes 4-coumaroyl-CoA and malonyl-CoA as substrates; is synthesized by a series of enzymes encoded by the structural genes chalcone synthase (CHS), chalcone isomerase (CHI), flavonoid 3-hydroxylase (F3H), flavonoid 3′-hydroxylase (F3'H), flavonoid 3′5′-hydroxylase (F3'5'H), dihydroflavonol 4-reductase (DFR) and anthocyanidin synthase (ANS); and is transported to vacuoles after glycosylation, methylation, and acylation (Tanaka et al., 2010). In recent years, these structural genes and modification- and transport-related genes are mainly regulated by R2RMYB transcription factors, basic helix-loop-helix proteins (bHLH), and WD40 protein (MBW complexes) (Koes et al., 2005). Among them, MYB and bHLH can specifically bind cis-elements of structural gene promoter regions named MRE (MYB-recognizing element) and BRE (bHLH-recognizing element) (Zhu et al., 2015). In contrast, some studies have demonstrated that *Arabidopsis CPC* and *MYBL2* can inhibit anthocyanin biosynthesis by inhibiting the assembly of ternary MBW (Matsui et al., 2008; HF et al., 2009). In addition, other transcription factors also participate in anthocyanin regulation including WRKY protein, NAC protein, etc (Morishita et al., 2009; Verweij et al., 2016). The molecular mechanism of fruit color mutation has always been an important issue for researchers, because fruit color is closely related to the nutritional value, appearance, and taste of the fruit (Paauw et al., 2019). Anthocyanin pathway biosynthesis or transcription regulation encoding gene mutations are related to color phenotype. In red pulp apples, multiple repeats of the transcription factor MYB10 promoter fragment result in automatic regulation of transcription factors (Espley et al., 2009). In strawberries, sequence changes in the upstream regulatory region of FnMYB10 resulted in low expression levels of the FnMYB10 gene, which most likely resulted in the white fruit phenotype of *F. nilgerrensis* (Zhang J. et al., 2020). For *L. ruthenicum*, an important economic crop as a source of anthocyanins, genetic control of the fruit color and exploration of germplasm resources are of particular concern.

Lycium ruthenicum Murry. is a small bush belonging to the Solanaceae family the fruits of which are used widely as ethnic medicine and nutraceutical food (Hu et al., 2014). The genus *Lycium* (Solanaceae) approximately 80 species and widely grows in arid to semi-arid environments of the temperate zones, and it is distributed worldwide in very different habitats (Levin and Miller, 2005). The fruits behave as functional components and are beneficial for human health; their compounds are mainly polyphenols such as flavonols, anthocyanins, and catechins, strong natural antioxidants (Wang et al., 2018). Modern medical experiments show that *L. ruthenicum* fruit extracts can resist pancreatic ductal adenocarcinoma cell activity (Zhang S. et al., 2019) and cure diabetic cardiomyopathy (Xue et al., 2014). Recent evidence suggests that the main active ingredient in this species is petunidin-3-*O*-rutinoside (trans-p-coumaroyl)-5-*O*-glucoside (Zheng et al., 2011). In the latest study, 49 compounds including anthocyanins, alkaloids, hydroxycinnamic acid derivatives, flavonoids and amino acids were initially identified in the four developmental stages of *L. ruthenicum* fruit (Yang et al., 2019). In recent years, a wild white berry variety without anthocyanin of *L. ruthenicum* has been found in Qinghai Province China (Yin-Yan et al., 2017). However, there are fewer report illustrating the ; secondary metabolites and molecular mechanism in the white fruits of *L. ruthenicum*. Previous research has identified the anthocyanin-related structural genes and transcription factors including MYB protein (*LrAN2* and *LrMYB113*), bHLH protein (*LrJAF13* and *LrAN1b*), and WD40 protein (*LrAN11*) in *L. ruthenicum* (Zeng et al., 2014). Recent studies claim that the functional diversity and high expression levels of *LrAN2* may be responsible for the high content of anthocyanins in *L. ruthenicum* fruits (Zong et al., 2019).

In this study, to obtain insights into the changes of flavonoid metabolites and transcriptional regulation of the anthocyanin synthetic pathways in black and white fruits of *L. ruthenicum*. We performed flavonoid metabolites of mature fruits of two colors using liquid chromatography tandem mass spectrometry (LC-MS/MS). Transcriptome sequencing of two fruit colors was performed on five distinct fruit development stages, and differentially expressed flavonoid and anthocyanin structural genes were identified. Furthermore, we obtained anthocyanin-related co-expression gene modules and screened some key synthetic and regulatory genes involved in anthocyanin synthesis by WGCNA. Accordingly, the expression levels of biosynthetic genes and regulatory genes of two different fruit colors were analyzed by quantitative real-time polymerase chain reaction (qRT-PCR). Here, first, we used yeast two-hybrid and tobacco transient expression experiments to prove that the candidate genes *LrAN2-like* and *LrAN1b* can interact and combine with AN11 to form an MBW complex. In addition, the *LrDFR* and *LrANS* gene promoters of kiwifruit were used in the plant *in vivo* screening test to identify R2R3 MYB (*LrAN2-like*) that binds Lr. We characterized the *LrAN1b* gene in Lr and analyzed its overexpression in model plant tobacco. The results show that *LrAN2-like* and *LrAN1b* can interact and act on anthocyanin accumulation. This work not only interprets the metabolomic flux change of flavonoids in *L. ruthenicum* fruits with black or white colors but also further refines the *L. ruthenicum* anthocyanin regulatory network.

MATERIALS AND METHODS

Plant Material and Sampling

Two *L. ruthenicum* varieties were collected at five different developmental stages (S1, S2, S3, S4, and S5) from the *Lycium* resource nursery of the Ningxia Academy of Agricultural and Forestry Sciences in Xixia district, the Ningxia Hui Autonomous Region, China. S1–S5 are shown in **Figure 1A**, and the method of time selection referred to the articles of Zeng et al. (2014) S1 refers to the small and green fruits 3 days before breaker; S2 refers to the

color breaker stage of fruits; S3 refers to period of complete discoloration (3 days after breaker), S4 refers to the early expansion and deepening of pigmentation (6 days after breaker), and S5 refers to the period of complete maturity (10 days after breaker). We collected fruits from three *Lycium ruthenicum* Murray trees separately, and each tree collected about 30 fruits from July to August, 2018. The fruits on each tree are treated as an independent biological repeat, so we have three independent biological repeats for each kind fruit. Similarly, for the metabolome samples, we collected about 50 mature fruits (S5 period) from the three trees for analysis respectively. The fruit of each tree was treated as an independent biological repeat, for a total of three biological repeats. Thus, all data were obtained based on three independent biological replicates. All fruits were snap-frozen in liquid nitrogen, and then kept at -80°C for subsequent metabolite extraction, transcriptome sequencing, and real-time PCR analysis.

Metabolite Extraction

Samples flavonoid content were performed according to previously methods (Chen et al., 2013). The freeze-dried fruit was crushed using a mixer mill (MM 400, Retsch) with a zirconia bead for 1.5 min at 30 Hz. 100 mg powder was weighted and extracted overnight at 4°C with 1.0 ml 70% aqueous methanol. Following centrifugation at 10,000 g for 10 min, the extracts were absorbed (CNWBOND Carbon-GCB SPE Cartridge, 250mg, 3 ml; ANPEL, Shanghai, China, www.anpel.com.cn/cnw) and filtrated (SCAA-104, 0.22μm pore size; ANPEL, Shanghai, China, http://www.anpel.com.cn/) before LC-MS analysis. Three biological replicates and three technical assays for each biological repetition were analyzed.

HPLC Conditions and ESI-Q TRAP-MS/MS

The sample extracts were analyzed using an LC-ESI-MS/MS system (HPLC, Shim-pack UFLC SHIMADZU CBM30A system, www.shimadzu.com.cn/; MS, Applied Biosystems 6500 Q TRAP, www.appliedbiosystems.com.cn/). The analytical conditions of HPLC and MS were referred to previous method (Chen et al., 2013). HPLC: column, Waters ACQUITY UPLC HSS T3 C18 (1.8 μm, 2.1 mm*100 mm); solvent system, water (0.04% acetic acid): acetonitrile (0.04% acetic acid); gradient program,100:0V/V at 0min, 5:95V/V at 11.0min, 5:95V/V at 12.0min, 95:5V/V at 12.1min, 95:5V/V at 15.0 min; flow rate, 0.40 ml/min; temperature, 40°C; injection volume: 2 μl. The effluent was alternatively connected to an ESI-triple quadrupole-linear ion trap (Q TRAP)-MS.

Linear ion trap (LIT) and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer (Q TRAP), API 6500 QTRAP LC/MS/MS System, equipped with an ESI Turbo Ion-Spray interface,

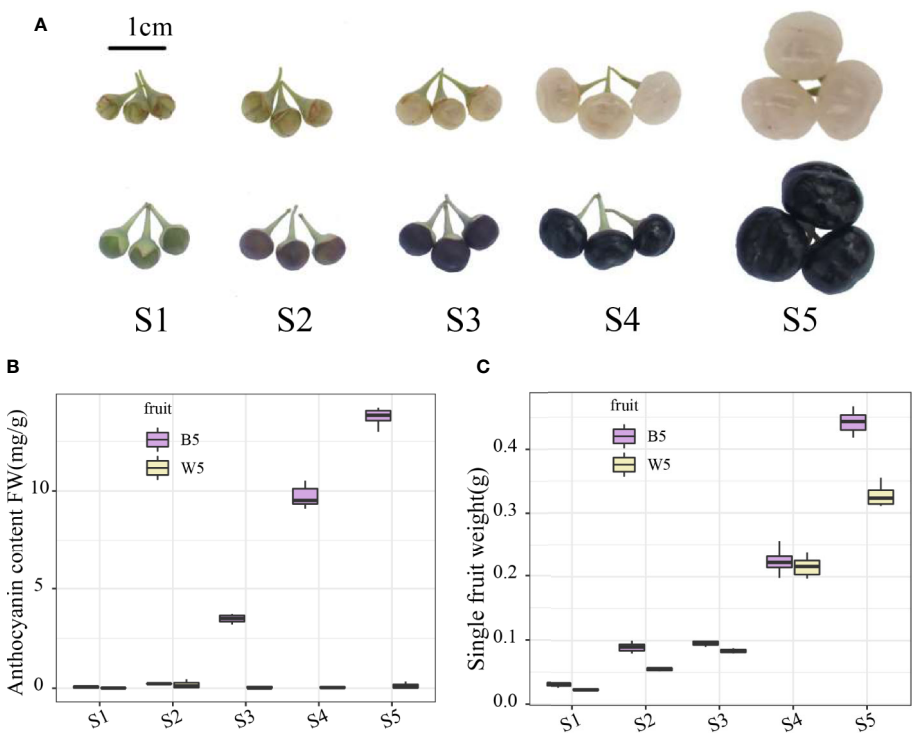

FIGURE 1 | Fruit colors and phenotypic characterization of different developmental stages between two varieties. **(A)** Fruit colors of the two varieties at S1 (3 days before breaker); S2 (breaker); S3 (3 days after breaker); and S4 (6 days after breaker); and S5 (10 days after breaker). **(B)** Anthocyanin content in S1–S5. **(C)** The average fruit weight in S1–S5. Three independent biological repeats were used for the determination of anthocyanin. Fruit weight is the average weight of every ten fruits, with three biological repeats.

operating in a positive ion mode and controlled by Analyst 1.6 software (AB Sciex). The ESI source operation parameters were as follows: ion source, turbo spray; source temperature 500°C; ion spray voltage (IS) 5500 V; ion source gas I (GSI), gas II (GSII), curtain gas (CUR) were set at 55, 60, and 25.0 psi, respectively; the collision gas(CAD) was high. Instrument tuning and mass calibration were performed with 10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as MRM experiments with collision gas (nitrogen) set to 5 psi. Declustering potential (DP) and collision energy (CE) for individual MRM transitions was done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period.

Flavonoids Identification and Quantification

Qualitative analysis of primary spectrum and secondary spectrum data of mass spectrometry detection, based on self-built database MWDB (metware database) and public database of metabolite information. The isotope signal and the K+ ion, Na+ ion, and NH4+ ion repetitive signal removed when some substances were quantified. The metabolite structure analysis refers to MassBank (http://www.massbank.jp/), KNAPSAcK (http://kanaya.naist.jp/KNApSAcK/), HMDB (http://www.hmdb.ca/) (Wishart et al., 2013), MoTo DB (http://www.ab.wur.nl/moto/) and METLIN (http://metlin.scripps.edu/index.php) (Zhu et al., 2013) mass spectrometry Public database.

Metabolite quantification is accomplished using multiple reaction monitoring mode (MRM) analysis using triple quadrupole mass spectrometry. The quadrupole first screens the precursor ions of the target substance and excludes ions corresponding to other molecular weight substances to preclude interference. The precursor ions are fragmented after induced ionization in the collision chamber to form a lot of fragment ions. The fragment ions are then filtered through a triple quadrupole to select a desired fragment ion to eliminate interference from non-target ions. After obtaining the metabolite mass spectrometry data of different samples, integrate the peak area of the mass spectrum peaks of all substances, and integrate and correct the peaks of the same metabolite in different samples (Fraga et al., 2010).

RNA-Seq Analysis

The total RNA was extracted from frozen fruits, the mRNA library of each sample was constructed and sequenced in the Illumina HiSeq platform, and paired-end reads were generated. There are three biological repeats in each sample. Raw data of fastq format were first processed through in-house Perl scripts. Clean data were obtained by removing reads containing adapter, reads containing ploy-N and low-quality reads from raw data. Transcriptome assembly was accomplished based on left.fq and right.fq using Trinity (r20140413p1) (MG et al., 2011). Gene function was annotated based on the following databases: Nr (NCBI non-redundant protein sequences) (v0.2.28) (Nguyen et al., 2007); Nt (NCBI non-redundant nucleotide sequences) (v2.2.28+) (Nguyen et al., 2007); Pfam (Protein family) (HMMER 3) (Finn et al., 2008); KOG/COG (Clusters of Orthologous Groups of proteins) (v0.2.28); Swiss-Prot (A manually annotated and reviewed protein sequence database) (v0.2.28); KO (KEGG Ortholog database) (r140224) (Kanehisa et al., 2008); and GO (Gene Ontology) (b2g4pipe_v2.5) (Götz et al., 2008). The gene expression levels were estimated by RSEM (v1.2.15) (Bo and Dewey, 2011) for each sample. A differential expression analysis of two samples was performed using the DEGseq R package (1.12.0) (Likun et al., 2010). The p-value was adjusted using the q value (Storey and Tibshirani, 2003), qvalue<0.005 & |log2 (fold change) |>1 was set as the threshold for significantly differential expression. The Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the GOseq R package (1.10.0) (MD et al., 2010) based Wallenius non-central hypergeometric distribution. We used KOBAS software (v2.0.12) (Mao et al., 2005) to test the statistical enrichment of differential expression genes in KEGG pathways. The sequences of the DEGs were blasted to the genome of a related species (the protein interaction of which exists in the STRING database: http://string-db.org/) to obtain the predicted PPI of these DEGs. Then, the PPI of these DEGs was visualized in Cytoscape (v3.7.2) (Shannon et al., 2003). The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Wang et al., 2017) in National Genomics Data Center (Zhang Z. et al., 2020), Beijing Institute of Genomics (China National Center for Bioinformation), Chinese Academy of Sciences. The accession number(s) is CRA002484 that are publicly accessible at https://bigd.big.ac.cn/gsa/s/VCuA768g.

Co-expression Network Construction of Metabolome and Transcriptome

A weighted gene co-expression network analysis (WGCNA) was analyzed using the WGCNA R package (v1.68) (Langfelder and Horvath, 2008) based on FPKM values. Network construction and module identification were conducted using the topological overlap measure (TOM). The calculation parameters "soft thresholding power" =14, "minModuleSize" = 30 and "mergeCutHeight" = 0.25 were selected for analysis of the transcriptome data sets. The modules were used to calculate the relationships among modules, anthocyanin content, and representative genes in the 10 samples. The crucial module networks were visualized using Cytoscape with threshold = 0.25.

Sequence Alignment and Phylogenetic Analysis

To investigate the mechanism regulating anthocyanin biosynthesis in *L. ruthenicum*, the deduced amino acid sequences of MYBs: LrAN2, LrAN2-like, MYB113, bHLHs: LrAN1b, and LrJAF13 were obtained from transcriptome data. 40 MYBs and 31 bHLHs of other species were downloaded from NCBI database. The amino acids of the MYB proteins and the conserved domains of bHLH proteins were engaged to perform phylogenetic analysis using MEGA (version 6) with the neighbor-joining statistical method and 1000 bootstrap replicates.

Yeast Two-Hybrid Analysis

According to the Matchmaker® Gold Yeast Two-Hybrid System (Clontech, HTTP://www.clontech.com/), the Full-length cDNAs of the candidate genes *LrAN2-like*, *LrAN1b*, *LrJAF13*, *LrAN11* were respectively cloned in the pGBKT7 and pGADT7, and then co-transformed into the yeast strain AH109. The primers listed in **Supplemental Table 1**. The transformants were selected on SD/-Leu/-Trp medium and tested on SD/-Leu/-Trp/- His/-Ade medium by increasing amounts of 3-amino-1,2,4-triazole (3-AT) and X-α-Gal. Meanwhile, pGADT7 and pGBKT7 Co-transformed with candidate genes as negative control. Growth was scored after 2 d at 30°C.

Yeast One-Hybrid Analysis

The method of Y1H based on previous report (Shim et al., 2013). The *LrAN2-like* CDS recombinded with GAL4 activation domain of pGADT7. A DNA fragment consisting of three copies of the DFR (-804 to -835) and ANS (-1385 to -1415) promoter sequence containing the R2R3-MYB core binding domain sequence was chemically synthesized and homologous recombination into the pHis2 vector. The AD- *LrAN2-like*, were respectively Co-transformated with pHis2-DFR^{MRE}, pHis2-ANS^{MRE}and grown on the SD/- Leu/-Trp medium. Then spot the yeast suspension on the SD/-Leu/-Trp/-His medium, with or without 3-AT (0 or 50 mM). The pGADT7 and four pHis2-DNA fragments respectively Co-transformated for the negative control. The primers listed in **Supplemental Table 1**.

Transient Over-Expression in Tobacco Leaves

The vector pEAQ-HT with 35S: *LrAN2-like* and 35S: *LrAN1b* construct were transformed into Agrobacterium tumefaciens strain GV3101 by heat shock method followed by incubation on plate before infiltration. The cultures of Agrobacterium were resuspended to OD600 reached about 0.75 using buffer (10 mM MES, 10 mM MgCl2, 200 mM acetosyringone, pH 5.6) before infiltration to the leaves. *N. tabacum* plants were grown in glasshouse with 26°C, 16 h light, and 8 h dark conditions. The leaves of 6-week-old N. tabacum with three biological replicates were used for infiltration. The Agrobacterium cultures with empty vector pEAQ-HT acted as a negative control. All primers used for this experiment were listed in **Table S1**. Leaves were harvested 7 days after infiltration to further phenotypic observation, anthocyanin determination and RT-qPCR. All samples were analyzed from at least three biological replicates.

Plant Expression Vectors and Stable Tobacco Transformation

In order to verify the function of *LrAN1b*, the ORF sequence for *LrAN1b* was inserted into pCM1307. The primers were listed in **Table S1**. Thus, *LrAN1b* was promoted by the 35S promoter. The recombinant or empty vector plasmid was introduced into GV3101 by the freeze–thaw method.

Tobacco plants (*N. tabacum*) were transformed with Agrobacterium as described previously (Horsch et al., 1985). The transformed tobacco plants were screened using 25 mg/L^{-1} hygromycin antibiotic, which was used as the plant selective marker. Two transgenic tobacco lines were obtained and the transformation was confirmed using qRT-PCR, which designed OE#1 and OE#2. The seeds of the transgenic plants were harvested separately. The next stage of analysis used T1 generation transgenic plants of overexpressing OE- *LrAN1b* that showed obvious color changes in the florals and capsules.

Measurement of Total Anthocyanin Content

Total anthocyanin content was determined using revised method referring to the previous method (Shin et al., 2007). 0.1g sample was ground into powder by liquid nitrogen and incubated in 600μL of extraction buffer (methanol containing 1% HCl) overnight at 4°C in the dark. After extraction, 400 μL of water and 400 μl of chloroform were added to each sample to remove chlorophyll. Then samples were centrifugated at 14,000 g for 5min at 4°Cto sediment the plant material and absorbances were read at 530 and 657 nm. The quantity of anthocyanin was determined A530 – 0.33×A657, and each sample was extracted and measured in three independent experiments.

QRT-PCR Analysis

High-quality total RNA was used for reverse transcription PCR using the Takara reverse transcription kit (PrimeScript™ RT reagent kit, Dalian, China). The qRT-PCR 20 μL reaction system was constructed using the AceQ qPCR SYBR Green Master Mix (Vazyme, Nanjing, China). All primers used in this study are listed in **Supplementary Table S1**. Real-time assays were performed with SYBR Green Dye using CFX96 Touch™ Real-Time PCR (Bio-Rad, Hercules, CA, USA) detecting platform. The primer amplification effectiveness was analyzed using CFX Manager™ (v3.0) software and the expression of genes was calculated by the $2^{-\Delta\Delta Ct}$ method. The values were expressed as means ± standard error (SE). Three biological replicates and technical duplication were used in RT-qPCR analysis.

Statistical Analysis and Data Availability

All experimental data from three independent biological replicates were subjected to a one-way analysis of variance (ANOVA). Significant differences were calculated by the Student's t-test: tests ($p<0.05$) using SPSS 22.0 Statistics (SPSS Inc., Chicago, IL, USA).

RESULTS

Phenotypic Characterization of Two-Color Fruits of *L. ruthenicum* in Different Developmental Stages

L. ruthenicum fruits of two different colors were investigated to study the phenotypic characterization. For black-fruited *L. ruthenicum*, fruit development was separated into green (S1), turning (S2), purple (S3), black (S4), and fully inflated black (S5)

stages (**Figure 1A**). In contrast, the white-fruited *L. ruthenicum* turned white at the S2 stage and changed from green to transparent-white fruit without pigment accumulation (**Figure 1A**). The further fruit weight results indicated no difference in the size of the two fruits from S1 to S4; however, in the S5 period the black fruit was larger than the white berry by a difference of approximately 0.1425 g (**Figure 1C**). Moreover, the anthocyanin content analysis revealed that the anthocyanin content of black fruits increased sharply in the S3 period, reaching the highest in S5 (full maturity) (13.47 mg g^{-1} of FW). In contrast, almost no anthocyanins were detected in the white fruits, which is compatible with the phenotype (**Figure 1B**). The results show that the black fruit of *L. ruthenicum* may be a good choice as an anthocyanin supplement for consumers, highlighting the prodigious differences in anthocyanin accumulation in black and white fruits of *L. ruthenicum*.

Identification of Flavonoids in the Fruits of Two *L. ruthenicum* Varieties

To compare the differences in flavonoid metabolites in two *L. ruthenicum* varieties, B5 (S5 of black fruit) and W5 (S5 of white fruit) were subjected to LC-ESI-Q- TRAP-MS/MS analysis due to the most abundant metabolite content. The fruits of W5 are purple-black and rich in pigments, while the fruits of B5 are white (**Figure 1A**). In this work, 147 flavonoid metabolites were identified and quantified in two *L. ruthenicum* fruits (**Tables S2** and **S3**). The principal component analysis (PCA) analysis showed that B5 and W5 were separated in the PC1×PC2 score plots with 73.12% of PC1 and 14.69% of PC2 (**Figure S1**). The detected flavonoid species mainly included 8 catechin derivatives, 15 anthocyanins, 41 flavones, 30 flavone *C*-glycosides, 30 flavonols, 19 flavanones, and 6 isoflavones. Furthermore, we set |Log2 (fold change) |≥2, p-value < 0.05, and VIP (variable importance in the project) ≥ 1 as the thresholds for differentially accumulated flavonoids. All differential metabolites are integrated in **Figure S2** and **Table S4**. Furthermore, the box diagrams of the main accumulated components in *L. ruthenicum* are depicted in **Figure 2B**. B5 was twice as high as W5 in the main substances upstream of flavonoid metabolism, naringenin chalcone and naringenin. As is known, naringenin, a key substance in flavonoid biosynthesis, is catalyzed to form downstream of DHK (Dihydrokaempferol), DHM (dihydromyricetin), DHQ (Dihydroquercetin) by a series of enzymes (Tanaka et al., 2010). The results indicate that B5 has more upstream substrates to produce downstream flavonoids compared with W5. Otherwise, DHK had little difference in the component accumulation of the two fruits, and DHM and DHQ of W5 were 3 times higher than B5 (**Figure 2B**). These results suggest that although B5 has more upstream substrate naringenin chalcone and naringenin to produce downstream flavonoids, B5 also has more DHM and DHQ consumed to synthesize downstream anthocyanins compared with W5.

Previous research shows that 37 anthocyanins have been detected from *L. ruthenicum* fruits, including derivatives of peonidin, petunidin, pelargonidin, cyanidin, malvidin, and delphinidin (Wang et al., 2018). In this study, 15 kinds of anthocyanins were identified, including pelargonidin, cyanidin, delphinidin, peonidin, petunidin, malvidin, and rosinidin derivatives. Among these anthocyanins, delphinidin 3-*O*-rutinoside and petunidin 3-*O*-glucoside were the most significant and abundant metabolites; they accumulated in B5 but not in W5. The secondary mass spectrum of delphinidin 3-*O*-rutinoside is shown in **Figure S3A**. This finding is consistent with previous data finding that petunidin derivatives account for 95% of the total anthocyanins in the fresh fruit (Zheng et al., 2011). Moreover, a small amount of rosinidin *O*-hexoside, malvidin 3-*O*-glucoside, malvidin 3-*O*-galactoside, and peonidin *O*-malonylhexoside was only detected in W5. This finding suggests that very trace amounts of anthocyanins are still synthesized in W5. Moreover, although delphinidin was 4 times higher than in B5, W5 did not produce more delphinidin derivatives (**Figure 2B**). It is likely that the delphinidin of W5 lacked modified enzymes to synthesize downstream metabolites.

Similar to anthocyanins, other flavonoids also have extraordinary antioxidant, antitumor, antibacterial, and anticancer activities (Falcone Ferreyra et al., 2012). In our study, 23 flavonoid metabolites were found only in black fruits and not detected in white fruits. These flavonoid metabolites include 3 ', 4', 5'-dihydrotricetin *O*-hexosyl-*O*-hexoside, chrysoeriol *O*-hexosyl-*O*-rutinoside, tricin *O*-rhamnoside, acacetin *O*-acetyl hexoside, chrysin, and pinocembrin in lignin metabolism, 5 myricetin derivatives, 3 catechin derivatives, 3 luteolin glycoside, 4 apigenin glycosides, 2 hesperidin glycosides, and two isoflavones of hydroxygenistein and sissotrin. On the other hand, sakuranetin, kumatakenin, 6-*C*-hexosyl-apigenin *O*-hexosyl-*O*-hexoside, isosakuranetin, daidzin, and glycitin only accumulated in W5 and were not detected in B5. The results show that B5 and W5 only accumulated small amounts of isoflavones and proanthocyanidins, and there was not much difference between the two fruits. Furthermore, the decrease in flavonoids species in W5 may be to reduce the number of branches and synthesize more other flavonoids. Although there are fewer flavonol species in W5, the resulting peak area integration showed that the four most common and representative flavonols in *L. ruthenicum* fruits were kaempferol 3-*O*-rutinoside, kaempferol 3-*O*-robinobioside, quercetin 4'-*O*-glucoside and luteolin 7-*O*-glucoside. The secondary mass spectra of the main four flavonoids in white fruits are shown in **Figures S3B, C,** and **D**, respectively. These four flavonols were detected at a 2-3-fold higher level in W5 than in B5 (**Figure 2B**). In addition, to determine why certain steps of the anthocyanin biosynthetic pathway in W5 were blocked, the intermediates involved in the metabolic process and their main branches were compared. A schematic diagram of the anthocyanin metabolism process and its core metabolites and enzymes in black and white *L. ruthenicum* fruits is shown in **Figure 2A**. The result indicated that purple compounds and arrows represented an accumulation of delphinidin derivatives and small amounts of cyanidin derivatives in B5. Yellow compounds and arrows indicate that several flavonols were accumulated in large amounts in W5. However, despite a sharp decrease in anthocyanins, W5 still contained all other core metabolites detected in B5 (**Figure 2A**). Notably, the concentrations of kaempferol, quercetin and luteolin

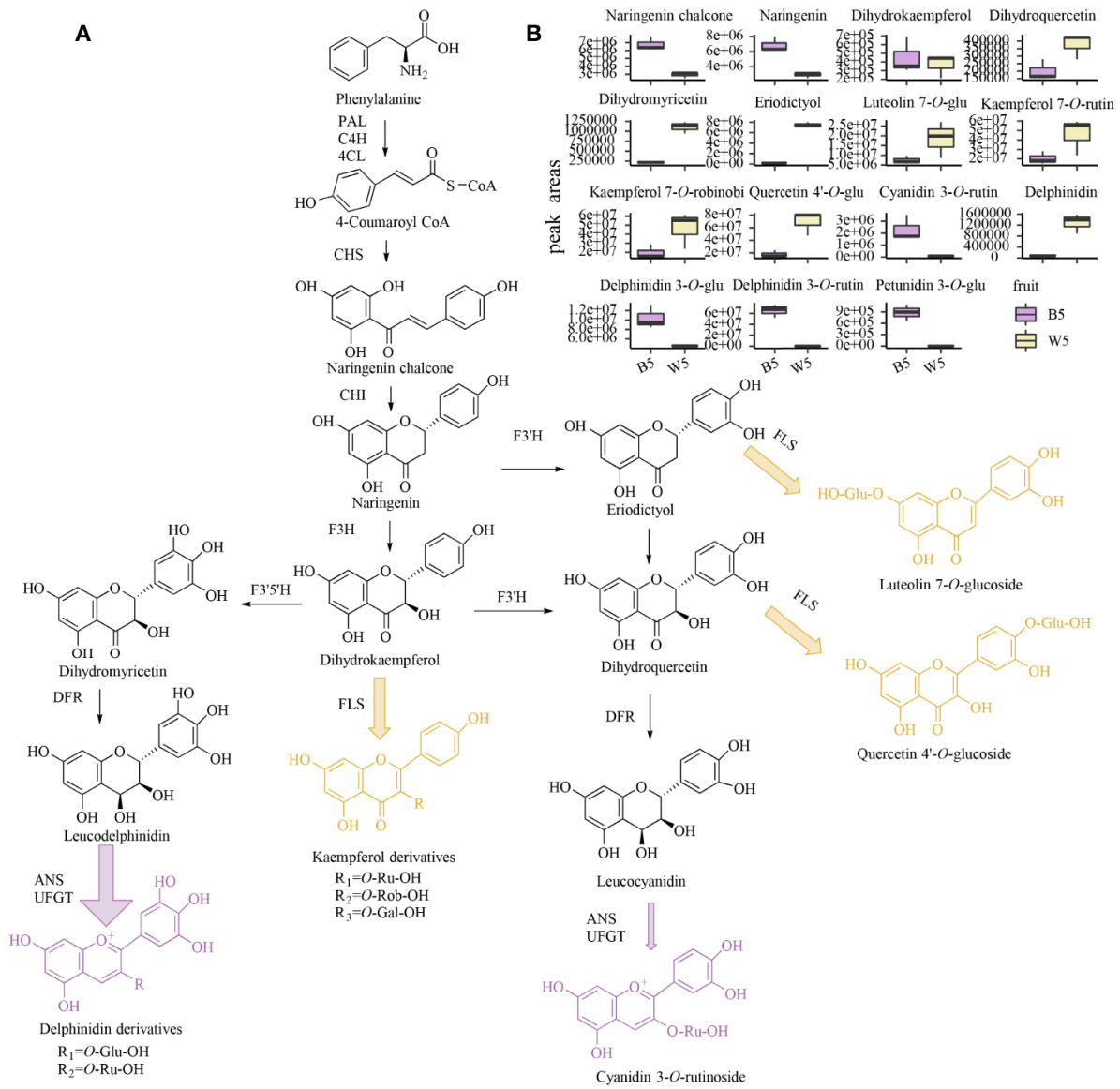

FIGURE 2 | Flavonoid metabolome of two fruits at S5. **(A)** a diagram of the putative anthocyanin metabolic process in black or white fruits; purple represents the main flavonoid component of black fruit, and yellow represents the main flavonoid substance that metabolizes white berries. **(B)** Box plots of the relative content of 14 major flavonoids in two fruits; center lines denote medians, bounds of boxes denote first and third quartiles. HPLC analyses were performed on three biological replicates.

derivatives detected in W5 were significantly higher than in B5 (**Figure 2B**). These results suggest that both W5 and B5 possess the same flavonoid upstream pathway, but the W5 anthocyanin pathway is blocked downstream of naringenin. Then, DHK, DHQ, DHM, and eriodictyol are catalyzed by FLS to synthesize more flavonols instead of continuing to synthesize anthocyanins. Despite no anthocyanin pigmentation in W5, flavonols preferably accumulate in W5. These findings are very similar to previous research on *grape hyacinth*; when anthocyanin synthesis is restricted in white flowers, upstream flux must flow into other branches of the flavonoid pathway (Lou et al., 2014).

RNA-Seq and Assembly

To understand the molecular basis of fruit color polymorphism in *L. ruthenicum*, three biological replicates from each of five developmental stages of black fruits and white fruits of *L. ruthenicum* were used to construct cDNA libraries for high-throughput sequencing, including B1, B2, B3, B4, B5, W1, W2, W3, W4, and W5. After cleaning and quality assessment of the libraries, each library generated 47,356,860–73,227,696 clean reads of sequencing error rates <0.01% (**Table S5**). These reads were assembled into 274,634 unigenes with a mean size of 1,190 bp and an N50 of 1,860 bp (**Figure S4A**). The number of the

unigene length distributions ranging from 500-2,000 was 140,901 (**Figure S4B**), indicating that the proportion of full-length gene assembly in this study was sufficiently high for our further analysis. Because no genomic information is available for *L. ruthenicum*, to obtain comprehensive gene function information, all unigenes were blasted to seven databases including Nr, Nt, Pfam, KOG/COG, Swiss-prot, KEGG, and GO. Finally, 188,308 unigenes were annotated, accounting for 68.56% of all unigenes (**Table S6**). According to the results of the Nr library comparison annotation, the species distribution chart showed that the unigenes had the highest homology to genes of *Solanum tuberosum* (28.2%), *Nicotiana sylvestris* (19.5%), *Nicotiana tomentosiformis* (19.3%) and *Solanum Lycopersicum* (14.3%), implying that the function of the gene of *L. ruthenicum* may be very similar to that of the main model plants of Solanaceae (**Figure S5**). A total of 31,302 DEGs were identified in Transcriptome with padj < 0.05 and log2|fold change| > 1, which is shown in the Venn diagram (**Figure S6A**). Compared with white fruits, upregulated differential genes in black fruits are described by heat maps (**Figure S6B**). The DEGs were mainly divided into three clusters: W1-W5 were divided into cluster1, B1 and B2 were divided into cluster2, and B3-B5 were divided into cluster3. This result indicates that, similar to the anthocyanin phenotype of fruits, the upregulated differentially expressed genes in black fruits are mainly in the B3-B5 period, whereas the early B1-B2 period has less of a difference.

DEGs Involved in Flavonoid Metabolism Between Two Fruit Varieties

To explore the molecular mechanisms leading to the differential fruit coloration in the two varieties, enzyme-encoding genes of the flavonoid pathways were investigated (**Figure 3** and **Table S7**). The analysis of transcriptome data revealed that 72 key candidates exerted direct influence over fourteen enzymes that were known to be involved in flavonoid and anthocyanin biosynthesis in our study (**Table S7**). The expression level of some phenylpropane biosynthetic genes, *LrPAL*, *LrC4L*, and *Lr4CH*, showed no differences in two fruit varieties (**Figure 3**). The result indicates no difference between the two fruits in the early phenylpropane pathway. Moreover, both the transcriptome analysis and qRT-PCR showed that most of the anthocyanin biosynthetic pathway genes, *LrCHS*, *LrCHI*, *LrF3H*, *LrF3'5'H*, *LrDFR*, and *LrANS*, were significantly downregulated in white fruits (**Figures 3** and **5A**). The three enzymes of *LrCHS*, *LrCHI*, and *LrF3H* are important and

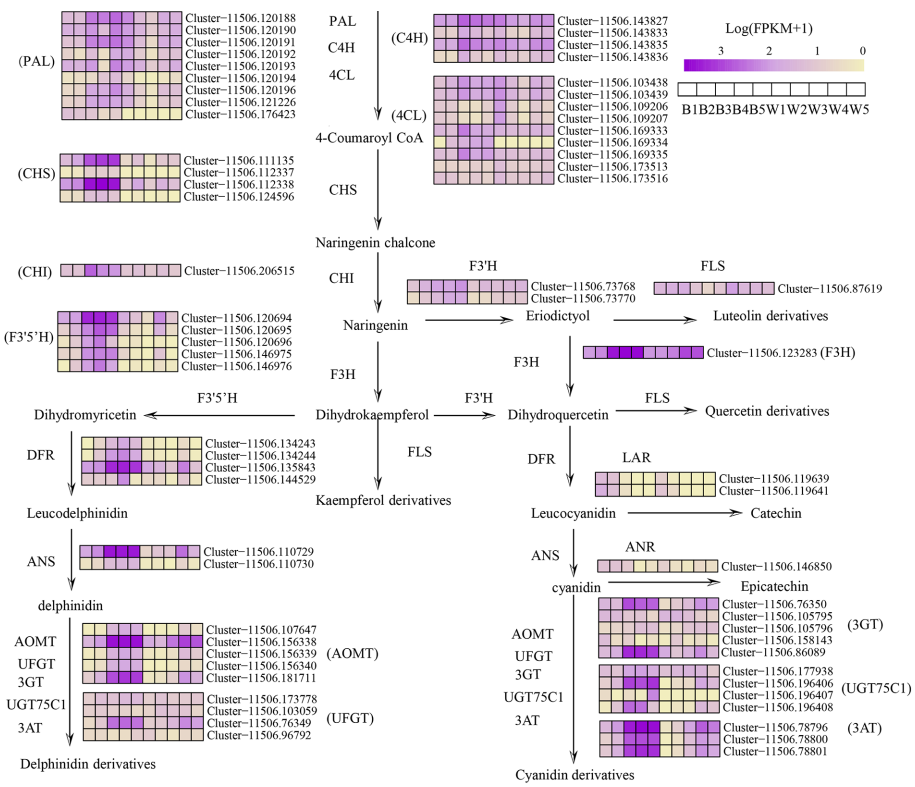

FIGURE 3 | Simplified scheme and heat map of anthocyanin biosynthesis genes in fruit color development of L. ruthenicum. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase;4CL,4-coumarate-CoAligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H flavonoid 3-hydroxylase; F3'H, flavonoid 3′-hydroxylase; F3'5'H, flavonoid 3′5′-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; FLS, flavonol synthase; UFGT, UDP-glucose flavonoid 3-O-glucosyltransferase; 3GT, anthocyanidin 3-O-glucosyltransferase; UGT75C1, anthocyanidin 3-O-glucoside 5-O-glucosyltransferase; 3AT, anthocyanin acyltransferase; AOMT, Anthocyanin methyltransferase; LAR, leucoanthocyanidin reductase; ANR, anthocyanidin reductase. The color scale represents the log-transformed FPKM+1 value. Purple indicates high expression, and yellow indicates low expression.

are considered to be the restriction enzymes to the synthesis of flavonoids (Zhang et al., 2014). These results explain that the difference between black and white fruits starts from the flavonoid pathway, which is also consistent with the differences between naringenin and naringenin chalcone. On the other hand, compared with black fruits, there was no significant difference in the gene expression of transcripts *LrF3'H* and *LrFLS* in white fruits. However, most plants hydroxylate the B-ring using either *LrF3'H* to synthesize red cyanidin or *LrF3'5H* to form a purple delphinium (Zhang et al., 2014). This shows that more delphinium is produced in the black fruit than cyanidin. Previous studies have shown that *LrDFR* and *LrFLS* share the same upstream substrates (e.g., DHM, DHQ, and DHK) and take a large proportion of total flavonoids in plants (Huang et al., 2019). In black fruits, the rate-limiting enzyme *LrDFR* specifically binds DHQ to synthesize delphinidin and its derivatives and has little binding to other substrates. This is accordant with previous studies, which means that DFR can selectively catalyze substrates, which correlates with the amino acid 134 being related to the 134th amino acid (Liu et al., 2019). However, in white fruits extremely low expression of *LrDFR* can no longer combine with the substrate DHQ to synthesize anthocyanins. Therefore, more substrates of DHK, DHQ, and eriodictyol were separately synthesized flavonols by *LrFLS* in the white fruits. Our study shows that the downregulation of *LrDFR* in white fruits results in *LrFLS* synthesizing more flavonols than anthocyanins, which is similar to the findings for *grape hyacinth* (Lou et al., 2014).

The structure of anthocyanins is very unstable, and anthocyanins are generally glycosylated, methylated and acylated before transport to vacuole storage. In this study, the transcripts of anthocyanin glycosyltransferases (*LrUFGT*, *Lr3GT*, and *LrUGT75C1*), anthocyanin methyltransferase (*LrAOMT*) and anthocyanin acyltransferase (*Lr3AT*) were identified, and they were highly expressed in black fruits and only slightly expressed in white fruits (**Figures 3** and **4D**). As explained in a previous study, those glycosyl groups are usually linked to the carbon 3 position and sometimes the carbon 5 and carbon 7 position of the anthocyanin to increase the stability of the pigment (Zhang et al., 2014). *LrUFGT* and *Lr3GT* are mainly involved in the glycosylation of carbon 3, and *LrUGT75C1* is related to the glycosylation of carbon 3 and carbon 5. The result suggests that the glycosylated anthocyanin of carbon 3 is necessary for the stable storage of delphinidin in *L. ruthenicum* fruits. Furthermore, previous research has demonstrated that most of the methylation occurs on the *C3*' and *C5*' hydroxyl groups of the anthocyanin molecules and sometimes on the *C5* or *C7* hydroxyl groups to reduce the chemical activity of the whole molecule and increase its water solubility (Mol et al., 1998). In the fruit of *L. ruthenicum*, *LrAOMT* methylated the delphinidin on its 3' hydroxyl group to form petunidin or both its 3' and 5' hydroxyl groups to form malvidin. Previous experiments have shown that petunidin derivatives accounted for 95% of the total anthocyanins in *L. ruthenicum* fruit, and most of the anthocyanins were acylated by coumaric acid (Zheng et al., 2011). In our study, we identified three transcripts of *Lr3AT* similar to *Gentiana trifloral* belonging to the BAHD acyltransferase family to catalyze anthocyanin acylation (Fujiwara et al., 1998). These results prove that a variety of anthocyanin modifications are found in black fruits but not in white fruits.

Identification of WGCNA Modules Related to Anthocyanin Biosynthesis

WGCNA is now commonly used to analyze entire transcriptome, proteomics, and metabolomics data to identify feature-related co-expression modules (Langfelder and Horvath, 2008). To investigate the gene regulatory network of anthocyanin synthesis in *L. ruthenicum* fruits, a weighted gene co-expression network analysis (WGCNA) was conducted using 23,642 filtered DEGs (**Table S8**). These DEGs identified 35 distinct co-expression modules corresponding to clusters of correlated transcripts (**Figure 4B** and **Table S9**). The total anthocyanin contents of the two kinds of fruits and significant structural genes in transcriptome including *LrCHS*, *LrF3'5'H*, *LrDFR*, *LrANS*, *LrUFGT*, and *Lr3GT* acted as the trait data for a module-trait relationship analysis. The MEdeeppink2 module (4447 genes) presented the highest correlation with anthocyanin ($r = 0.98$, $p = 4e\text{-}07$) and all structural genes ($r<0.05$) in the 35 modules (**Figure 4A**). Furthermore, the expression profile of the deeppink2 module indicated that all DEGs were up-regulated with anthocyanin accumulation in black fruits, but almost no expression was detected in white fruits (**Figure 4C**). Based on the analysis above, the deeppink2 modules were selected for further study. The Gene Ontology (GO) enrichment analysis of the "deeppink2" module showed significantly enriched terms associated with binding, protein binding, and catalytic activity (**Figure S7A**). The KEGG pathway enrichment analysis of the "deeppink2" module showed that the first five significant metabolic enrichment pathways were regulation of autophagy, peroxisome, flavonoid biosynthesis, flavone and flavonol biosynthesis, and endocytosis (**Figure S7B**). Cytoscape representation of the genes with WGCNA edge weight >0.25 indicated that these genes were highly positively connected in the "deeppink2" module. In the interaction network diagram, the outer layer consists of 24 hub genes, including phenylpropanoid-biosynthesis genes *PAL*, *BGLU11*; flavonoid synthesis genes *CHS*, *F3H*, *F3'H*, and *F3'5'H*; and anthocyanin synthesis genes *DFR*, *ANS*, *AOMT*, *3GT*, and *3AT* (**Figure 4D**). In the middle of the network diagram, 10 anthocyanin transporter hub genes showed the highest node connectivity with the 24 structural genes (**Figure 4D**). The transcripts of 7 hub genes of transcription factors were identified in the center of the network diagram including *LrMYB3*(Cluster-11506.171206 and Cluster-11506.161621), *LrETC1*(Cluster-11506.49726), *LrMYBX* (Cluster-11506.134566), *LrAN1b* (Cluster-11506.84096 and Cluster-11506.84101), *LrbHLH 128* (Cluster-11506.180271), *LrTTG2* (Cluster-11506.82488), and *LrNAC78* (Cluster-11506.127677 and Cluster-11506.127679), which are orthologous genes in Arabidopsis and petunia (**Figure 4D**). All Hub genes are listed in **Table S10**. These results further illustrate that the deeppink2 module is very relevant to anthocyanin synthesis.

Confirmation of the Transcriptome Data Using qRT-PCR

To visualize the reliability of the RNA-seq data, 19 genes involved in phenylpropane, flavonoid and anthocyanin synthesis pathways were used for qRT-PCR, and the results were consistent with transcriptome data (**Figure 5A**). At the same time, we screened

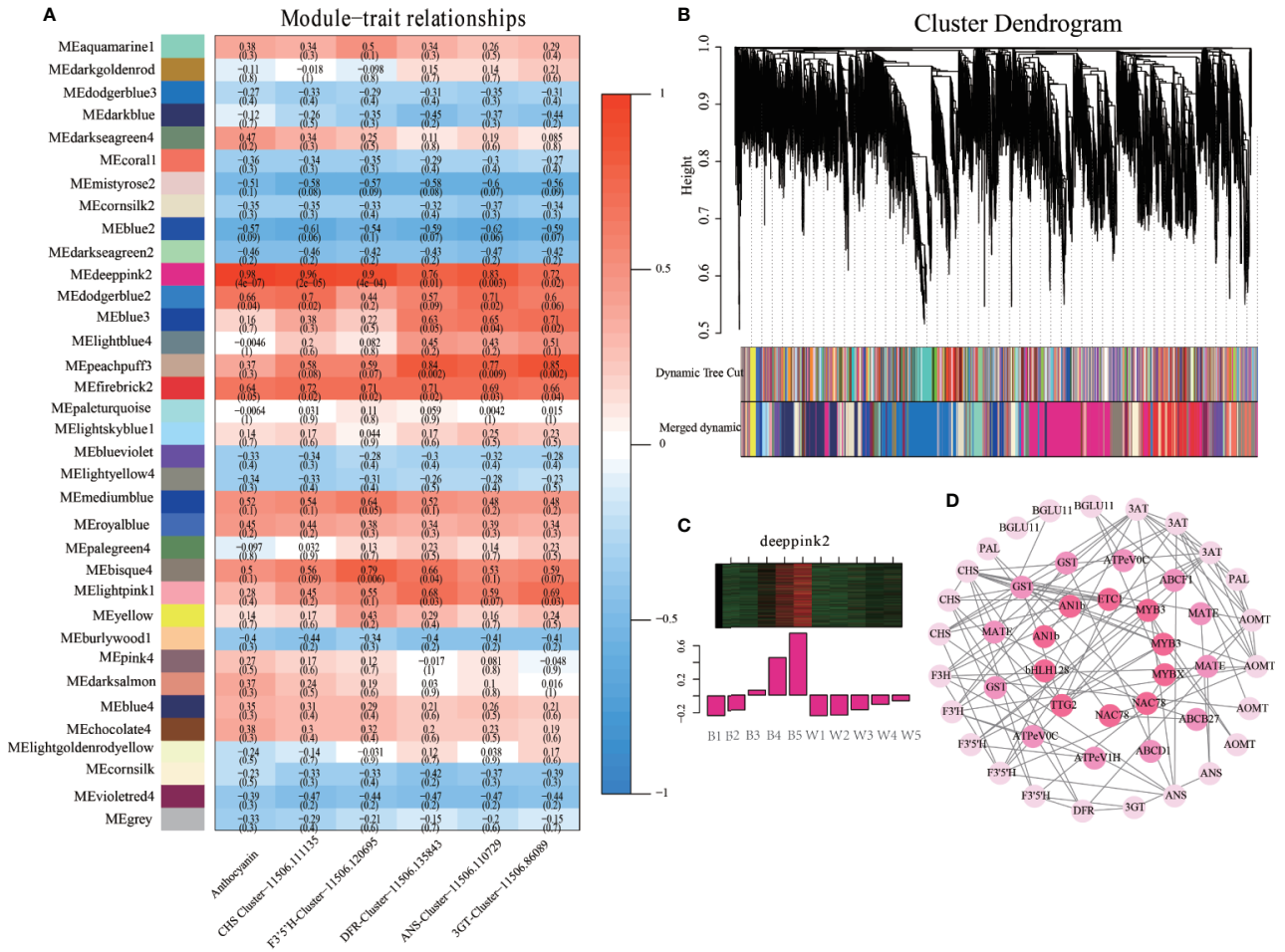

FIGURE 4 | Weighted gene co-expression network analysis (WGCNA) of differentially expressed genes (DEGs) identified black and white fruits. **(A)** Module–trait correlations and the corresponding P-values are in parentheses. The left panel shows 35 modules. The color scale on the right indicates the correlation of module characteristics from -1 (blue) to 1 (red). The panel labeled "Anthocyanins" represents the biosynthetic properties of anthocyanins. The other panels represent changes in gene expression levels. **(B)** Hierarchical clustering tree showing 35 co-expressed gene modules. 23,642 DEGs are clustered into branches, and each module is represented by the main tree branch. The lower panel displays the module in the specified color. **(C)** Eigengene expression profile for the deeppink2 in different samples. The heat map above shows the expression profiles of all co-expressed genes in the module. The bar graph below shows the common expression pattern of genes co-expressed in the module. **(D)** Cytoscape representation of co-expressed anthocyanin metabolism-related genes with edge weights ≥0.25 in the "deeppink2" module. The correlation network diagram is divided into three layers, the outermost genes are phenylpropane, flavonoids, and anthocyanin synthesis genes, the second layer is anthocyanin transporter genes, and the most central layer is transcription factors.

important 4 Hub genes *LrMYB3*, *LrETC1*, *LrAN1b*, *LrTTG2*, 5 anthocyanin related transcription factors *LrAN2*, *LrMYB113*, *LrJAF13*, *LrAN11* identified from previous study, A newly identified MYBs transcription factor is designed as *LrAN2-like* (Cluster-11506.45226) through transcriptome data related to anthocyanin metabolism were selected to analyze their expression using qRT-PCR (**Figure 5B**). All genes used for qRT-PCR was listed in **Table S11**. The results showed that the expression level of Hub genes gradually increased in black fruits, and then the expression of these genes during the development of white fruits was extremely low. In particular, *LrETC1* and *LrAN1b* were hardly expressed. In addition, the expression levels of *LrJAF13* and *LrAN11* in black fruits were significantly higher than those in white fruits in the later stages of fruit development. In contrast, the expression levels of *LrAN2* and *LrAN2-like* are significantly higher in white fruits than in black fruits, and *LrMYB113* is not different between the two fruits. Among them, *LrAN2-like* has the highest expression level among the three genes. Overall, the qRT-PCR results of all candidate genes are very consistent with the corresponding transcriptional abundance of RNA-Seq (**Figures 3** and **5B**), verifying that the RNA-Seq results expressed by these candidate genes are reliable.

Sequence Alignment and Phylogenetic Analysis

Phylogenetic analysis was employed to analysis the anthocyanin-related R2R3-MYB proteins LrAN2 (Cluster-11506.201474), LrAN2-like (Cluster-11506.45226), LrMYB113(Cluster-

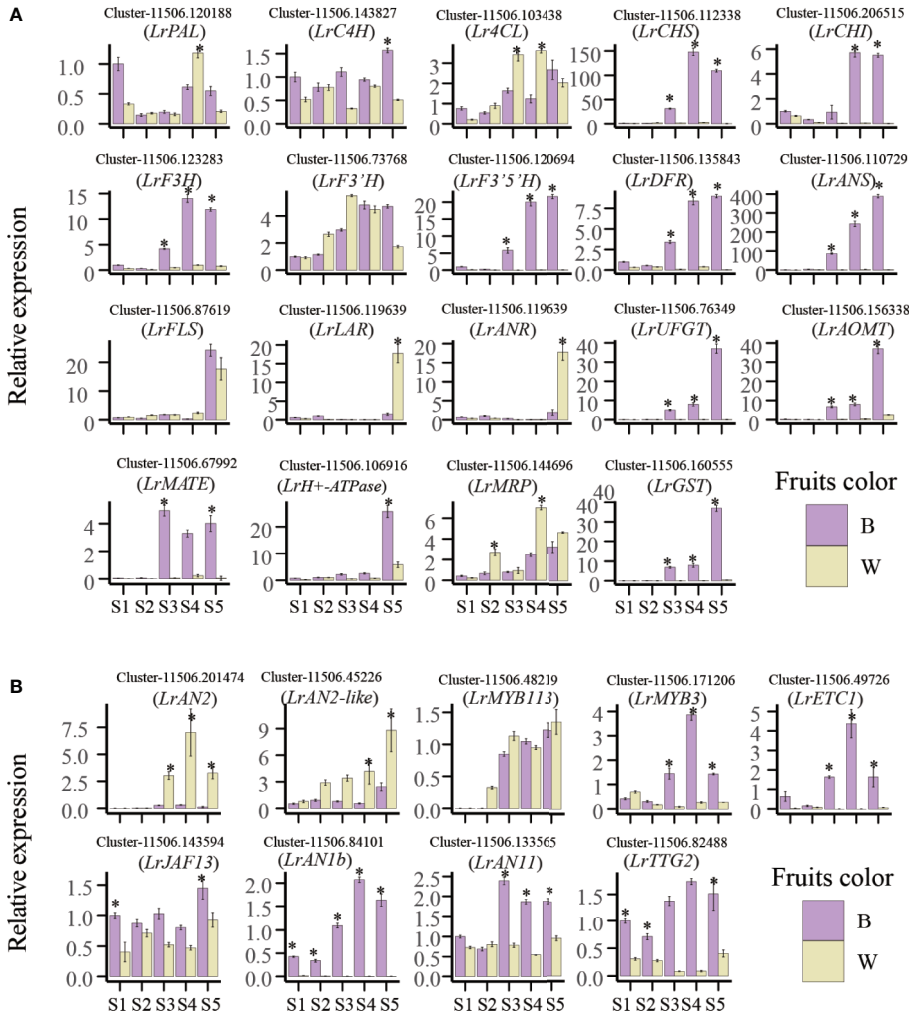

FIGURE 5 | Quantitative real-time polymerase chain reaction (qRT-PCR) verify anthocyanin-related genes. **(A)** Anthocyanin biosynthesis and **(B)** anthocyanin-regulated transcription factor expression at different developmental stages of black and white fruits (S1–S5). Each value is the average of three replicates, and error bars represent SEM (standard error of the mean). Statistical significance was determined by the Student's t-test: *$p < 0.05$.

11506.48219) and bHLH proteins LrAN1b (Cluster-11506.84101) and LrJAF13 (Cluster-11506.143594) in *L. ruthenicum*. At the same time, homologous proteins of other different plants were downloaded to analyze the phylogenetic relationship with these genes (**Figures 6A, B**). Three proteins were identified cluster with anthocyanin regulators PhAN2 and PhPHZ (*Petunia* x *hybrida*), AtPAP1(*Arabidopsis thaliana*) (Borevitz et al., 2000; Albert et al., 2011). Meanwhile, we found that LrAN1b and LrJAF13 were clustered into two clades, namely TT8 and GL3 clades, as previously described (Zhang B. et al., 2019). LrAN1b belongs to the TT8 clade, most homologous with StbHLH (*Solanum tuberosum*) and PhAN1(*Petunia* x *hybrida*), while another LrJAF13 was clustered with NtJAF13a, NtJAF13a (*Nicotiana tabacum*) and PhJAF13 (*Petunia* x *hybrida*) (Quattrocchio et al., 1998). Previous studies show that both branches of bHLH group have the function of regulating anthocyanin, but there is no functional redundancy, and TT8 clade has a more direct role in regulating anthocyanin (Montefiori et al., 2015).

LrAN2-Like Interact With LrAN1b and the WD40 Protein LrAN11

The MBW complex regulates the expression of genes involved in anthocyanin biosynthesis, analysis of the protein fragments shows that R2R3 MYB and WD40 proteins all bind to different regions of the bHLH protein (Payne et al.). Yeast two-hybrid (Y2H) experiments were performed to determine if LrAN1b could bind other TFs involved in anthocyanin regulation (**Figure 7A**). The molecular interactions candidate gene LrAN2-like were performed with the bHLH proteins LrAN1b and LrJAF13, and the WD40 protein LrAN11 (**Figure 7A**). The result showed that LrAN2-like tested interacted with LrAN1b and LrJAF13, but with varying strengths; LrAN2-like shows a higher affinity for

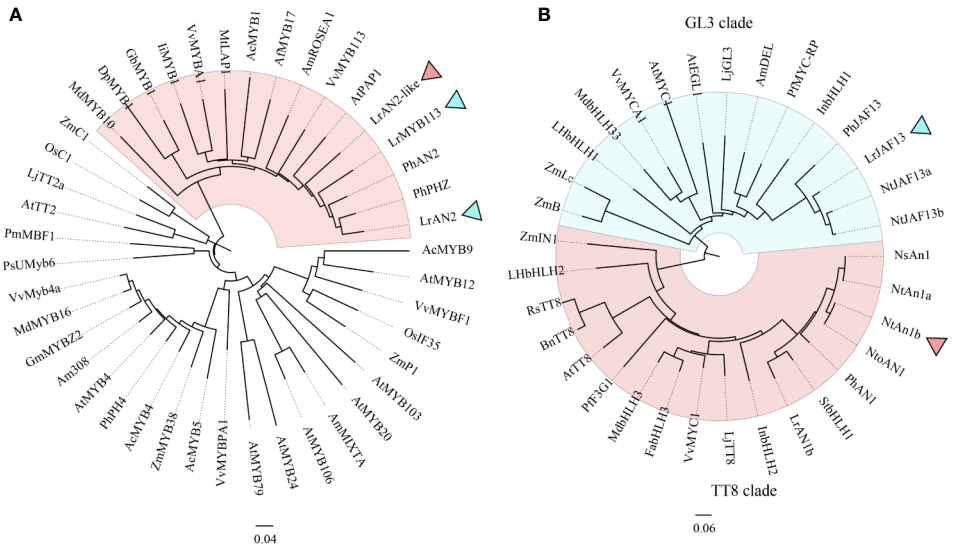

FIGURE 6 | The neighbor-joining phylogenetic tree of plant MYB and bHLH TF sequences. Numbers next to the nodes are bootstrap values from 1,000 replications. The tree is drawn to scale and its branch length is the same as the unit used to infer the evolutionary distance of the phylogenetic tree (scale bar, 0.1 amino acid substitutions per site). The NCBI database retrieves the following deduced amino acid sequences: **(A)** AcMYB1(AQP25672.1), AcMYB4(AQP25673.1), AcMYB5(AQP25674.1), AcMYB9(AQP25675.1) in *Allium cepa*; AmROSEA1(ABB83826.1), AmMIXTA (CAA55725.1) and Am308(P81393.1) in *Antirrhinum majus*; AfMYB17(ACQ82820.1)in *Aquilegia Formosa*; AtMYB4(AAP13410.1), AtMYB12(ABB03913.1), AtMYB20(NP_176797.1), AtMYB24(AAD53092.1), AtPAP1 (AAG42001.1), AtMYB79(AEE83284.1), AtMYB103(AED96722.1), AtMYB106(AEE73615.1) and AtTT2(CAC40021.1) in *Arabidopsis thaliana*; DpMYB1 (BAJ33513.1) in *Dahlia pinnata*; GmMYBZ2(NP_001235092.2) in *Glycine max*, GbMYB1(BAJ17661.1) in *Gynura bicolor*; IiMYB1(BAE94391.1) in *Ipomoea nil*; LjTT2a (BAG12893.1) in Lotus japonicus; MdMYB10(ACQ45201.1) and MdMYB16(ADL36756.1) in *Malus domestica*; MtLAP1(ACN79541.1) in *Medicago truncatula*; OsC1(BAD04024.1), OsIF35(BAB64301.1) in *Oryza sativa*; PhAN2 (AAF66727.1), PhPH4(ADX33331.1),PhPHZ(ADW94951.1) in *Petunia x hybrida*; PsUMyb6 (ACH95792.1) in *Phalaenopsis schilleriana*; PmMBF1(AAA82943.1) in *Picea mariana*; VvMYBA1(BAD18977.1),VvMyb4a(ABL61515.1),VvMYBPA1(CAJ90831.1),VvMYBF1 (ACT88298.1),VvMYB113(RVW78283.1) in *Vitis vinifera*; ZmC1(AAA33482.1, ZmP1(AAC49394.1), ZmMYB38(P20025.1) *Zea mays*; LrAN2(MK125045.1), LrAN2-like (MT749386), and LrMYB113 (MT773444) *Lycium ruthenicum*. **(B)** AmDEL(AAA32663) in *Antirrhinum majus*, BnTT8 (NP_001302903) in *Brassica napus*,FabHLH3 (AFL02463) in *Fragaria* × *ananassa*, (BAE20057) and LhbHLH2 (BAE20058) in *Lilium hybrid*,InbHLH1(BAE94393) and InbHLH2(BAE94394) in *Ipomoea nil*; MdbHLH3 (ADL36597) and MdbHLH33 (ABB84474) in *Malus domestica*, PhAN1 (AAG25928), and PhJAF13 (AAC39455) in *Petunia* ×*hybrida*, VvMYC1 (ACC68685) and VvMYCA1 (ABM92332) in *Vitis vinifera*, ZmB (CAA40544), ZmIN1 (AAB03841), and ZmLc (P13526) in *Zea mays* LjGL3 (AB492284) and LjTT8 (AB490778) in *Lotus japonicus*; NsAN1 (HQ589210) in *Nicotiana sylvestris*; NtAN1a (HQ589208), NtAN1b (HQ589209), NtJAF13a (KF305768), and NtJAF13b (KF298397) in *Nicotiana tabacum*, NtoAN1 (HQ589211) in *Nicotiana tomentosiformis* PfF3G1 (AB103172) and MYC-RP (AB024050) in *Perilla frutescens*, RsTT8 (KY651179) in *Raphanus sativus*;StbHLH1 (ALA13578.1)in *Solanum tuberosum*; AtEGL3 (Q9CAD0), AtGL3 (NP_680372), AtMYC1 (Q8W2F1), and AtTT8 (Q9FT81) in *Arabidopsis thaliana*; LrJAF13(KF768076), LrAN1b(KF768077) in *Lycium ruthenicum*.

LrAN1b than for LrJAF13. It was observed that the bHLH factors LrAN1b and LrJAF13 not only interacted with themselves, but also that these two transcription factors interacted. In addition, BD-LrAN11 has a strong interaction relationship with all MYBs and bHLHs, which prove WD40 is a crucial component of the MBM protein complex. The Y2H assays contained 25 mM of 3-amino-1,2,4-triazole (3-AT), which inhibits the action of the HIS3 reporter gene to reduce the self-activation of transcription factors. Adding chromogenic substrate X-α-Gal to the selection medium (-Leu-TrP-His-Ade) can activate the LacZ reporter gene.

LrAN2-Like Proteins Bind to the Promoters of Anthocyanin Biosynthetic Genes

To verify the speculation that anthocyanin biosynthetic genes might be regulated by *LrAN2-like* in *L. ruthenicum*, a yeast one-hybrid (Y1H) assay was employed to explore their ability to bind the promoters of key structural gene *LrDFR* and *LrANS*. Thirty nucleotides from the *LrDFR* (-804 to -835) and *LrANS* (-1385 to -1415) core binding sequence were repeated three times, which were placed upstream of the *HIS3* selectable marker gene (**Figure 7B**). The MRE binding site-*HIS3* construct and full-length LrAN2-like fused to the GAL4 activation domain were co-transformated in Y187. In the selection medium (-Leu-TrP-His-Ade) with 50mM 3AT, LrAN2-like could activate *LrDFR* and *LrANS*, but the negative controls were not (**Figure 7B**). Above all, we deduced that the anthocyanin biosynthetic genes *LrDFR* and *LrANS* might be candidate target genes of LrAN2-like.

Transient Overexpression of LrAN2-Like and LrAN1b in Tobacco Leaves

To further verify the function of newly identified genes *LrAN2-like* and candidate gene *LrAN1b* in anthocyanin synthesis, we transiently transformed these genes into tobacco leaves. As shown

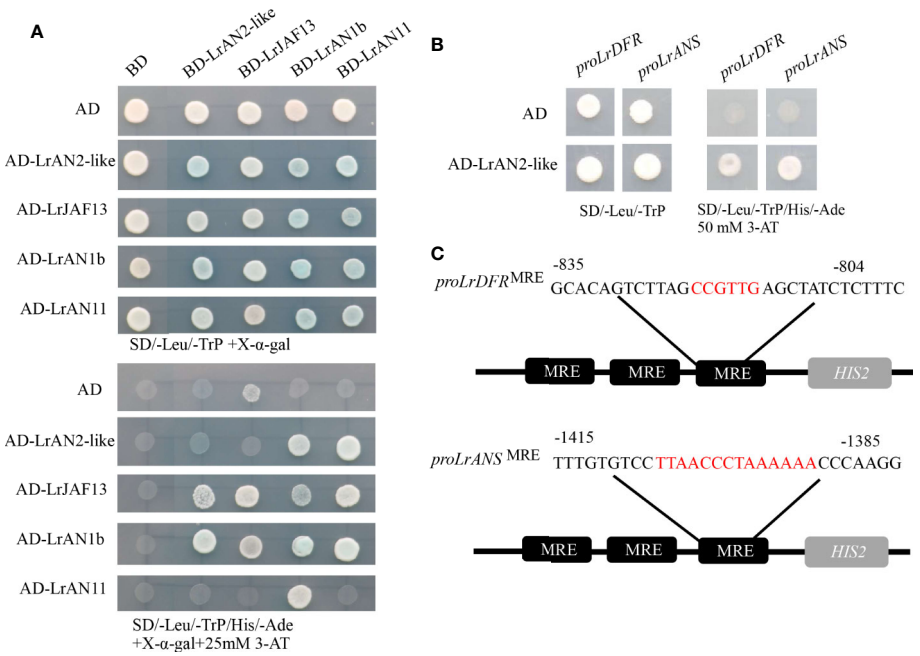

FIGURE 7 | LrAN2-like interacts with other transcription factors and regulates downstream target genes. **(A)** Yeast two-hybrid assay of LrAN2-like and main regulators bHLH protein LrAN1b, LrJAF13 and WD40 protein LrAN11in anthocyanin biosynthesis. The assay shown represents growth on selective media lacking Leu, Ade, Trp, and His, supplemented with 25 mM 3-AT. AD, GAL4 activation domain; BD, GAL4 DNA binding domain. **(B)** Yeast one-hybrid indicates LrAN2-like directly interacts with the MRE in *LrDFR* and *LrANS* promoter. **(C)** The *LrDFR* (-835 to -804) and *LrANS* (-1415 to -1385) promoter sequence containing the MYB-recognizing element (MRE) motif was repeated three times and fused to the *HIS2* reporter gene in Y1H assay.

in **Figure 8A**, anthocyanin was not observed when the empty vector pEAQ-HT and *LrAN1b* alone was transformed in tobacco leaves. However, A large amount of pigment accumulated in the tobacco leaves, when *LrAN2-like* was injected into the leaves. This shows that *LrAN2-like* can activate the anthocyanin pathway of tobacco leaves. Moreover, the pigment was largely enhanced when *LrAN2-like* was co-transformed with *LrAN1b*. The anthocyanin content measurement results show that *LrAN2-like* can cause tobacco leaves to produce a large amount of pigment content, and co-transformation of *LrAN1b* can produce the most pigment (**Figure 8B**). To further investigate the mechanism of *LrAN2-like* and *LrAN1b* in regulating anthocyanin metabolism, we evaluated the gene expression of three key anthocyanins regulate transcription factors *NtAN2*, *NtAn1a*, *NtAn1b* and 8 anthocyanin biosynthesis genes *NtCHS*, *NtCHI*, *NtF3H*, *NtF3'H*, *NtF3'5'H*, *NtDFR*, *NtANS*, *NtUFGT* in tobacco by RT-qPCR analysis (**Figure 8C**). As shown in **Figure 8C**, *LrAN2-like* could stimulate the expression of tobacco transcription factors and structural genes, activating anthocyanin synthesis in tobacco leaves. In addition, when *LrAN2-like* and *LrAN1b* were co-expressed, the related genes relative tobacco anthocyanin had a higher upregulation compared to the expression of *LrAN2-like* alone. QRT-PCR analysis suggested that co-transformation of TFs, *LrAN2-like* and *LrAN1b*, enhanced the anthocyanin synthesis and transport by upregulating the expression of the *NtAN2*, *NtAn1a*, *NtAn1b* to activate tobacco anthocyanin synthesis structural. The results indicated that *LrAN2-like* was essential for regulating anthocyanin biosynthesis while *LrAN1b* could facilitate increased anthocyanin accumulation.

Overexpression of *LrAN1b* Elevated Anthocyanin Accumulation in Florals and Capsules of Transgenic Tobacco

We generated two T1 over-expressing *LrAN1b* transgenic lines, designed OE#1, OE#2 and used for further analysis (**Figure 9A**). Only increased pigmentation was observed in the florals and capsules of transgenic plants. Anthocyanins in florals tissues accumulate in the corolla, filaments, and anthers (**Figures 9A, B, D**). In addition, in the epidermis of capsules, immature seeds will accumulate large amounts of anthocyanins (**Figure 9C**), but not in other plant tissues of transgenic plants observed. The total anthocyanins content of the transgenic *LrAN1b* OE#1 and OE#2 corollas was 3.6 times and 4.2 times higher than that of the control (**Figure 9E**). We used corollas of two *LrAN1b* overexpression lines OE#1, OE#2 and empty vector control to extract mRNA for qRT-PCR analysis. The results showed that *LrAN1b* of the two lines were significantly increased but not expressed in the control (**Figure 9F**). In addition, qRT-PCR analysis was performed on three key transcription factors *NtAN2*, *NtAn1a*, *NtAn1b* and 8 structural genes involved in anthocyanin synthesis in tobacco

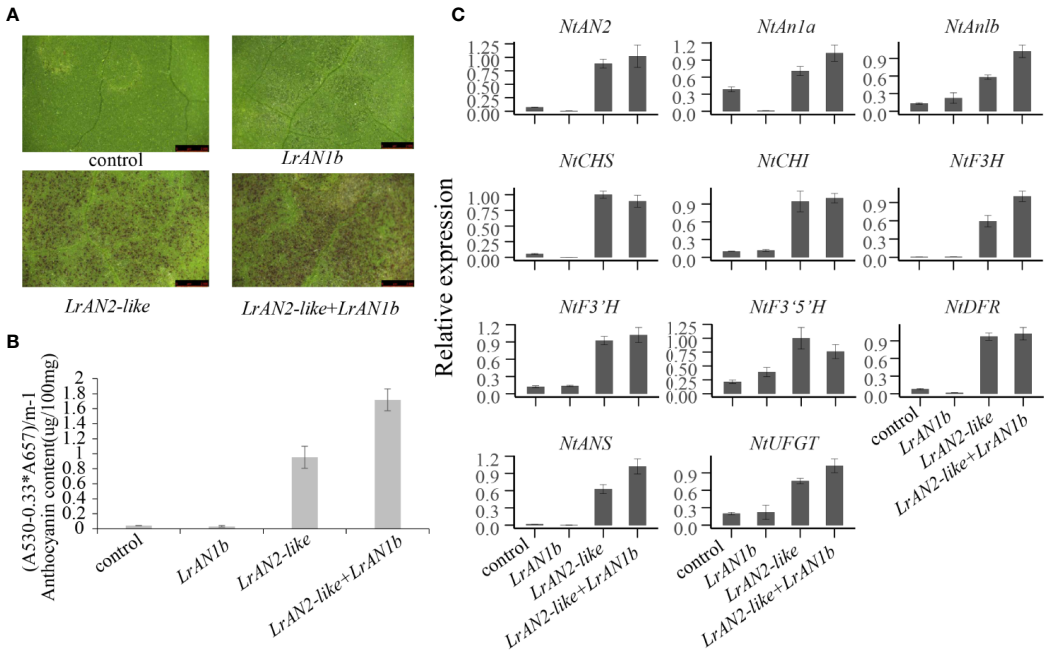

FIGURE 8 | Anthocyanin contents in transiently transformed tobacco leaves infiltrated with Agrobacterium strains carrying *LrAN2-like* and *LrAN1b*. **(A)** Images of transiently transformed tobacco leaves 5 d after agroinfiltration. Four different assays are indicated control: empty vector, *LrAN1b*, *LrAN2-like*, *LrAN2-like*, and *LrAN1b*. **(B)** Anthocyanin contents of four different samples with three biological replicates. **(C)** Relative expression levels of tobacco anthocyanin structural and regulatory genes determined by qPCR analysis. Results represent mean values ± SD from three biological replicates.

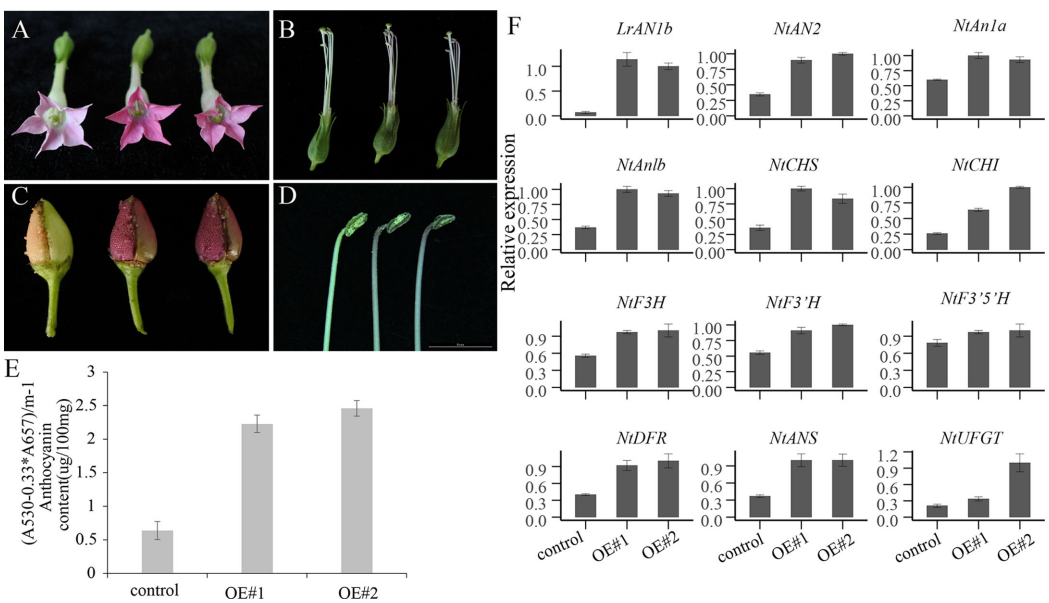

FIGURE 9 | Phenotypic observation and quantitative real-time polymerase chain reaction (qRT-PCR) analyses of the anthocyanin pathway genes in empty vector (control) and two OE-*LrAN1b* line (OE#1, OE#2). the phenotypic images control(left) and OE#1 and OE#2 (right) of **(A)** flower, **(B)** stamen and pistil, **(C)** capsule, **(D)** anther. **(E)** anthocyanin extracted from corollas of the control (left) and OE#1 and OE#2 plants. **(F)** Relative expression of *LrAN1b*, anthocyanin biosynthetic genes, and tobacco transcription factor (TF) genes in the corollas of the control in OE#1 and OE#2 plants. Error bars indicate the standard deviations (SD) of average results. qRT-PCR and anthocyanin determination analyses were performed on three biological replicates.

anthocyanins pathway. As shown in **Figure 9F**, the transcription levels of *NtAN2*, *NtAn1a*, *NtAn1b* and structural genes *NtCHS*, *NtCHI*, *NtF3H*, *NtF3'H*, *NtF3'5'H*, *NtDFR*, *NtANS*, *NtUFGT* were 2.0–3.0 -fold higher than the control, respectively. In summary, *LrAN1b* is a homologous gene of *NtAn1a* and *NtAn1b* (**Figure 6B**), and overexpression of *LrAN1b* can greatly up-regulate the anthocyanin of tobacco florals and capsules.

DISCUSSION

LrAN2-Like Interacts With LrAN1b to Participate in Anthocyanin Regulation

MYB proteins are key factors in regulatory networks controlling secondary metabolism and responses to biotic and abiotic stresses (Dubos et al., 2010). Previous studies identified anthocyanin-activated MYB transcription factor *LrAN2*, *LrMYB113* in *L. ruthenicum* (Zeng et al., 2014). In this study, a new gene MYB transcription factor homologous to *LrAN2* was identified and described as *LrAN2-like*. Recent research indicates that the functional diversity and high expression level of *LrAN2* may be the reason for the high content of anthocyanins in *L. ruthenicum* fruit (Zong et al., 2019). However, the qRT-PCR results indicated that those transcription factors including *LrAN2*, *LrAN2*-like, and *LrMYB113* were still highly expressed in white fruits compared with black fruits (**Figure 5B**). Among them, *LrAN2*-like is the highest expression transcription factor. Sequence alignment and phylogenetic analysis showed that *LrAN2-like* is very homologous to *LrAN2* and *LrMYB113*. Besides, LrAN2-like has extremely high homology to the petunia anthocyanin regulatory gene *PhAN2* (Albert et al., 2011). Y2H indicate that LrAN2-like can interact strongly with LrAN1b and LrAN11, and weakly interact with LrJAF13; indicating that LrAN2-like was able to form MBW complex with bHLH and WD40 proteins. In addition, it was proved that LrAN2-like can directly bind to the MRE of the *LrDFR* and *LrANS* by the Y1H experiment. In order to certify the function of anthocyanin *LrAN2*-like more intuitively, *LrAN2*-like was instantaneously injected into tobacco leaves. It was found that *LrAN2*-like could produce anthocyanins on the leaves in a short time. Several studies indicate that *LrAN2*-like is a new transcription factor regulating anthocyanins different from *LrAN2* and *LrMYB113* in *L. ruthenicum*.

LrAN1b was identified by WGCNA and it was associated with anthocyanin accumulation in two fruits and was not expressed in white fruits. The abnormal expression of *LrAN1b* paralleled the lack of anthocyanin accumulation in the white fruit of *L. ruthenicum*. Zeng et al. Showed that *LrAN1b* may be involved in the regulation of anthocyanins through functional prediction and qRT-PCR in *L. ruthenicum*, but there is no direct proof to prove it (Zeng et al., 2014). Therefore, in this study, Y2H and tobacco injection experiments showed that *LrAN1b* can interact strongly with *LrAN2-like* and *LrAN11*. Co-expression of *LrAN1b* and *LrAN2-like* could significantly increase the content and activation of anthocyanin in tobacco leaves Anthocyanin synthesis structural gene. To further verify the gene function of *LrAN1b*, it was stably transformed into tobacco. It was observed that there is significant accumulation of pigment in florals and capsules. In addition, the anthocyanin-regulated transcription factors *NtAN2*, *NtAnla*, *NtAnlb*, and anthocyanin synthesis genes in the corolla of tobacco were up-regulated about 2 times. This is similar to the phenotype of *NtAnla* and *NtAnlb* overexpression in tobacco (Bai et al., 2011). Phylogenetic analysis showed that LrAN1b and LrJAF13 belong to TT8 clade and GL3 clade transcription factors, respectively. In fact, previous studies have proposed *AN1* is directly involved in the activation of biosynthetic genes, and *JAF13* is involved in the regulation of *AN1* transcription in the anthocyanin pathway of Solanaceae plants (Montefiori et al., 2015). The LrAN1b homologous protein PhAN1 in petunia, which interacts with all anthocyanin-regulated MYBs to regulate the anthocyanin pathway (Albert, 2015). PhAN1 mutations will affect pigment synthesis, vacuole pH and seed coat development in petunia (Spelt et al., 2000; Spelt et al., 2002; Zhang S. et al., 2020). In summary, LrAN1b interacts with LrAN2-like and LrAN11 to form a complex to regulate anthocyanins in *L. ruthenicum*. And abnormal expression of *LrAN1b* in white fruits may cause the loss of anthocyanins. This is also the case in other plants, The bHLH transcription factor *RLL1* gene of lettuce loses its function of activating anthocyanin biosynthesis due to a 5 bp deletion in some varieties. (Su et al., 2020). Furthermore, the genome-wide association analysis showed that part of the non-purple rice leaf phenotype was due to an insertion of a 6.5-kb Copia-like retrotransposon found in the 5'UTR (-49-bp) of bHLH transcription factor *OsRb* (Zheng et al., 2019). Above all, these findings imply that the LrAN1b is indispensable for regulation of the anthocyanin pathway in *L. ruthenicum* fruits.

Identification of Anthocyanin Transport and Other Hub Transcription Factors Involved in Anthocyanin Biosynthesis

Anthocyanins are eventually transported to the vacuole storage after being synthesized. In this study, three transcripts of glutathione S-transferase (GST) (Cluster-11506.122999, Cluster-11506.160555, Cluster-11506.160556) were identified homologous to the *TT19* gene in Arabidopsis. Research declares that Arabidopsis *TT19* encodes a GST protein catalyzing glutathione and covalently binding to anthocyanins to form glutathione S-conjugate transports to the vacuolar membrane (Sun et al., 2012). Interestingly, in strawberry, overexpression of *GST* can cause proanthocyanidin biosynthetic enzymes to synthesize anthocyanins at an early developmental stage, leading to early fruit ripening (Gao et al., 2019). Thus, *GST* is not only involved in anthocyanin transport but also a potential candidate gene for color breeding. In addition, three genes were predicted to encode the ABC (ATP-binding cassette) transporter superfamily including D family member 1 (*ABCD1*, Cluster-11506.146202), B family member 27 (*ABCB27*, Cluster-11506.94554) and F family member 1 (*ABCF1*, Cluster-11506.144849). ABCC1 is thought to

be a transporter located on the vacuolar membrane responsible for transporting cyanidin 3-*O*-glucoside to the vacuole in grapes (Francisco, 2013). Additionally, three multidrug and toxic compound extrusion (*MATE*) (Cluster-11506.55162, Cluster-11506.67992, Cluster-11506.67991) genes were identified in this module, with extremely high expression in black ripe fruits and low expression in white fruits (**Figure 5A**). In grapes, *MATE* can only transport acylated anthocyanins, but not pelargonidin 3-*O*-glucoside or cyanidin 3-*O*-glucoside, meaning that acylation is necessary for MATE transport (Gomez et al., 2009). Furthermore, three V-type H + -ATPase genes (Cluster-11506.124451, Cluster-11506.128173, Cluster-11506.148228) were identified in the deeppink2 module. Previous research has found that the functions and activities of *MATE* largely depend on different types of H^+ - ATPase to provide and maintain the H^+ concentration gradients on both sides of the vacuole membrane (Zhao and Dixon, 2010). Above all, three types of mechanisms for anthocyanin transport are found in *L. ruthenicum* fruits: GST-mediated transport, MRP and MATE-mediated transport and membrane vesicle-mediated transport.

In addition to the members of the MBW complex, other transcription factors also indirectly regulate anthocyanin biosynthesis. In this study, a homologous protein of PH3 was identified in WGCNA designed as *LrTTG2*, which was up-regulated in black fruits according to anthocyanin accumulation but exhibits extremely low expression in white fruits (**Figures 4D** and **5B**). PH3 encodes a WRKY protein. It is not only the target gene of AN11-AN1-PH4 complex, but also can be combined with AN11 alone to regulate the vacuolar acidification of petunia (Verweij et al., 2016). Furthermore, two transcripts of *LrNAC78* were detected in this module, and their expression levels were higher in black fruits than in white fruits (**Figure 4D**). LrNAC78 is homologous to *ANAC078* in *Arabidopsis thaliana*, which regulates flavonoid production under light (Morishita et al., 2009). Furthermore, *LrMYBX* and *LrbHLH 128* are homologous to MYB-like protein X in *Solanum lycopersicum* and *StbHLH128* in *Solanum tuberosum*, respectively, encoding a protein with unknown function.

On the other hand, we identified a putative R2R3-MYB TF *LrMYB3* and another R3 TF MYB transcription factor *LrETC1* in the center of the network. These two genes were predicted to be negative regulators of anthocyanin biosynthesis (**Figure 4D**). The transcription of *LrMYB3* and *LrETC1* was highly correlated with that of *LrAN1b* and with the degree of fruit coloration. The gene function prediction results showed that *LrMYB3* was most homologous to the petunia transcription factor *PhMYB27*. PhMYB27 contains a C-terminal EAR motif to bind MBW complexes and targets the anthocyanin pathway genes, and its expression increases under light induction to inhibit excessive anthocyanin accumulation and prevent damage to plants (Albert et al., 2014). Additionally, *LrETC1* is very similar to Arabidopsis CAPRICE (*CPC*), TRIPTYCHON (*TRY*), and ENHANCER OF TRY and CPC 1 (*ETC1*). Although these R3MYBs do not have DNA-binding domains, a previous study indicated that *CPC* retains the motifs responsible for binding to bHLH proteins to participate in inhibiting anthocyanin synthesis under different stress conditions (HF et al., 2009). What emerges from the results reported here is that *LrMYB3* and *LrETC1* play a repressor role in the regulation of anthocyanin biosynthesis in *L. ruthenicum*. The specific regulatory mechanism of these two genes is under study, and more work is needed to prove it.

Research on Anthocyanin Regulation Network *L. ruthenicum* Fruits

In our research, the black mature fruits mainly accumulated delphinidin derivatives while flavonols were detected in white mature fruits to be 2–3 times higher than in black fruits (**Figure 2**). The metabolic difference of the white fruit phenotype is due to the drastic reduction of anthocyanins, and the metabolic flow shifts to the flavonol products. Meanwhile, a large number of anthocyanin metabolism synthesis genes, modified genes, transport genes and transcription factors were identified as candidate genes for pigment changes in two fruit by RNA-seq and WGCNA analysis.

Through gene function verification, we proposed a new working model that explains the specific accumulation of anthocyanins and flavonols in black and white fruits on a metabolic and molecular level (**Figure 10**). In black fruits, LrAN2-like interacts with LrAN1b and LrAN11 to form an MBW complex, which regulates the downstream target genes *LrDFR* and *LrANS* promoters to regulate anthocyanin synthesis. When the accumulation of anthocyanins is too high in the later stages of fruit development, the MBW activation complex activates the genes of the *LrMYB3* and *LrETC1* repressors to achieve feedback inhibition. On the other hand, however, in white fruits, the expression of *LrAN1b* is abnormal, so this dynamic balance mechanism is broken. The MBW complex that depends on *LrAN1b* cannot effectively activate its target genes, especially the key genes *LrDFR* and *LrANS*. Anthocyanin metabolism and flavonols are important branches of flavonoid metabolism. The decrease in *LrDFR* expression leads to the substrate leading to the flavonol branch, as a result, the proportion of flavonols in white is higher than that in black fruits. In other words, *LrMYB3* and *LrETC1* cannot be effectively activated, resulting in the failure of the feedback regulation mechanism during fruit pigmentation. Therefore, abnormal expression of *LrAN1b* will destroy the flavonoid homeostasis network, and the content and proportion of flavonol and anthocyanin in white fruits will be different (**Figure 10**). A similar result was also reported in the white flesh of radish, which was positively correlated with low or undetectable *RsbHLH4* expression (Lai et al., 2020). Another latest study shows abnormal *bHLH3* expression disrupts the homeostatic regulatory network, causing differences in pigment composition among mulberry fruits (Li et al., 2020). Anthocyanin regulation is a complex and variable process. Changes in the synthesis of a gene and transcription factor will affect the composition of the tissue (Paauw et al., 2019). Therefore, the regulation of anthocyanin

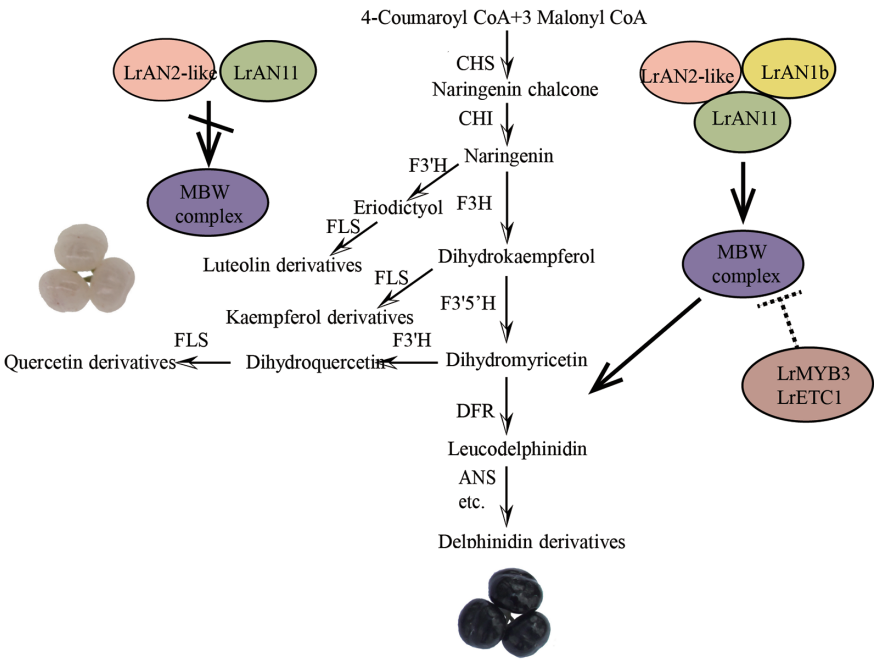

FIGURE 10 | A working model explains the reasons for the specific accumulation of anthocyanins and flavonols in black and white fruits, respectively.

metabolism of *L. ruthenicum* is worth further investigation and analysis. In this study, our results initially revealed the reasons for the different color of *L. ruthenicum* fruit, and laid the foundation for the analysis of the anthocyanin regulatory network of *L. ruthenicum*.

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

LT and YC conceived and designed the study. TL and YF performed the experiment, analyzed the data, and drafted the article. LT and YC reviewed and edited the manuscript. DL, YL, YY, and XQ contributed to data acquisition. HQ, GL, FC, and JW participated in the data analysis and interpretation. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by the Agro-Technical Independent Innovation Special Project of Ningxia Hui Autonomous Region, China (QCYL-2018-04,YES-16-0403), Foreign cooperation project of ningxia academy of agricultural and forestry sciences (DW-X-2018011) and the National Natural Science Foundation of China (31260351).

ACKNOWLEDGMENTS

We are grateful to Professor Yang Zhang (College of Life Sciences, Sichuan University, Chengdu 610065, China) for giving the PEAQ-HT vector.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.01256/full#supplementary-material

REFERENCES

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

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

Albert, N. W. (2015). Subspecialization of R2R3-MYB Repressors for Anthocyanin and Proanthocyanidin Regulation in Forage Legumes. *Front. Plant Sci.* 6, 1–13. doi: 10.3389/fpls.2015.01165

Bai, Y., Pattanaik, S., Patra, B., Werkman, J. R., Xie, C. H., and Yuan, L. (2011). Flavonoid-related basic helix-loop-helix regulators, NtAn1a and NtAn1b, of tobacco have originated from two ancestors and are functionally active. *Planta* 234, 363–375. doi: 10.1007/s00425-011-1407-y

Bo, L., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 12, 323. doi: 10.1186/1471-2105-12-323

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

Butelli, E., Titta, L., Giorgio, M., Mock, H. P., Matros, A., Peterek, S., et al. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.* 26, 1301–1308. doi: 10.1038/nbt.1506

Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., et al. (2013). A Novel Integrated Method for Large-Scale Detection, Identification, and Quantification of Widely Targeted Metabolites: Application in the Study of Rice Metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/sst080

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Espley, R. V., Brendolise, C., Chagne, D., Kutty-Amma, S., Green, S., Volz, R., et al. (2009). Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. *Plant Cell* 21, 168–183. doi: 10.1105/tpc.108.059329

Falcone Ferreyra, M. L., Rius, S. P., and Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3, 222. doi: 10.3389/fpls.2012.00222

Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, S. J., Hotz, H. R., et al. (2008). The Pfam protein families database. *Nucleic Acids Res.* 36, 281–288. doi: 10.1093/nar/gkm960

Fraga, C. G., Clowers, B. H., Moore, R. J., and Zink, E. M. (2010). Signature-discovery approach for sample matching of a nerve-agent precursor using liquid chromatography-mass spectrometry, XCMS, and chemometrics. *Anal. Chem.* 82, 4165–4173. doi: 10.1021/ac1003568

Francisco, R. (2013). ABCC1, an ATP Binding Cassette Protein from Grape Berry, Transports Anthocyanidin 3-O-Glucosides. *Plant Cell* 25, 1840–1854. doi: 10.1105/tpc.112.102152

Fujiwara, H., Tanaka, Y., Yonekura-Sakakibara, K., Fukuchi-Mizutani, M., and Kusumi, T. (1998). cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from Gentiana triflora. *Plant J.* 16, 421–431. doi: 10.1046/j.1365-313x.1998.00312.x

Gao, Q., Luo, H., Li, Y., Liu, Z., and Kang, C. (2019). Genetic modulation of RAP alters fruit coloration in both wild and cultivated strawberry. *Plant Biotechnol. J.* 18, 1550–1561. doi: 10.1111/pbi.13317

Gomez, C., Terrier, N., Torregrosa, L., Vialet, S., Fournier-Level, A., Verries, C., et al. (2009). Grapevine MATE-Type Proteins Act as Vacuolar H+-Dependent Acylated Anthocyanin Transporters. *Plant Physiol.* 150, 402–415. doi: 10.1104/pp.109.135624

Götz, S., JM, G.-G., Terol, J., TD, W., SH, N., MJ, N., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176

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

Horsch, R. B., Fry, J. E., Hoffmann, N. L., Eichholtz, D., Rogers, S. G., Fraley, R. T., et al. (1985). and General Method for hybridization revealed the expected. *Sci. (80-.).* 227, 1229–1231. doi: 10.1126/science.227.4691.1229

Hu, N., Zheng, J., Li, W., and Suo, Y. (2014). Isolation, Stability, and Antioxidant Activity of Anthocyanins from Lycium ruthenicum Murray and Nitraria Tangutorum Bobr of Qinghai-Tibetan Plateau. *Sep. Sci. Technol.* 49, 2897–2906. doi: 10.1080/01496395.2014.943770

Huang, J., Xing, M., Li, Y., Cheng, F., Gu, H., Yue, C., et al. (2019). Comparative Transcriptome Analysis of the Skin-Specific Accumulation of Anthocyanins in Black Peanut (Arachis hypogaea L.). *J. Agric. Food Chem.* 67, 1312–1324. doi: 10.1021/acs.jafc.8b05915

Jain, P. K., Kharya, M. D., Gajbhiye, A., Sara, U. V. S., and Sharma, V. K. (2008). Flavonoids as nutraceuticals. A review. *Herba Pol.* 7, 1089–1099. doi: 10.4314/tjpr.v7i3.14693

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, 480–484. doi: 10.1093/nar/gkm882

Koes, R., Verweij, W., and Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* 10, 236–242. doi: 10.1016/j.tplants.2005.03.002

Lai, B., Cheng, Y., Liu, H., Wang, Q., Wang, Q., Wang, C., et al. (2020). Differential anthocyanin accumulation in radish taproot: importance of RsMYB1 gene structure. *Plant Cell Rep.* 39, 217–226. doi: 10.1007/s00299-019-02485-z

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9:559. doi: 10.1186/1471-2105-9-559

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

Li, H., Yang, Z., Zeng, Q., Wang, S., Luo, Y., Huang, Y., et al. (2020). Abnormal expression of bHLH3 disrupts a flavonoid homeostasis network, causing differences in pigment composition among mulberry fruits. *Hortic. Res.* 7, 1–19. doi: 10.1038/s41438-020-0302-8

Likun, W., Zhixing, F., Xi, W., Xiaowo, W., and Xuegong, Z. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26, 136–138. doi: 10.1093/bioinformatics/btp612

Liu, H., Lou, Q., Ma, J., Su, B., Gao, Z., and Liu, Y. (2019). Cloning and Functional Characterization of Dihydroflavonol 4-Reductase Gene Involved in Anthocyanidin Biosynthesis of Grape Hyacinth. *Int. J. Mol. Sci.* 20, 4743. doi: 10.3390/ijms20194743

Lou, Q., Liu, Y., Qi, Y., Jiao, S., Tian, F., Jiang, L., et al. (2014). Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J. Exp. Bot.* 65, 3157–3164. doi: 10.1093/jxb/eru168

Mao, X., Cai, T., JG, O., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.2307/1592215

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

MD, Y., MJ, W., GK, S., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11, R14. doi: 10.1186/gb-2010-11-2-r14

MG, G., BJ, H., Yassour, M., JZ, L., DA, T., 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

Mol, J., Grotewold, E., and Koes, R. (1998). How gens paint flowers and seeds. *Trends Plant Sci.* 3, 212–217. doi: 10.1016/S1360-1385(98)01242-4

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

Morishita, T., Kojima, Y., Maruta, T., Nishizawa-Yokoi, A., Yabuta, Y., and Shigeoka, S. (2009). Arabidopsis NAC transcription factor, ANAC078, regulates flavonoid biosynthesis under high-light. *Plant Cell Physiol.* 50, 2210–2222. doi: 10.1093/pcp/pcp159

Nguyen, P. D., Ho, C.-L., Harikrishna, J. A., Wong, M. C. V.-L., and Abdul Rahim, R. (2007). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Trees (Berl. West)* 21, 3389–3402. doi: 10.5511/plantbiotechnology.19.145

Paauw, M., Koes, R., and Quattrocchio, F. M. (2019). Altgieration of Flavonoid Pigmentation Patterns During Domestication of Food Crops. *J. Exp. Bot.* 70, 3719–3735. doi: 10.1093/jxb/erz141

Payne, C. T., Zhang, F., and Lloyd, a. M. GL3 encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with GL1 and TTG1. *Genetics* 156, 1349–1362. doi: 10.1002/1526-968X(200011/12)28:3/4<167::AID-GENE120>3.0.CO;2-N

Quattrocchio, F., JF, W., van der W, J. N.M., and Koes, R. (1998). Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. *Plant J.* 13, 475–488. doi: 10.1046/j.1365-313x.1998.00046.x

Shannon, P., Markiel, A., Ozier, O., NS, B., JT, W., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Shim, J. S., Jung, C., Lee, S., Min, K., Lee, Y. W., Choi, Y., et al. (2013). AtMYB44 regulates WRKY70 expression and modulates antagonistic interaction between salicylic acid and jasmonic acid signaling. *Plant J.* 73, 483–495. doi: 10.1111/tpj.12051

Shin, J., Park, E., and Choi, G. (2007). PIF3 regulates anthocyanin biosynthesis in an HY5-dependent manner with both factors directly binding anthocyanin biosynthetic gene promoters in Arabidopsis. *Plant J.* 49, 981–994. doi: 10.1111/j.1365-313X.2006.03021.x

Spelt, C., Quattrocchio, F., Mol, J. N., and Koes, R. (2000). anthocyanin1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* 12, 1619–1632. doi: 10.1105/tpc.12.9.1619

Spelt, C., Quattrocchio, F., Mol, J., and Koes, R. (2002). ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. *Plant Cell* 14, 2121–2135. doi: 10.1105/tpc.003772

Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9440–9445. doi: 10.1073/pnas.1530509100

Su, W., Tao, R., Liu, W., Yu, C., Yue, Z., He, S., et al. (2020). Characterization of four polymorphic genes controlling red leaf colour in lettuce that have undergone disruptive selection since domestication. *Plant Biotechnol. J.* 18, 479–490. doi: 10.1111/pbi.13213

Sun, Y., Li, H., and Huang, J.-R. (2012). Arabidopsis TT19 Functions as a Carrier to Transport Anthocyanin from the Cytosol to Tonoplasts. *Mol. Plant* 5, 387–400. doi: 10.1093/mp/ssr110

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2010). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749. doi: 10.1111/j.1365-313X.2008.03447.x

Verweij, W., Spelt, C. E., Bliek, M., de Vries, M., Wit, N., Faraco, M., et al. (2016). Functionally Similar WRKY Proteins Regulate Vacuolar Acidification in Petunia and Hair Development in Arabidopsis. *Plant Cell* 28, 786–803. doi: 10.1105/tpc.15.00608

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017). GSA: Genome Sequence Archive*. *Genomics Proteomics Bioinforma.* 15, 14–18. doi: 10.1016/j.gpb.2017.01.001

Wang, H., Li, J., Tao, W., Zhang, X., Gao, X., Yong, J., et al. (2018). Lycium ruthenicum studies: Molecular biology, Phytochemistry and pharmacology. *Food Chem.* 240, 759–766. doi: 10.1016/j.foodchem.2017.08.026

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., et al. (2013). HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Res.* 41, 1–7. doi: 10.1093/nar/gks1065

Xue, J., Liu, Z., Zhu, H., Niu, X., and Ning, J. (2014). PW289 Anthocyanins Extracted From Lycium Ruthenicum Murray Alleciate Cardiac Cardiomyopathy In Experimental Diabetic Rats. *Glob. Heart* 9, e317–e317. doi: 10.1016/j.gheart.2014.03.2369

Yang, X., Lin, S., Jia, Y., Rehman, F., Zeng, S., and Wang, Y. (2019). Anthocyanin and spermidine derivative hexoses coordinately increase in the ripening fruit of Lycium ruthenicum. *Food Chem.* 311, 125874. doi: 10.1016/j.foodchem.2019.125874

Yin-Yan, Q.II, Chen, X. Y., Chen, W. S., Deng, L., and Zhu, C. Y. (2017). Clone and activity analysis of promoter fragment of the F3′5′H gene of Lycium ruthenicum Marr.and its albino fruits. *J. Agric. Univ. Hebei* 40, 51–59. doi: 10.13320/j.cnki.jauh.2017.0125

Zeng, S., Wu, M., Zou, C., Liu, X., Shen, X., Hayward, A., et al. (2014). Comparative analysis of anthocyanin biosynthesis during fruit development in two Lycium species. *Physiol. Plant* 150, 505–516. doi: 10.1111/ppl.12131

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

Zhang, B., Chopra, D., Schrader, A., and Hülskamp, M. (2019). Evolutionary comparison of competitive protein-complex formation of MYB, bHLH, and WDR proteins in plants. *J. Exp. Bot.* 70, 3197–3209. doi: 10.1093/jxb/erz155

Zhang, S., He, F., Chen, X., and Ding, K. (2019). Isolation and structural characterization of a pectin from Lycium ruthenicum Murr and its anti-pancreatic ductal adenocarcinoma cell activity. *Carbohydr. Polym.* 223, 115104. doi: 10.1016/j.carbpol.2019.115104

Zhang, J., Lei, Y., Wang, B., Li, S., Yu, S., Wang, Y., et al. (2020). The high-quality genome of diploid strawberry (Fragaria nilgerrensis) provides new insights into anthocyanin accumulation. *Plant Biotechnol. J.* doi: 10.1111/pbi.13351

Zhang, S., Chen, Y., Zhao, L., Li, C., Yu, J., Li, T., et al. (2020). A novel NAC transcription factor, MdNAC42, regulates anthocyanin accumulation in red-fleshed apple by interacting with MdMYB10. *Tree Physiol.* 40, 413–423. doi: 10.1093/treephys/tpaa004

Zhang, Z., Zhao, W., Xiao, J., Bao, Y., He, S., Zhang, G., et al. (2020). Database Resources of the National Genomics Data Center in 2020. *Nucleic Acids Res.* 48, D24–D33. doi: 10.1093/nar/gkz913

Zhao, J., and Dixon, R. A. (2010). The 'ins' and 'outs' of flavonoid transport. *Trends Plant Sci.* 15, 0–80. doi: 10.1016/j.tplants.2009.11.006

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

Zheng, J., Wu, H., Zhu, H., Huang, C., Liu, C., Chang, Y., et al. (2019). Determining factors, regulation system, and domestication of anthocyanin biosynthesis in rice leaves. *New Phytol.* 223, 705–721. doi: 10.1111/nph.15807

Zhu, Z. J., Schultz, A. W., Wang, J., Johnson, C. H., Yannone, S. M., Patti, G. J., et al. (2013). Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database. *Nat. Protoc.* 8, 451–460. doi: 10.1038/nprot.2013.004

Zhu, Z., Wang, H., Wang, Y., Guan, S., Wang, F., Tang, J., et al. (2015). Characterization of the cis elements in the proximal promoter regions of the anthocyanin pathway genes reveals a common regulatory logic that governs pathway regulation. *J. Exp. Bot.* 66, 3775–3789. doi: 10.1093/jxb/erv173

Zong, Y., Zhu, X., Liu, Z., Xi, X., Li, G., Cao, D., et al. (2019). Functional MYB transcription factor encoding gene AN2 is associated with anthocyanin biosynthesis in Lycium ruthenicum Murray. *BMC Plant Biol.* 19, 169. doi: 10.1186/s12870-019-1752-8

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 September 2020
doi: 10.3389/fpls.2020.573531

High Altitude Is Beneficial for Antioxidant Components and Sweetness Accumulation of Rabbiteye Blueberry

Qilong Zeng[1†], *Gangqiang Dong*[2†], *Liangliang Tian*[1], *Han Wu*[3], *Yongjun Ren*[4], *Guy Tamir*[5], *Wuyang Huang*[3*] and *Hong Yu*[1*]

[1] Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China, [2] Amway (China) Botanical R&D Center, Wuxi, China, [3] Institute of Farm Product Processing, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [4] School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, China, [5] Agricultural Research and Development, Central Mountain Region, Tekoa, Israel

OPEN ACCESS

Edited by:
Pedro Mena,
University of Parma, Italy

Reviewed by:
José Juan Ordaz-Ortiz,
National Polytechnic Institute of Mexico (CINVESTAV), Mexico
Qian Shen,
Shanghai Jiao Tong University, China

***Correspondence:**
Wuyang Huang
wuyanghuang@hotmail.com
Hong Yu
njyuhong@vip.sina.com

[†]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: 03 July 2020
Accepted: 10 September 2020
Published: 25 September 2020

Citation:
Zeng Q, Dong G, Tian L, Wu H, Ren Y, Tamir G, Huang W and Yu H (2020) High Altitude Is Beneficial for Antioxidant Components and Sweetness Accumulation of Rabbiteye Blueberry. Front. Plant Sci. 11:573531. doi: 10.3389/fpls.2020.573531

To better understand the effect of growing location on the phytochemical compounds and sensory properties of blueberry (*Vaccinium* spp.), here we investigated rabbiteye blueberry 'Brightwell' (*Vaccinium ashei* cv. 'Brightwell') grown in 10 locations of China. Significant differences in terms of total soluble solids, titratable acidity, flavonoids, phenols, as well as proanthocyanidins and anthocyanins, were found in the fruits (berries) of blueberry plants among the different sampled locations. Furthermore, their sensory properties, which evaluated by the electronic tongue method, also significantly differed among the 10 locations. The content of flavonoids, phenols, proanthocyanidins, and anthocyanins all had significant correlations with sensory properties, except that of aftertaste-astringency. A key finding to emerge was that blueberry plants grown at high altitude locations harbored a high content of total soluble solids, flavonoids, phenols, proanthocyanidins, and anthocyanins along with high scores for the sweetness. These results suggested cultivating blueberry at high altitude can produce fruit that not only possess pronounced beneficial health effects but also good taste.

Keywords: ***Vaccinium ashei*, antioxidant compounds, anthocyanin, electronic tongue, sensory properties**

INTRODUCTION

The cultivation of blueberry (*Vaccinium* spp.) in China has garnered dramatically rising interest over the last 15 years, with its production area having reached 22,000 hm^2, or 16.3% that of world's total (Li et al., 2016; Brazelton et al., 2017). Meanwhile, global blueberry acreage is continuing to increase due to their nutritional and health benefits (Scalzo et al., 2013; Upadhaya and Dwivedi, 2019).

Blueberry is a richer source of antioxidant compounds (flavonoid, polyphenolic compounds, and especially anthocyanins) when compared to other berries, such as strawberry or raspberry (Cardeñosa et al., 2016). Blueberry cultivars high in antioxidant compounds could command a price premium when marketing them to consumers (Gilbert et al., 2015; Kraujalytė et al., 2015;

Gallardo et al., 2018). Furthermore, the sensory quality of the harvested berries is another important aspect influencing the behavior of their would-be consumers. Consumer preference studies have suggested that key sensory properties, such as sweetness, freshness, and juiciness, were another main psychological drive for purchasing blueberries (Saftner et al., 2008; Gilbert et al., 2014; Yue and Wang, 2016).

To better accommodate consumer preferences, hundreds of blueberry cultivars have been bred by scientists worldwide. Among them, phytochemical compounds were found to vary widely (Connor et al., 2002; Gilbert et al., 2014). Similarly, there exists much variation in the sensory scores for intensity of blue color, juiciness, sweetness, and blueberry-like flavor among different blueberry cultivars currently on the market (Saftner et al., 2008).

Phytochemical compounds in blueberry fruits were also considerably affected by the plant's growing locations and the interaction between environment and cultivars, although contradictory results have been reported (Zoratti et al., 2015; Correia et al., 2016). Sensory properties are induced by a variety of phytochemical compounds, such as sugars, acids, polyphenol composition, total soluble solids (TSS), titratable acidity (TA), and other quality measurements related with sweetness, sourness, and blueberry-like flavor (Saftner et al., 2008; Barrett et al., 2010; Bett-Garber et al., 2015). Therefore, it is reasonable to speculate that blueberry's sensory properties may also be influence by growing location. Yet, surprisingly, very little information is available on the effect of growing location upon blueberry sensory properties, and their correlation with phytochemical compounds (Saftner et al., 2008; Bett-Garber et al., 2015).

Traditionally, the taste of blueberries has been evaluated using human sensory panels, which provides an integrated, direct measurement of perceived intensities of target taste attributes (Hanson et al., 1993; Trigo et al., 2004; Moreno et al., 2007). However, this approach is quite time-consuming and expensive and it is confined to certain conditions (Liang et al., 2003; Liu et al., 2017). Electronic tongue ("E-tongue") systems have been successfully developed to mimic the organization of human taste buds, which are sensor arrays (Wei and Wang, 2013; Phat et al., 2016). This instrument is widely used for testing liquid foodstuffs, such as water (Carbó et al., 2018), wine (Buratti et al., 2007), fruit juice (Qiu et al., 2015), and tea (Chen et al., 2008). It has proven to be a promising, rapid, and low-cost technique for the discrimination of blueberry juices produced from different plant genotypes (Kraujalytė et al., 2015). Accordingly, in this study, sensory taste variability of blueberry from different locations in China were evaluated using the E-tongue, with the aim of exploring the effect of growing location on blueberry sensory properties, and to clarify the relationship between sensory properties and phytochemical compounds in blueberry. Knowledge gained in our study may provide a useful guide for selecting orchard locations with a view towards improving fruit berry quality.

MATERIALS AND METHODS

Plant Materials

Rabbiteye blueberry (*Vaccinium ashei* cv. 'Brightwell') fruits were sampled from five major blueberry-growing provinces of China, in July 2017 (**Figure 1**). Ten blueberry orchards known for their abundant yields were selected as sampling sites, whose location and altitude are listed in **Table 1**.

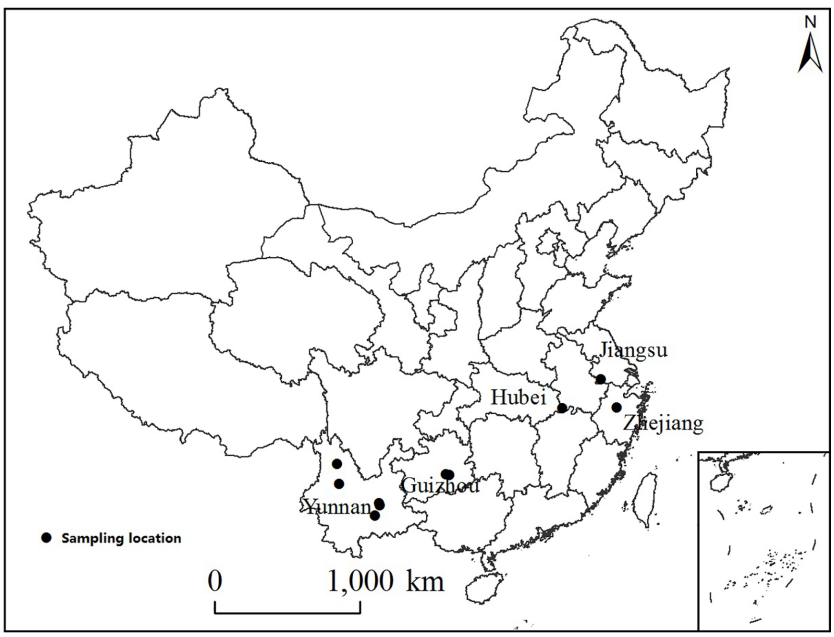

FIGURE 1 | The locations of 10 fields sampled in five provinces of China.

TABLE 1 | Geographical details for the 10 sampled sites of blueberry in China.

Location	Abbreviation	Latitude (N)	Longitude (E)	Altitude (m)
Caota town, Zhejiang province	CT	29° 38' 30"	120° 1' 8"	199
Baima town, Jiangsu province	BM	31° 30' 20"	119° 8' 51"	49
Dushan town, Hubei province	DS	30° 2' 24"	116° 4' 21"	37
Xuanwei town, Guizhou province	XW	26° 22' 56"	107° 44' 24"	750
Xianchang town, Guizhou province	XC	26° 25' 21"	107° 31' 54"	912
Shigu town, Yunnan province	SG	26° 59' 40"	99° 56' 30"	1,950
Yinqiao town, Yunnan province	YQ	25° 45' 28"	100° 7' 13"	2,010
Yousuo town, Yunnan province	YS	24° 38' 21"	102° 55' 51"	1,731
Haikou town, Yunnan province	HK	24° 31' 28"	102° 56' 52"	1,781
Longpeng town, Yunan province	LP	23° 51' 24"	102° 39' 58"	1,840

Blueberries, that are the fruits, were manually picked at the commercial ripening stage, when the berries were fully developed in size and blue in color. Five plants (replicates) were randomly chosen in each orchard, and with, approximately, 120 g of blueberries randomly collected from each plant. These samples of harvested blueberries were then pooled and stored in a sealed foam box with ice to prevent their degradation while moving them from field to the laboratory, where they were stored at −20°C. All metabolic analyses were performed within 3 months from this field collection.

Chemicals and Reagents

Folin–Ciocalteu reagent, catechin, and cyanidin 3-glucoside were purchased from Sigma-Aldrich (St. Louis, MO, USA), while rutin and gallic acid were purchased from J&K chemicals Ltd (Beijing, China). The HPLC solvents were bought from EMD (Darmstadt, Germany). Other chemicals and reagents used in this study were of analytical grade and were obtained from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China).

Sample Preparation and Extraction

All analytical determinations were applied to a fruit homogenate, obtained by crushing about 50 g of frozen berries in a food blender (Philip, Shanghai, China). A portion of each sample's homogenate was used to determine TSS content, TA, and sensory properties; the other part was used to extract phenols, flavonoids, proanthocyanidin, and anthocyanin

The extraction method was adapted from Wang et al. (2015), with some modifications. Approximately 1 g of each homogenized blueberry sample was put into a 50 ml centrifuge tube, in which it was mixed with 15 ml of 85% methanol containing 0.5% formic acid, and then shaken in a vortexer for 15 s to ensure even mixing. Each sample was sonicated at 20°C, for 20 min, in a KQ-5200E ultrasonic cleaner (Shumei, Kunshan, China), and then centrifuged at 5,000 rpm for 10 min. The supernatant was removed, and another 15 ml of the methanol extraction solution was added to the residue. This extraction operation was repeated three times per sample, until the berries had become colorless. These extracts were combined and stored at −20°C for later chemical analyses.

Total Soluble Solid (TSS) and Titritable Acid (TA) Content

The TSS of berries was measured with a hand refractometer (Co. Atago, Japan) on the juice obtained by squeezing the homogenized blueberries. To determine TA, a 2.5 g sample of homogenized blueberries was first diluted with 25 ml of distilled water, after which TA was quantified with a neutralizing 0.1 M NaOH solution to an end-point of pH 8.1. Values of TA were expressed as the percent citric acid (mass/mass) on a fresh weight (FW) basis of fruit; the ratio between TSS and TA ratio was calculated as a taste indicator.

Total Phenolic Content

Total phenolic content (TP) was estimated using the Folin–Ciocalteu colorimetric method, as described by Li et al. (2013). Briefly, appropriate dilutions of the extracted samples (0.2 ml) were oxidized with 2 ml of 0.5 mol L^{-1} Folin–Ciocalteu reagent, for 4 min, at room temperature. Then, the reactive solution was neutralized, by 2 ml of 75 mg ml^{-1} saturated sodium carbonate, and its absorbance was measured at 760 nm after incubation for 2 h at room temperature in the dark. Quantifications were made based on the standard curve of gallic acid.

Total Flavonoid Content

Total flavonoid content (TF) was measured by a modified colorimetric method (Huang et al., 2012). The appropriate dilutions of the extracted samples (1 ml) were mixed with 0.1 ml of 0.05 g ml^{-1} $NaNO_2$. After 6 min, to each 0.1 ml of 0.1 g ml^{-1} $AlCl_3$ solution was added; 5 min later, 1 ml of 1 mol L^{-1} NaOH was added to the reactive solution. This was then mixed well and allowed to stand for 15 min before measuring its absorbance at 510 nm. The TF was calculated and expressed as rutin equivalents.

Proanthocyanidin Content

Proanthocyanidin content (PAC) of the blueberry extract was determined by the colorimetric assay developed by Price et al. (1978). Briefly, a 0.5% (w/v) solution of vanillin-HCl reagent (0.5% vanillin in 4% concentrated HCl, in methanol; 2.5 ml) was added to 0.5 ml of diluted extracts and standard (catechin) solutions, mixed thoroughly, and incubated at 30°C in the dark for 20 min. The absorbance of this solution was recorded at 500 nm against the corresponding blanks. The PAC was calculated and expressed as catechin equivalents.

Total Anthocyanin Content

The blueberry extracts were filtered through a 0.22 μm filter (Millipore), and then injected to high performance liquid

chromatography (HPLC) for analysis. This HPLC analysis was performed in an Aglient 1100 HPLC system (Agilent Technologies, USA) equipped with a binary pump and a diode-array detector (DAD). For the chromatographic analysis, a 250 mm × 4.6 mm, 5-μm particle size, end-capped reverse-phase Zorbax SB-C18 column was used (Agilent Technologies, USA). The running temperature was 35°C, with an injection volume of 10 μl. The detections were made at 520 nm, at a flow rate of 0.6 ml/min. Mobile phase A was a mixture of 6% HAc (ethanoic acid) and ultrapure water, whereas mobile phase B was a mixture of 6% HAc and acetonitrile. The gradient used was as follows: 5 to 10% B (from 0 to 5 min), 10 to 15% B (from 5 to 20 min), 15 to 20% B (from 20 to 35 min), 20 to 40% B (from 35 to 40 min), 40 to 80% B (from 40 to 45 min), 80 to 85% B (from 45 to 50 min), 85 to 5% B (from 50 to 55 min), and 5% B (from 55 to 60 min). To quantify anthocyanins, their peak areas were compared to the absorbance of a cyanidin 3-glucoside external standard (Sigma-Aldrich, Shanghai, China). Total anthocyanin content (hereon, ACs) was calculated as the sum of these peaks and expressed on a FW basis of fruit.

E-Tongue Analysis

The overall taste profile of the blueberry was measured by an E-tongue (TS-SA402B, Intelligent Sensor Technology Inc., Kanagawa, Japan). This device consisted of an automatic sampler, a selective sensor array of six sensors (sourness, bitterness, astringency, umami, saltiness, richness, aftertaste-astringency, and aftertaste-bitterness), a signal acquisition instrument, and pattern recognition software. The surface of the selective sensors is combined with artificial lipid membranes has different response properties to chemicals on the basis of their taste. The selective sensors used in this study were: GL1 specific for sweetness, CA0 for sourness, C00 for bitterness and aftertaste-bitterness, AE1 for astringency and aftertaste-astringency, AAE for umami, CT0 for saltiness. The measurement principle of the E-tongue is based on the capability of tasty substances to change the potential of the sensors through electrostatic or hydrophobic interaction with the hydrophilic and hydrophobic groups of the lipid membranes. The sensors were dipped into the calibrated solution, which was 30 mmol L^{-1} potassium chloride and 0.3 mmol L^{-1} tartaric acid, then the electric potential measured for each sensor was defined as Vr. Then the sensors were dipped into the blueberry solution for 30 s, and the electric potential for each sensor was defined as Vs. The relative value (Rv) was represented by the differences between the potentials of blueberry sample and the reference solution (Vs-Vr). Then sensors were rinsed with fresh calibrated solution for 6 s and then dipped into the reference solution again. The new potential for the reference solution was defined as Vra. The difference Vra-Vr between the potentials of the calibrated solution after and before sample measurement is the "Change of Membrane Potential caused by Absorption value" (CPAv) and corresponds to the "aftertaste". Before a new cycle measurement, sensors were rinsed for 90 s with a washing solution (30% ethanol) and then for 180 s with the calibrated solution. The "taste values" were calculated by multiplying the Rv and CPAv of the sensors for coefficients based on Weber-Fechner law, which gives the intensity of sensation considering the sensor properties for tastes (Kobayashi et al., 2010). For preparing the blueberry samples, the blueberry homogenate was first mixed with distilled water (1:2, g/ml) by vortexing. This mixture was then put into a 50 ml centrifuge tube and centrifuged at 5,000 rpm for 10 min. Next, the supernatant was filtered through three layers of gauze and the ensuing clear liquid was poured into the special beaker of the E-tongue, and analyzed at room temperature. In conducting this data collection sequence, it alternated between the tested calibrated solution and a given blueberry sample and repeated four times for both. The average of the last three sets of obtained data was used for the statistical analysis.

Data Analysis

All data are expressed here as the mean ± SD (standard deviation). Statistical analysis of the above data relied on one-way analysis of variance (ANOVA), followed by Duncan's multiple test, to identify significant differences among samples of blueberry grown at different locations, using SPSS v18.0 (SPSS Inc., Chicago, USA). An alpha level of $p < 0.05$ or less was considered to be significant.

Correlation analysis was performed among phytochemical conpounds and sensory properties using SPSS v18.0 (SPSS Inc., Chicago, USA). The correlation coefficients were analyzed by Pearson method, and significance was tested by two-tailed method. An alpha level of $p < 0.05$ or less was considered to be significant.

In order to test for similarities among blueberries sampled from different locations, principal component analysis (PCA) was applied to the data set. Locations were inserted in rows and response variables including phytochemical properties, and sensory values were placed in columns. The PCA analysis was implemented in the Origin2019 software (Massachusetts, USA).

RESULTS AND DISCUSSION

Phytochemical Compounds of Blueberry Affected by Growing Locations

Significant differences in blueberry TSS, TA, and TSS/TA were found among the different growing locations (**Table 2**). The TSS among the 10 locations ranged from 9.4 to 14.2%, TA ranged from 0.33 to 0.55%, and TSS/TA ranged from 20 to 37.7. Blueberry grown in Shigu had the highest TSS compared with the other locations, having also the lowest TA content and highest TSS/TA. These results show that where blueberry is grown has a significant impact on its fruit berries' TSS and TA, which is in line with other results (Correia et al., 2016).

Growing location also had a significant effect on the TF, TP, PAC, and ACs of blueberry (**Table 3**). Across the 10 locations, the corresponding values for TF, TP, PAC, and ACs ranged from 0.94–2.07 mg rutin g^{-1} FW, 1.44–3.53 mg gallic acid g^{-1} FW, 2.40–9.53 mg catechin g^{-1} FW, and 0.26–1.34 mg Cyanidin-3-glucoside g^{-1} FW. Blueberry grown in Caota and Baima towns had the lowest TF, TP, PAC, and ACs content, whereas that grown in Shigu, Yinqiao, Yousuo, and Longpeng towns had the highest TF content, while Yinqiao and Longpeng towns had the

TABLE 2 | The content of total soluble solids (TSS), titratable acidity (TA), and TSS/TA in blueberry from different locations in China.

Location	TSS (%)	TA (%)	TSS/TA
CT[1]	10.3 ± 0.83de	0.52 ± 0.04ab	20.0 ± 1.84e
BM	12.1 ± 1.26b	0.55 ± 0.03a	22.0 ± 2.33de
DS	10.9 ± 0.73cd	0.39 ± 0.03e	28.3 ± 3.66bc
XW	10 ± 0.00de	0.33 ± 0.03f	30.7 ± 3.32b
XC	9.4 ± 0.88e	0.47 ± 0.04bc	20.2 ± 2.64e
SG	14.2 ± 1.27a	0.38 ± 0.04e	37.7 ± 5.75a
YQ	11.4 ± 0.77bc	0.42 ± 0.08de	28.3 ± 5.69bc
YS	10.6 ± 0.78cd	0.34 ± 0.02f	31.4 ± 2.78b
HK	10.1 ± 0.55de	0.49 ± 0.05bc	21.0 ± 2.03e
LP	11.4 ± 1.47bc	0.46 ± 0.03cd	25.1 ± 3.52cd

Values are the mean ± SD, n = 5; different letters within each column are significantly different at p < 0.05 according to Duncan' s test.

[1]CT, Caotao town; BM, Baima town; DS, Dushan town; XW, Xuanwei town; XC, Xianchang town; SG, Shigu town; YQ, Yinqiao town; YS, Yousuo town; HK, Haikou town; LP, Longpeng town.

highest TP content and Longpeng town had the highest PAC and ACs contents. Significant correlations among blueberry's TF, TP, PAC, and ACs content were found in the samples whose contents were significantly positively correlated with the altitude of their growing locations (**Table 4**). Studies involving altitude have provided contradictory results for the phenols content of both blueberries and bilberries (Martz et al., 2010; Zoratti et al., 2015). Some researchers also found no clear relationships, or negative ones, between altitude and the ACs content of blueberries (Åkerström et al., 2010; Martz et al., 2010). In later work, Wang et al. (2015) and Zoratti et al. (2015) reported the ACs in blueberries and bilberries increased when grown at higher altitudes. In our study, the results suggested the TF, TP, PAC, and ACs content of blueberry increased with altitudes rising from 37 to 2,010 m. This could be due to the plant's protective mechanisms against the stronger solar radiation and ultraviolet light generally prominent in high altitude regions (Jaakola et al., 2004; Zoratti et al., 2015). Such strong radiation and ultraviolet light could increase free radicals, such as reactive oxygen species (ROS), which activate the biosynthesis of flavonoids, especially anthocyanins to protect plant organs (Dudareva et al., 2020).

The TSS and TA significantly increased and decreased with altitude, respectively, meanwhile the TSS/TA was increased with the rising altitude (**Table 4**). This suggested blueberry grown at high altitude could be sweeter, since the TSS/TA is considered as a rough but reliable indicator of fruit sweetness and is recognized as a key factor contributing to the flavor of blueberries (Almutairi et al., 2017). Our results agree with reports on mandarin (Rokaya et al., 2016) and apricots (Naryal et al., 2019); however, Correia et al. (2016) reported that altitude did not affect the TSS in blueberry, and an inverse relation was observed in bilberry and pomegranate, in that a higher TSS content predominated at lower altitudes (Mikulic-Petkovsek et al., 2015; Mphahlele et al., 2016). From our results, we infer that high altitude climatic conditions are favorable for blueberry's accumulation of sugar contents. Sunlight at high

TABLE 3 | The content of total flavonoids (TF), phenols (TP), proanthocyanidin (PAC), and anthocyanins (ACs) in blueberry from different locations in China.

Location	TF (mg rutin g^{-1})	TP (mg gallic acid g^{-1})	PAC (mg catechin g^{-1})	ACs (mg cyanidin 3-glucoside kg^{-1})
CT[1]	0.94 ± 0.19d	1.44 ± 0.24e	2.40 ± 0.63f	0.26 ± 0.07g
BM	1.11 ± 0.14cd	1.67 ± 0.25de	3.37 ± 0.89ef	0.40 ± 0.12fg
DS	1.24 ± 0.22c	1.97 ± 0.36d	4.47 ± 1.07e	0.59 ± 0.17de
XW	1.56 ± 0.20b	2.440.34c	6.10 ± 1.21cd	0.83 ± 0.18bc
XC	1.26 ± 0.08c	1.98 ± 0.17d	3.48 ± 0.41ef	0.45 ± 0.07ef
SG	1.96 ± 0.34a	3.07 ± 0.60b	7.07 ± 1.91bd	0.90 ± 0.20bc
YQ	2.06 ± 0.23a	3.29 ± 0.49ab	7.21 ± 0.92bc	0.93 ± 0.09b
YS	2.07 ± 0.21a	3.10 ± 0.43b	7.67 ± 1.36b	1.02 ± 0.18b
HK	1.65 ± 0.13b	2.89 ± 0.38b	5.88 ± 0.91d	0.74 ± 0.13cd
LP	2.06 ± 0.28a	3.53 ± 0.47a	9.53 ± 1.75a	1.34 ± 0.27a

Values are the mean ± SD, n = 5; different letters within a column are significantly different at p < 0.05 according to Duncan's test.

[1]CT, Caotao town; BM, Baima town; DS, Dushan town; XW, Xuanwei town; XC, Xianchang town; SG, Shigu town; YQ, Yinqiao town; YS, Yousuo town; HK, Haikou town; LP, Longpeng town.

TABLE 4 | Pearson correlation coefficients for altitude and phytochemical contents of blueberry.

	Altitude	TSS[1]	TA	TSS/TA	TF	TP	PAC	ACs
Altitude	1							
TSS	0.23*	1						
TA	-0.30**	-0.04	1					
TSS/TA	0.36**	0.60**	-0.80**	1				
TF	0.83**	0.28*	-0.48**	0.52**	1			
TP	0.83**	0.26*	-0.37**	0.43**	0.95**	1		
PAC	0.73**	0.28**	-0.45**	0.50**	0.91**	0.96**	1	
ACs	0.70**	0.26*	-0.45**	0.47**	0.88**	0.93**	0.99**	1

*The * and ** indicate significance at p < 0.10 and p < 0.05, respectively.*

[1]TSS, total soluble solids; TA, titratable acidity; TSS/TA, the ratio between TSS and TA; TF, total flavonoids; TP, total phenols; PAC, proanthocyanidin; ACs, anthocyanins.

altitudes is of higher intensity than at lower altitudes, contributing to an accelerated photosynthetic rate and metabolite accumulation (Naryal et al., 2019). Another factor that could increase TSS content arises from reduced respiration caused by the low temperatures characterizing high altitude zones, which would promote the synthesis and accumulation of carbohydrates in fruits (MacKenzie et al., 2011).

Correlations Among Phytochemical Compounds of Blueberry

Significant correlations were also found between phenolic compounds and TAA, TA, and SS/TA (**Table 4**). These results were similar to findings from other studies, mainly on grape (Matsumoto et al., 2007) and pomegranate (Fawole and Opara, 2013). It may be that blueberry with high TSS are better able to sustain a high content of phenolic compounds, since physiologically they both increase simultaneously during the maturity stage of berries (Ribera et al., 2010).

Correlations Between Sensory Properties and Phytochemical Compounds of Blueberry

Significant differences were found among blueberry sensory scorings evaluated by the E-tongue among the 10 sampled locations (**Figure 2**). Furthermore, significant negative correlations were evident between altitude and sourness, astringency, aftertaste-bitterness, and richness, while altitude was significantly positively correlated with bitterness, umami, and sweetness of the berries (**Table 5**). Additionally, TSS was significantly positively correlated with bitterness, aftertaste astringency as well as umami and sweetness, yet negative correlated with sourness. The TA was also correlated with

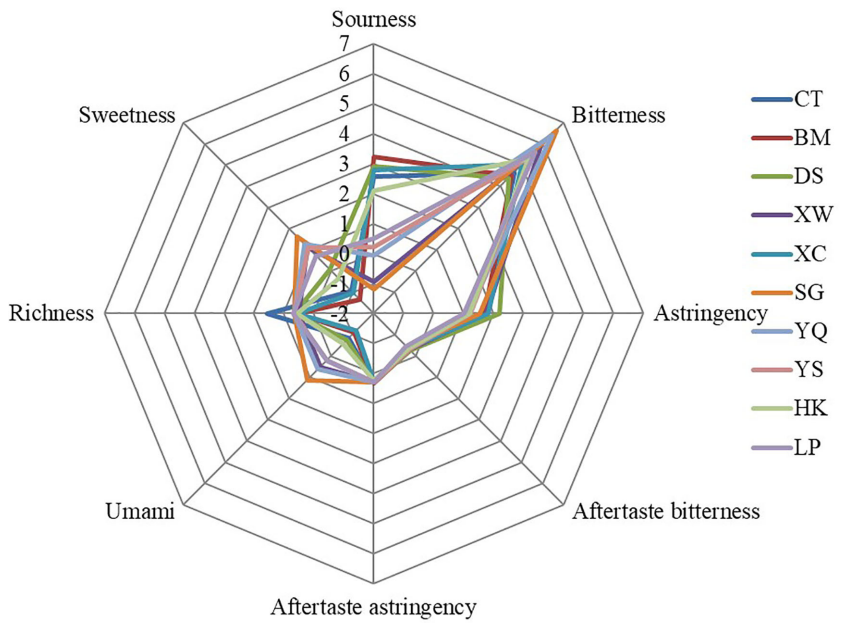

FIGURE 2 | Spider plot for sensory scores of blueberry sampled from different locations in China. CT, Caotao town; BM, Baima town; DS, Dushan town; XW, Xuanwei town; XC, Xianchang town; SG, Shigu town; YQ, Yinqiao town; YS, Yousuo town; HK, Haikou town; LP, Longpeng town.

TABLE 5 | Pearson correlation coefficients between altitude, phytochemical properties, and sensory scorings of blueberry.

	Sourness	Bitterness	Astringency	Aftertastebitter	Aftertaste astringency	Umami	Richness	Saltiness	Sweetness
Altitude	-0.57**	0.76**	-0.70**	-0.48**	-0.14	0.62**	-0.34**	-0.76**	0.51**
TSS[1]	-0.28**	0.37**	-0.14	-0.20	0.30**	0.31**	0.04	-0.26*	0.27**
TA	0.64**	-0.47**	-0.09	-0.02	-0.03	-0.63**	0.34**	0.34**	-0.70**
TSS/TA	-0.67**	0.58**	-0.02	-0.10	0.17	0.67**	-0.24*	-0.40**	0.69**
TF	-0.60**	0.66**	-0.62**	-0.30**	0.09	0.64**	-0.42**	-0.68**	0.61**
TP	-0.54**	0.62**	-0.66**	-0.34**	-0.01	0.59**	-0.42**	-0.71**	0.57**
PAC	-0.58**	0.57**	-0.60**	-0.30**	0.07	0.62**	-0.39**	-0.62**	0.62**
ACs	-0.56**	0.56**	-0.58**	-0.32**	0.03	0.60**	-0.36**	-0.59**	0.60**

*The * and ** values indicate significance at $p < 0.10$ and $p < 0.05$, respectively.*

[1]TSS, total soluble solids; TA, titratable acidity; TSS/TA, the ratio between TSS and TA; TF, total flavonoids; TP, total phenols; PAC, proanthocyanidin; ACs, anthocyanins.

sensory scores but its trend was the opposite of that shown by TSS. The TF, TP, PAC, and ACs all had a similar association with the sensory values as found for TSS, which was positively correlated with bitterness, umami, and sweetness, and negatively correlated with sourness, astringency, aftertaste-bitter, and richness of the blueberries. Unlike the findings of Drewnowski and Gomez-Carneros (2000), we found that a high concentration of anthocyanin and total phenolics increased the astringent tastes of blueberry, a result similar to that reported by Bett-Garber et al. (2015), who studied the relationship between physicochemical contents and sensory values of six blueberry cultivars, and found that a sweet taste was positively correlated with the TSS and anthocyanin content of blueberry, while its sour taste was positively correlated with its TA.

Sensory scorings and phytochemical analyses were also processed by PCA, for which the two-dimensional biplots of scores and loading of the different locations are presented in **Figure 3**. Principal component 1 (PC1) and PC2 respectively represented 53.3 and 14.2% of the total variance in the data set. The PC1 was dominated by sensory scores: umami, and sweetness, as well as phytochemical contents: TF, TP, PAC, and ACs. The PC2 was dominated by sensory scores: astringency, aftertaste-astringency and aftertaste-bitterness as well as some phytochemical content: TA and TSS/TA (**Figure 3**). The short distance among TF, TP, PAC, and ACs suggested they have a similar influence upon the evaluation of blueberry's food quality. The PCA plot showed that blueberry from different locations could be effectively discriminated, although the values for some locations were clustered and showed some overlap (**Figure 3**). Blueberry plants grown at high altitudes were located on the right of the x-axis, which suggested their fruits had high sensory scores for umami and sweetness, and a high content of TF, TP, PAC, and ACs. These results indicated that for the production of high-quality berries with respect to their sweetness and ACs content, it would be better to established blueberry fields at high altitudes. However, it was hard to discriminate the uality of blueberry from closing altitudes, due to the overlapping (**Figure 2B**). Therefore, looking ahead, more environmental factors should be considered, such as soil type and cultivation methods, since these two have an important role to play in determining blueberry properties and quality as food (Connor et al., 2002; Ehret et al., 2014; Zoratti et al., 2015).

CONCLUSION

A systematic evaluation of various quality characteristics and sensory values of blueberry 'Brightwell' from 10 locations in China demonstrated that both sets were affected by their location, while blueberry sensory aspects were markedly influenced by co-occurring phytochemical contents. Bitterness, umami, and sweetness of berries were positively correlated with TSS, TF, TP, PAC, and ACs in blueberry; their sourness was positively correlated with TA but negatively correlated with TSS, TF, TP, PAC, and ACs. In any case, the sensory scores and values of TSS, TA, TF, TP, PAC, and ACs clearly depended on altitude. Taken together, we suggest that blueberry plants grown at high altitude are apt to yield high-quality berries with regard to their sweetness and ACs content.

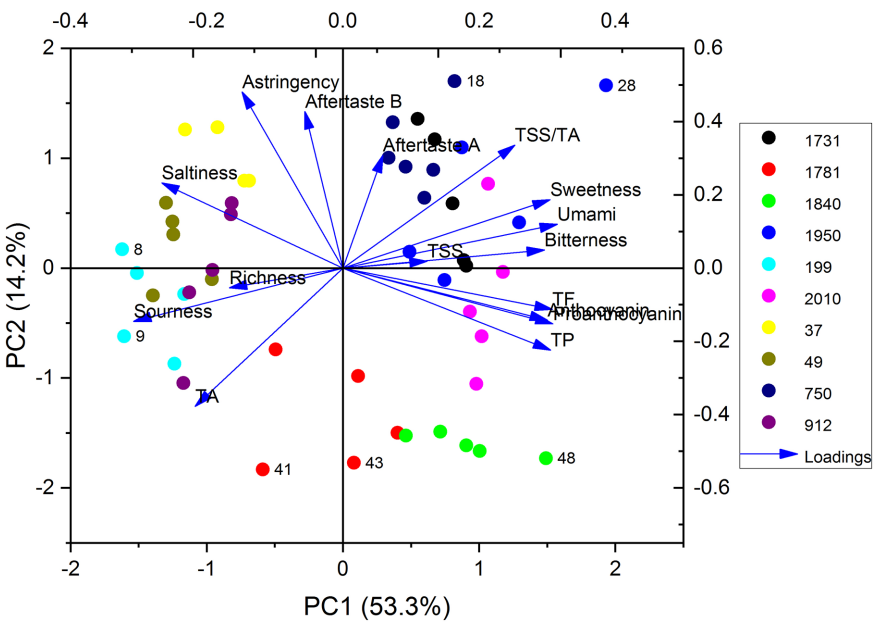

FIGURE 3 | Principal component analysis (PCA) of phytochemical analyses and sensory scores of blueberry from different locations in China.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

WH and HY conceived, designed the experiments, and revised the manuscript. QZ, GD, LT, and HW performed the experiments. QZ and GD wrote the manuscript. YR, QZ, and GD analyzed the data. GT revised the manuscript and improved the language. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the National Natural Science Foundation of China (Nos. 31301838, 31601709); the Natural Science Foundation of Jiangsu Province, China (Nos. BK20160597, BK20170615); and the Key Research and Development Plan (Modern Agriculture) of Jiangsu Province, China (No. BE2017373).

REFERENCES

Åkerström, A., Jaakola, L., Bång, U., and Jaderlund, A. (2010). Effects of latitude-related factors and geographical origin on anthocyanidin concentrations in fruits of *Vaccinium myrtillus* L. (bilberries). *J. Agric. Food Chem.* 58 (22), 11939–11945. doi: 10.1021/jf102407n

Almutairi, K. F., Bryla, D. R., and Strik, B. C. (2017). Potential of deficit irrigation, irrigation cutoffs, and crop thinning to maintain yield and fruit quality with less water in northern highbush blueberry. *HortScience* 52 (4), 625–633. doi: 10.21273/HORTSCI11533-16

Barrett, D. M., Beaulieu, J. C., and Shewfelt, R. (2010). Color, flavor, texture, and nutritional quality of fresh-cut fruits and vegetables: desirable levels, instrumental and sensory measurement, and the effects of processing. *Crit. Rev. Food Sci. Nutr.* 50 (5), 369–389. doi: 10.1080/10408391003626322

Bett-Garber, K. L., Lea, J. M., Watson, M. A., Grimm, C. C., Lloyd, S. W., Beaulieu, J. C., et al. (2015). Flavor of fresh blueberry juice and the comparison to amount of sugars, acids, anthocyanidins, and physicochemical measurements. *J. Food Sci.* 80 (4), 818–827. doi: 10.1111/1750-3841.12821

Brazelton, C., Young, K., and Bauer, N. (2017). Global blueberry statistics and intelligence report. *Int. Blueberry Organ.* 2017, 84.

Buratti, S., Ballabio, D., Benedetti, S., and Cosio, M. S. (2007). Prediction of Italian red wine sensorial descriptors from electronic nose, electronic tongue and spectrophotometric measurements by means of Genetic Algorithm regression models. *Food Chem.* 100 (1), 211–218. doi: 10.1016/j.foodchem.2005.09.040

Carbó, N., López Carrero, J., Garcia-Castillo, F., Tormos, I., Olivas, E., Folch, E., et al. (2018). Quantitative determination of spring water quality parameters via electronic tongue. *Sensors* 18 (1):40. doi: 10.3390/s18010040

Cardeñosa, V., Girones-Vilaplana, A., Muriel, J. L., Moreno, D. A., and Moreno-Rojas, J. M. (2016). Influence of genotype, cultivation system and irrigation regime on antioxidant capacity and selected phenolics of blueberries (*Vaccinium corymbosum* L.). *Food Chem.* 202, 276–283. doi: 10.1016/j.foodchem.2016.01.118

Chen, Q., Zhao, J., and Vittayapadung, S. (2008). Identification of the green tea grade level using electronic tongue and pattern recognition. *Food Res. Int.* 41 (5), 500–504. doi: 10.1016/j.foodres.2008.03.005

Connor, A. M., Luby, J. J., Tong, C. B., Finn, C. E., and Hancock, J. F. (2002). Genotypic and environmental variation in antioxidant activity, total phenolic content, and anthocyanin content among blueberry cultivars. *J. Am. Soc. Hortic. Sci.* 127 (1), 89–97. doi: 10.21273/JASHS.127.1.89

Correia, S., Gonçalves, B., Aires, A., Silva, A., Ferreira, L., Carvalho, R., et al. (2016). Effect of harvest year and altitude on nutritional and biometric characteristics of blueberry cultivars. *J. Chem.* 2016, ID 8648609. doi: 10.1155/2016/8648609

Drewnowski, A., and Gomez-Carneros, C. (2000). Bitter taste, phytonutrients, and the consumer: a review. *Am. J. Clin. Nutr.* 72 (6), 1424–1435. doi: 10.1093/ajcn/72.6.1424

Dudareva, L., Tarasenko, V., and Rudikovskaya, E. (2020). Involvement of photoprotective compounds of a phenolic nature in the response of *Arabidopsis thaliana* leaf tissues to low-intensity laser radiation. *Photochem. Photobiol.* doi: 10.1111/php.13289

Ehret, D. L., Frey, B., Forge, T., Helmer, T., Bryla, D. R., and Zebarth, B. J. (2014). Effects of nitrogen rate and application method on early production and fruit quality in highbush blueberry. *Can. J. Plant Sci.* 94 (7), 1165–1179. doi: 10.4141/cjps2013-401

Fawole, O. A., and Opara, U. L. (2013). Effects of maturity status on biochemical content, polyphenol composition and antioxidant capacity of pomegranate fruit arils (cv.'Bhagwa'). *South Afr. J. Bot.* 85, 23–31. doi: 10.1016/j.sajb.2012.11.010

Gallardo, R. K., Zhang, Q., Dossett, M., Polashock, J. J., Rodriguez-Saona, C., Vorsa, N., et al. (2018). Breeding trait priorities of the blueberry industry in the United States and Canada. *HortScience* 53 (7), 1021–1028. doi: 10.21273/HORTSCI12964-18

Gilbert, J. L., Olmstead, J. W., Colquhoun, T. A., Levin, L. A., Clark, D. G., and Moskowitz, H. R. (2014). Consumer-assisted selection of blueberry fruit quality traits. *HortScience* 49 (7), 864–873. doi: 10.21273/HORTSCI.49.7.864

Gilbert, J. L., Guthart, M. J., Gezan, S. A., de Carvalho, M. P., Schwieterman, M. L., Colquhoun, T. A., et al. (2015). Identifying breeding priorities for blueberry flavor using biochemical, sensory, and genotype by environment analyses. *PloS One* 10 (9), e0138494. doi: 10.1371/journal.pone.0138494

Hanson, E. J., Beggs, J. L., and Beaudry, R. M. (1993). Applying calcium chloride postharvest to improve highbush blueberry firmness. *HortScience* 28 (10), 1033–1034. doi: 10.21273/HORTSCI.28.10.1033

Huang, W. Y., Zhang, H. C., Liu, W. X., and Li, C. Y. (2012). Survey of antioxidant capacity and phenolic composition of blueberry, blackberry, and strawberry in Nanjing. *J. Zhejiang Univ. Sci. B* 13 (2), 94–102. doi: 10.1631/jzus.B1100137

Jaakola, L., Määttä-Riihinen, K., Kärenlampi, S., and Hohtola, A. (2004). Activation of flavonoid biosynthesis by solar radiation in bilberry (*Vaccinium myrtillus* L.) leaves. *Planta* 218 (5), 721–728. doi: 10.1007/s00425-003-1161-x

Kobayashi, Y., Habara, M., Ikezazki, H., Chen, Y., Naito, Y., and Toko, K. (2010). Advanced taste sensors based on artificial lipids with global selectivity to basic taste qualities and high correlation to sensory scores. *Sensors* 10 (4), 3411–3443. doi: 10.3390/s100403411

Kraujalytė, V., Venskutonis, P. R., Pukalskas, A., Česonienė, L., and Daubaras, R. (2015). Antioxidant properties, phenolic composition and potentiometric sensor array evaluation of commercial and new blueberry (*Vaccinium corymbosum*) and bog blueberry (*Vaccinium uliginosum*) genotypes. *Food Chem.* 188, 583–590. doi: 10.1016/j.foodchem.2015.05.031

Li, C., Feng, J., Huang, W. Y., and An, X. T. (2013). Composition of polyphenols and antioxidant activity of rabbiteye blueberry (*Vaccinium ashei*) in Nanjing. *J. Agric. Food Chem.* 61 (3), 523–531. doi: 10.1021/jf3046158

Li, Y., Sun, H., and Chen, L. (2016). The blueberry industry of China: the past 10 years and the future. *XI Int. Vaccinium Symposium* 1180, 531–536. doi: 10.17660/ActaHortic.2017.1180.75

Liang, Y., Lu, J., Zhang, L., Wu, S., and Wu, Y. (2003). Estimation of black tea quality by analysis of chemical composition and colour difference of tea infusions. *Food Chem.* 80 (2), 283–290. doi: 10.1016/S0308-8146(02)00415-6

Liu, D., Li, S., Wang, N., Deng, Y., Sha, L., Gai, S., et al. (2017). Evolution of taste compounds of Dezhou-braised chicken during cooking evaluated by chemical analysis and an electronic tongue system. *J. Food Sci.* 82 (5), 1076–1082. doi: 10.1111/1750-3841.13693

MacKenzie, S. J., Chandler, C. K., Hasing, T., and Whitaker, V. M. (2011). The role of temperature in the late-season decline in soluble solids content of strawberry

fruit in a subtropical production system. *HortScience* 46 (11), 1562–1566. doi: 10.21273/HORTSCI.46.11.1562

Martz, F., Jaakola, L., Julkunen-Tiitto, R., and Stark, S. (2010). Phenolic composition and antioxidant capacity of bilberry (*Vaccinium myrtillus*) leaves in Northern Europe following foliar development and along environmental gradients. *J. Chem. Ecol.* 36 (9), 1017–1028. doi: 10.1007/s10886-010-9836-9

Matsumoto, K., Kim, B., Oahn, V. T. K., Seo, J., Yoon, H., Park, M., et al. (2007). Comparison of sugar compositions and quality parameters during berry ripening between grape cultivars. *Korean J. Hortic. Sci. Technol.* 25 (3), 230–234.

Mikulic-Petkovsek, M., Schmitzer, V., Slatnar, A., Stampar, F., and Veberic, R. (2015). A comparison of fruit quality parameters of wild bilberry (*Vaccinium myrtillus* L.) growing at different locations. *J. Sci. Food Agric.* 95 (4), 776–785. doi: 10.1002/jsfa.6897

Moreno, M. A., Castell-Perez, M. E., Gomes, C., Da Silva, P. F., and Moreira, R. G. (2007). Quality of electron beam irradiation of blueberries (*Vaccinium corymbosum* L.) at medium dose levels (1.0–3.2 kGy). *LWT-Food Sci. Technol.* 40 (7), 1123–1132. doi: 10.1016/j.lwt.2006.08.012

Mphahlele, R. R., Caleb, O. J., Fawole, O. A., and Opara, U. L. (2016). Effects of different maturity stages and growing locations on changes in chemical, biochemical and aroma volatile composition of 'Wonderful' pomegranate juice. *J. Sci. Food Agric.* 96 (3), 1002–1009. doi: 10.1002/jsfa.7186

Naryal, A., Acharya, S., Bhardwaj, A. K., Kant, A., Chaurasia, O. P., and Stobdan, T. (2019). Altitudinal effect on sugar contents and sugar profiles in dried apricot (*Prunus armeniaca L.*) fruit. *J. Food Compos. Anal.* 76, 27–32. doi: 10.1016/j.jfca.2018.11.003

Phat, C., Moon, B., and Lee, C. (2016). Evaluation of umami taste in mushroom extracts by chemical analysis, sensory evaluation, and an electronic tongue system. *Food Chem.* 192, 1068–1077. doi: 10.1016/j.foodchem.2015.07.113

Price, M. L., Van Scoyoc, S., and Butler, L. G. (1978). A critical evaluation of the vanillin reaction as an assay for tannin in sorghum grain. *J. Agric. Food Chem.* 26 (5), 1214–1218. doi: 10.1021/jf60219a031

Qiu, S., Wang, J., and Gao, L. (2015). Qualification and quantisation of processed strawberry juice based on electronic nose and tongue. *LWT-Food Sci. Technol.* 60 (1), 115–123. doi: 10.1016/j.lwt.2014.08.041

Ribera, A. E., Reyes-Diaz, M., Alberdi, M., Zuñiga, G. E., and Mora, M. L. (2010). Antioxidant compounds in skin and pulp of fruits change among genotypes and maturity stages in highbush blueberry (*Vaccinium corymbosum* L.) grown in southern Chile. *J. Soil Sci. Plant Nutr.* 10 (4), 509–536. doi: 10.4067/S0718-95162010000200010

Rokaya, P. R., Baral, D. R., Gautam, D. M., Shrestha, A. K., and Paudyal, K. P. (2016). Effect of altitude and maturity stages on quality attributes of mandarin (*Citrus reticulata* Blanco). *Am. J. Plant Sci.* 7 (06), 958–966. doi: 10.4236/ajps.2016.76091

Saftner, R., Polashock, J., Ehlenfeldt, M., and Vinyard, B. (2008). Instrumental and sensory quality characteristics of blueberry fruit from twelve cultivars. *Postharvest Biol. Technol.* 49 (1), 19–26. doi: 10.1016/j.postharvbio.2008.01.008

Scalzo, J., Stevenson, D., and Hedderley, D. (2013). Blueberry estimated harvest from seven new cultivars: fruit and anthocyanins. *Food Chem.* 139, 44–50. doi: 10.1016/j.foodchem.2013.01.091

Trigo, M. J., Sousa, M. B., Sapata, M. M., Ferreira, A., Curado, T., Andrada, L., et al. (2004). Quality of gamma irradiated blueberries. *VIII Int. Symposium Vaccinium Culture* 715, 573–578. doi: 10.17660/ActaHortic.2006.715.88

Upadhaya, S., and Dwivedi, P. (2019). The role and potential of blueberry in increasing deforestation in southern Georgia, United States. *Agric. Syst.* 173, 39–48. doi: 10.1016/j.agsy.2019.01.002

Wang, L. J., Wu, J., Wang, H. X., Li, S. S., Zheng, X. C., Du, H., et al. (2015). Composition of phenolic compounds and antioxidant activity in the leaves of blueberry cultivars. *J. Funct. Foods* 16, 295–304. doi: 10.1016/j.jff.2015.04.027

Wei, Z., and Wang, J. (2013). The evaluation of sugar content and firmness of non-climacteric pears based on voltammetric electronic tongue. *J. Food Eng.* 117 (1), 158–164. doi: 10.1016/j.jfoodeng.2013.02.007

Yue, C., and Wang, J. (2016). Consumer preferences for fresh blueberry attributes. *XI Int. Vaccinium Symposium* 1180, 1–8. doi: 10.17660/ActaHortic.2017.1180.1

Zoratti, L., Jaakola, L., Häggman, H., and Giongo, L. (2015). Modification of sunlight radiation through colored photo-selective nets affects anthocyanin profile in *Vaccinium* spp. berries. *PloS One* 10 (8), e0135935. doi: 10.1371/journal.pone.0135935

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.

frontiers in Pharmacology

REVIEW
published: 29 September 2020
doi: 10.3389/fphar.2020.01265

Therapeutic Potential of Hydroxysafflor Yellow A on Cardio-Cerebrovascular Diseases

Xue Bai, Wen-Xiao Wang, Rui-Jia Fu, Shi-Jun Yue, Huan Gao, Yan-Yan Chen and Yu-Ping Tang**

Key Laboratory of Shaanxi Administration of Traditional Chinese Medicine for TCM Compatibility, and State Key Laboratory of Research & Development of Characteristic Qin Medicine Resources (Cultivation), and Shaanxi Key Laboratory of Chinese Medicine Fundamentals and New Drugs Research, and Shaanxi Collaborative Innovation Center of Chinese Medicinal Resources Industrialization, Shaanxi University of Chinese Medicine, Xi'an, China

The incidence rate of cardio-cerebrovascular diseases (CCVDs) is increasing worldwide, causing an increasingly serious public health burden. The pursuit of new promising treatment options is thus becoming a pressing issue. Hydroxysafflor yellow A (HSYA) is one of the main active quinochalcone *C*-glycosides in the florets of *Carthamus tinctorius* L., a medical and edible dual-purpose plant. HSYA has attracted much interest for its pharmacological actions in treating and/or managing CCVDs, such as myocardial and cerebral ischemia, hypertension, atherosclerosis, vascular dementia, and traumatic brain injury, in massive preclinical studies. In this review, we briefly summarized the mode and mechanism of action of HSYA on CCVDs based on these preclinical studies. The therapeutic effects of HSYA against CCVDs were presumed to reside mostly in its antioxidant, anti-inflammatory, and neuroprotective roles by acting on complex signaling pathways.

Keywords: hydroxysafflor yellow A, quinochalcone *C*-glycoside, cardio-cerebrovascular diseases, Carthami Flos, ischemia

OPEN ACCESS

Edited by:
M. Carmen González-Mas, University of Valencia, Spain

Reviewed by:
Li Yu, Zhejiang Chinese Medical University, China
Ya-nan Yang, Chinese Academy of Medical Sciences and Peking Union Medical College, China

****Correspondence:***
Shi-Jun Yue
shijun_yue@163.com
Yu-Ping Tang
yupingtang@sntcm.edu.cn

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: *06 June 2020*
Accepted: *30 July 2020*
Published: *29 September 2020*

Citation:
Bai X, Wang W-X, Fu R-J, Yue S-J, Gao H, Chen Y-Y and Tang Y-P (2020) Therapeutic Potential of Hydroxysafflor Yellow A on Cardio-Cerebrovascular Diseases. Front. Pharmacol. 11:01265. doi: 10.3389/fphar.2020.01265

INTRODUCTION

Cardio-cerebrovascular diseases (CCVDs) are characterized by ischemic or hemorrhagic lesions of the heart, brain, and peripheral circulatory tissues (Liu Z. et al., 2018). It is the high incidence, recurrence, and disability rates of CCVDs that directly aggravate the global burden of public health and hinder socio-economic development (Minno et al., 2019). Although much progress has been

made in understanding the pathological mechanisms of CCVDs, there is still no effective therapy to prevent or stop the epidemic trend of CCVDs, resulting in the urgent need to identify novel therapeutic options (Kazantsev and Outeiro, 2010; Upadhyay, 2014).

Traditional Chinese medicine (TCM), a cost-effective and safe remedy, has been widely used in China and surrounding countries (including Japan and Korea) for the treatment and management of CCVDs with exact and prominent efficacy. *Carthamus tinctorius* L. (Compositae) (**Figure 1A**) seeds are known to be rich in α-linoleic acid and have been used since ancient times as a source of cooking oil. Meanwhile, its flowers are widely used for coloring and flavoring foods and manufacturing dyes (Hu et al., 2018; Guo et al., 2019). Notably, the medical use of Carthami Flos (**Figure 1B**, the dried florets of *C. tinctorius*) was first documented in the Golden Chamber Synopsis (Han Dynasty, ~2000 years ago) (Ma et al., 2014), and also described in the Compendium of Materia Medica (Ming Dynasty, ~500 years ago) as being able to "invigorate the blood circulation", suggestive of its potential uses against circulatory system diseases. In modern Chinese clinic, Honghua injection (made from the water extract of Carthami Flos) and Danhong injection (extracted and refined from *Salviae miltiorrhizae* Radix et Rhizoma and Carthami Flos herb pair) are widely used for the treatment of coronary heart disease, angina pectoris, myocardial infarction, ischemic encephalopathy, and cerebral thrombosis (Fan et al., 2019; Zhang et al., 2019).

The chemical constituents of Carthami Flos are plentiful, and include flavonoids (e.g. quinochalcone *C*-glycosides), alkaloids, phenolic acids, fatty acids, and more (Yue et al., 2013). Among them, hydroxysafflor yellow A (HSYA, **Figure 1C**) is both a representative water-soluble quinochalcone *C*-glycoside pigment and the quality marker of Carthami Flos. It produces remarkable pharmacological activities in CCVDs that have aroused great interest worldwide (Zhang et al., 2019), and massive preclinical studies have aimed to prove the pharmacological effects and dissect the mechanisms of actions of HSYA in treating CCVDs. Safflower yellow injection, a purified yellow pigment extract from Carthami Flos containing no less than 70% of HSYA, is commercially available for stable exertional angina pectoris of coronary heart disease with a marked curative effect in Chinese clinic (Liu et al., 2014; Xuan et al., 2018). Here, we briefly summarize the existing evidence to provide valuable references and implications for the clinical uses of HSYA.

Abbreviations: Akt, the protein kinase B; AMPK, 5'-monophosphate -activated protein kinase; Ang II, angiotensin II; AS, atherosclerosis; Bax, Bcl-2 associated X protein; BBB, blood-brain barrier; Bcl-2, B-cell lymphoma-2; BDNF, brain-derived neurotrophic factor; CCVDs, cardio-cerebrovascular diseases; CI, cerebral ischemia; CI/R, cerebral ischemia/reperfusion; CYP, cytochrome P450; eNOS, endothelial nitric oxide synthase; GSSG, oxidized glutathione; H/R, hypoxia/reoxygenation; HIF-1, hypoxia inducible factor-1; HO-1, heme oxygenase-1; HSYA, hydroxysafflor yellow A; HUVECs, human umbilical vein endothelial cells; IL, interleukin; JAK2, janus kinase 2/signal transducer; LDL, low-density lipoprotein; LPS, lipopolysaccharide; MCAO, middle cerebral artery occlusion; MDA, malondialdehyde; MI, myocardial ischemia; MI/R, myocardial ischemia/reperfusion; MMP, matrix metalloproteinase; mTOR, mammalian target of rapamycin; mtPTP, mitochondrial permeability transition pore; NF-κB, nuclear factor kappa beta; NLRP3, NOD-like receptor 3; NMDARs, NR2B-containing N-methyl-d-aspartate receptors; NO, nitric oxide; Nrf2, nuclear factor erythroid-2-related factor 2; OGD, oxygen-glucose deprivation; OGD/R, OGD/reoxygenation; ox-LDL, oxidized low-density lipoprotein; PAF, platelet activating factor; PI3K, phosphoinositide 3-kinase; RCT, randomized controlled clinical trial; ROS, reactive oxygen species; SCFA, short-chain fatty acid; SD, Sprague-Dawley; SOCS3, suppressor of cytokine signaling protein 3; SOD, superoxide dismutase; TBI, traumatic brain injury; TCM, traditional Chinese medicine; TLR, toll like receptor; TNF-α, tumor necrosis factor-α; VaD, vascular dementia; VEGF, vascular endothelial growth factor; VSMCs, vascular smooth muscle cells; WRP, washed rabbits platelet.

THERAPEUTIC POTENTIAL OF HSYA ON CCVDS

Effects on Myocardial Ischemia (MI)

It is acknowledged that MI results from insufficient blood-oxygen supply (Thiagarajan et al., 2017) and the improved vasomotor and circulatory functions exert beneficial effects on MI (Ribeiro et al., 2019). The vasoconstrictor endothelin and the vasodilator nitric oxide (NO) are known as two common regulators modulating vasomotor function. HSYA can reverse the circulating levels of both in acute MI animals (e.g., dogs and rats), thereby elevating myocardial blood-oxygen supply and reducing myocardial injury and apoptosis (Li et al., 2006; Wang et al., 2007). Other vasomotor function-related factors, such as 6-keto-prostaglandin F1α, thromboxane B2, and angiotensin II (Ang II), were also of great importance for HSYA (Wang et al., 2007). Angiogenesis participates in the circulatory function recovery from MI, and HSYA exerts the pro-angiogenic effects in two main ways: (1) nucleolin-mediated post-transcriptional regulation of vascular endothelial growth factor-A (VEGF-A) and matrix metalloproteinase (MMP) -9 expressions (Zou et al., 2018); and (2) the up-regulation of heme oxygenase-1 (HO-1)/VEGF-A/stromal cell-derived factor-1α cascade (Wei et al., 2017).

A specialized piece of *in vivo* research demonstrated that the antioxidant effect of HSYA was involved in the prevention of Ang II-induced myocardial hypertrophy (a compensatory response to MI), which may act through the activation of the nuclear factor erythroid-2-related factor 2 (Nrf2)/NAD(P)H: quinone oxidoreductase 1/HO-1 signaling pathway (Ni et al., 2018). Nrf-2, as the main regulator of the antioxidant system present on the cardiovascular system, is becoming a very promising pharmacological target for cardiovascular diseases (McSweeney et al., 2016). Our research group has found that HSYA possessed significant antioxidant activity *in vitro* (Yue et al., 2014). Thus, the antioxidant effect of HSYA may be essential to improve the outcomes of cardiovascular diseases.

Effects on Myocardial Ischemia/Reperfusion (MI/R) Injury

Thrombolytic or percutaneous coronary intervention reperfusion for acute myocardial infarction is favorable in most cases, but can also cause MI/R injury, resulting in excessive pro-

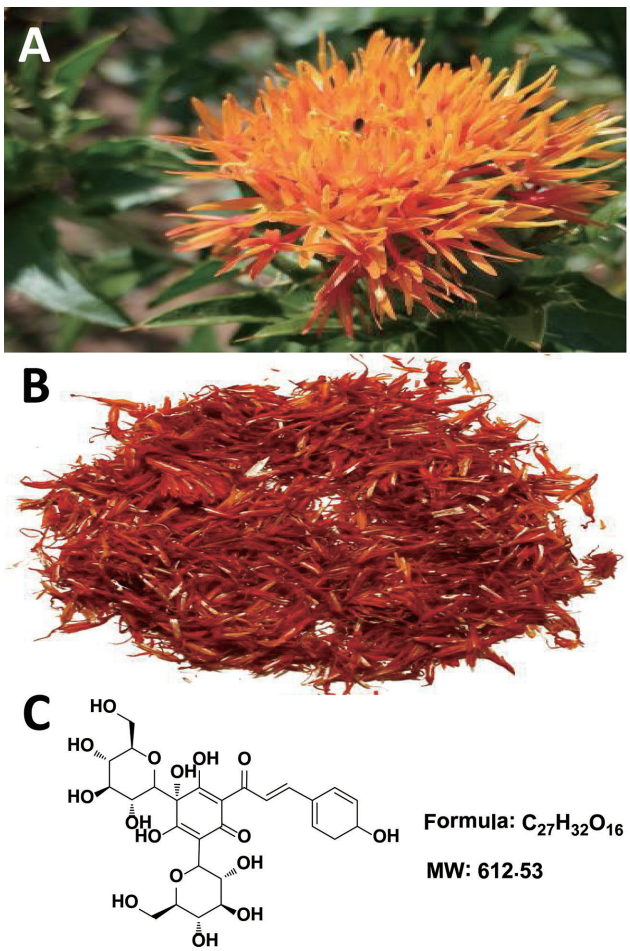

FIGURE 1 | *Carthamus tinctorius* L. **(A)**, Carthami Flos (the dried florets of *C. tinctorius*) **(B)**, and hydroxysafflor yellow A **(C)**.

inflammatory cytokines in myocardial tissue (Xie et al., 2018). HSYA has been reported to possess significant anti-inflammatory activity *in vitro* (Yue et al., 2016). In neonatal rat ventricular myocytes induced by hypoxia/reoxygenation (H/R) and lipopolysaccharide (LPS), HSYA not only inhibited the excessive secretion of pro-inflammatory cytokines but also suppressed the over-expression of toll like receptor (TLR) 4 and nuclear factor kappa beta (NF-κB). Importantly, HSYA was found to alleviate cardiac damage caused by MI/R in normal rats instead of TLR4-knockout mice (Han et al., 2016). It is important to note that TLR4 is the receptor of LPS, the component of the outer membrane of Gram-negative bacteria. There is evidence that NOD-like receptor 3 (NLRP3) inflammasome activation promotes myocardial injury and apoptosis *via* inducing the production of pro-inflammatory cytokines. HSYA improved H/R-induced H9c2 cell viability, maintained mitochondrial membrane potential, and inhibited NLRP3 inflammasome activation, while the adenosine 5'-monophosphate-activated protein kinase (AMPK) inhibitor partially abolished these observed effects of HSYA, which suggests that HSYA suppresses NLRP3 inflammasome activation *via* the AMPK pathway (Ye et al., 2020). Together, the TLR4 signaling pathway and NLRP3 inflammasome impact on the anti-inflammatory action of HSYA in MI/R.

Apoptosis is initiated shortly after the onset of myocardial infarction and is enhanced markedly during reperfusion. In H/R-induced H9c2 cells, the anti-apoptotic effect of HSYA not only depends on the up-regulation of HO-1 expression through the phosphoinositide 3-kinase (PI3K)/the protein kinase B (Akt)/Nrf2 signaling pathway, a compensatory mechanism limiting the apoptotic events in the presence of aggressive factors (Liu et al., 2012), but also targets the Akt/hexokinase II pathway to activate the hexokinase II protein and restore mitochondrial energy to reduce intracellular reactive oxygen species (ROS) generation (Min and Wei, 2017). Additionally, Zhou et al. reported that the anti-apoptotic effect of HSYA might be largely dependent on the Janus kinase 2/signal transducer (JAK2) and activation of the transcription 1 pathway (Zhou et al., 2019).

Beside inflammation and apoptosis, MI/R damages cardiomyocytes in part *via* the opening of the mitochondrial permeability transition pore (mtPTP), a non-selective pore that penetrates the inner and outer mitochondrial membranes (Bhosale and Duchen, 2019). HSYA has the capability to enter the cardiomyocytes and then inhibit mtPTP opening to alleviate H/R-induced myocardial injury through the enhanced endothelial nitric oxide synthase (eNOS)-produced NO (Liu et al., 2008; Huber et al., 2018).

Effects on Hypertension

Hypertension is a major global health challenge and an important risk factor of CCVDs. The blood pressure control rate of hypertensive patients in developing countries remains at unacceptably low levels (Mills et al., 2016). There is evidence that the conspicuous antihypertensive effect of HSYA is attributed to the inhibition of voltage-gated channels, the renin-angiotensin-aldosterone system, and the sympathetic nervous system. Specifically, HSYA inhibited the endotehlin-independent contraction of the thoracic aorta rings of rats through the blockade of inositol 1,4,5-triphosphate receptor in vascular smooth muscle cells (VSMCs), leading to the decrease of extracellular Ca^{2+} influx (Zhang et al., 2011). Beside VSMCs, endothelial cells participate in vasoconstriction and relaxation. Yang et al. found that oral HSYA has a concentration-dependent antihypertensive effect. It reversed the constriction of mesenteric arteries induced by a thromboxane A2 mimetic agent, the potential mechanism of which might be associated with the TRPV4 channel-dependent Ca^{2+} influx, protein kinase A-dependent eNOS phosphorylation, and NO production (Yang et al., 2020). A further study disclosed that HSYA could normalize blood pressure and heart rate dose-dependently in spontaneously hypertensive rats, which might be related to activating K_{ATP} and BK_{Ca} channels, inhibiting L-type Ca channels, decreasing Ca^{2+} influx, and subsequently inhibiting cardiac contractility (Nie et al., 2012; Wang et al., 2020). However, HSYA reduced blood pressure and heart rate in

normotensive rats, on which more focus needs to be placed in the future. Moreover, HSYA can increase the reduced diastolic response of the thoracic aortic to acetylcholine and sodium nitroprusside, and thus attenuate the vascular contractile effect of phenylephrine (Jin et al., 2013). HSYA can also inhibit proliferative activity and collagen synthesis of Ang II-induced adventitial fibroblasts, reduce the expressions of MMP-1, transforming growth factor-β1, α-smooth muscle actin, and NF-κB p65, and thereby decreasing vascular adventitia proliferation and hyperplasia during vascular remodeling (Yuan et al., 2014). Obviously, it could be drawn that HSYA might be potentially useful as an antihypertensive drug *via* multiple mechanisms mainly involved in both cardiac output and peripheral resistance.

Hypertension may cause ventricular hypertrophy, which produces mechanical stimulation to the heart, further causing arrhythmia, heart failure, coronary occlusion, and sudden death (Pearson et al., 1991). A study by Wang et al. indicated that oral HSYA exhibited anti-apoptotic effects on hypertensive ventricular hypertrophy in rats by increasing the B-cell lymphoma-2 (Bcl-2)/Bcl-2 associated X protein (Bax) ratio and blocking serum MMP-2 and MMP-9 levels (Wang et al., 2014). Moreover, pulmonary arterial hypertension is a common combination of congenital heart disease with systemic-to-pulmonary artery shunt diseases, leading to right ventricular heart failure and premature death (Huang et al., 2018). Voltage-gated K^+ channel (K_V channel) is an important channel for maintaining normal membrane potential and muscle tension of VSMCs, while HSYA could activate the Kv channel of pulmonary artery VSMCs to reduce vascular tension, suggesting that HSYA may be a potential medication for pulmonary arterial hypertension (Bai et al., 2012).

Effects on Atherosclerosis (AS)

AS could precipitate the onset of myocardial infarction and inflammation and is being increasingly recognized as the main pathogenic mechanism through the narrowing and blockage of arteries and the increased risk of blood vessel rupture (Kotla et al., 2017). HSYA could inhibit ROS-induced inflammation in THP-1 macrophages (Jiang et al., 2017), but could also suppress tumor necrosis factor-α (TNF-α)-induced inflammatory responses through inhibiting the TNF-α receptor type 1-mediated classical NF-κB pathway in arterial endothelial cells (Wang H. F. et al., 2016). Jiang et al. discovered that HSYA-mediated sonodynamic therapy induced an autophagic response to inhibit inflammation *via* the PI3K/Akt/mammalian target of rapamycin (mTOR) signaling pathway in THP-1 macrophages (Jiang et al., 2017). Nevertheless, it is very interesting to justify the potential biphasic effect of HSYA under excessive macrophage autophagy, which can drive the instability of atherosclerotic plaque (Petrovski et al., 2011).

In fact, the oxidation of low-density lipoprotein (LDL) and oxidized LDL (ox-LDL)-induced vascular damage are key events in early AS. Although its LDL-lowering effect remains unknown, HSYA was able to reduce the susceptibility of LDL to copper-induced lipid peroxidation *in vitro* (Bacchetti et al., 2020). It is interesting to identify the effect of HSYA against ox-LDL formation *in vivo*. In ox-LDL-induced foamy macrophages, HSYA displayed obvious repairing effects on the *de novo* fatty acid biosynthesis pathway, among which oleoyl-(acyl-carrier-protein) hydrolase was postulated to be a target of HSYA (Wei et al., 2018). HSYA also exerted protective effects against ox-LDL-induced VSMCs proliferation *via* increasing mitogen-activated protein kinase phospholipase-1 expression and the proportion of cells in the G0/G1 phase, followed by reducing p-extracellular regulated protein kinases activity (Sheng et al., 2012). Moreover, HSYA has been shown to significantly improve ox-LDL-induced human endothelial injury, partially *via* the anti-apoptotic effect of the mitochondrial membranous voltage-dependent anion-selective channel protein 2 (Ye et al., 2017). Recently, Miao et al. revealed that HSYA could inhibit the high ox-LDL-induced human coronary artery endothelial cell injury, possibly *via* increasing eNOS expression and NO release, while inhibiting LDL receptor-1 expression and lactate dehydrogenase release (Miao et al., 2019).

During the development of AS, platelets can accelerate activation and release a variety of active substances, such as platelet-activating factor (PAF) and thromboxane B2, conversely promoting platelet adhesion and aggregation, and even damaging vascular endothelial cells (Wang et al., 2008). HSYA was able to inhibit PAF-induced platelet aggregation in rabbits by blocking PAF-mediated washed rabbit platelets (WRPs) and polymorphonuclear leukocyte aggregation (Zang et al., 2002).

Collectively, the above *in vitro* studies have manifested the potential anti-AS effects of HSYA, which should go through additional *in vivo* studies to determine its clinical implications.

Effects on Vascular Injury and Remodeling Diseases

The vascular endothelium plays an important role in modulating numerous aspects of vascular homeostasis (Scarabelli et al., 2002). HSYA was capable of promoting the survival and proliferation of vascular endothelial cells under both normoxic and hypoxic conditions, and its effect was stronger under hypoxia *via* up-regulating the Bcl-2/Bax ratio and accumulating hypoxia-inducible factor-1 (HIF-1) α, which was related to VEGF and its receptor system (Song et al., 2005; Ji et al., 2008). Also, HSYA could protect human umbilical vein endothelial cells (HUVECs) from hypoxia-induced injury by reducing p53 expression in the cell nucleus and up-regulating eNOS expression to produce NO in cell supernatant (Ji et al., 2009).

Abnormal proliferation of VSMCs is a crucial cytopathological basis for the development and progression of vascular remodeling diseases (Ivey et al., 2008). HSYA could inhibit platelet-derived growth factor-BB induced VSMCs proliferation by decreasing proliferating cell nuclear antigen expression and blocking mitogen-activated protein kinase/extracellular regulated protein kinases and Akt signaling pathways (Song et al., 2014; Zhao et al., 2015). In the LPS-induced VSMCs proliferation and migration model, HSYA inhibits the up-regulation of TLR4 expression and the activation of Ras-related C3 botulinum toxin substrate 1/Akt pathway (Yang et al., 2015).

Effects on Cerebral Ischemia (CI)

CI is one of the leading causes of death worldwide, and patients who survive CI often experience paralysis, impaired speech, or loss of vision (Moskowitz et al., 2010). HSYA appears to treat focal CI injury in rats through its anti-coagulation effects on thrombosis formation and platelet aggregation, as well as beneficial regulation on prostacyclin/thromboxane and blood rheological changes (Zhu et al., 2005). HSYA could also preserve cortex mitochondrial function of CI rats *via* scavenging free radicals, reducing lipid peroxides, and antagonizing Ca^{2+} (Tian et al., 2004). Importantly, HSYA possessed a better effect on cerebrovascular vasodilatation than on cardiovascular vasodilatation (Sun Y. et al., 2018), but the differential molecular mechanism remains to be discovered.

The blood-brain barrier (BBB) essentially maintains a stable cerebral homeostasis, while the destruction or increased permeability of BBB are common pathological processes during many serious cerebrovascular diseases (Chen Z. X. et al., 2019). A study by Lv et al. revealed that HSYA significantly attenuated BBB dysfunction in anti-inflammatory patterns in ischemia stroke *via* the tight junction pathway, especially the NMMHC IIA, TLR4/PI3K/Akt/Jun N-terminal kinase 1/2/14-3-3ϵ pathway while inhibiting the expressions of occludin, claudin-5, and zonula occludens-1 (Lv and Fu, 2018). Since MMPs are the main endoproteinases involved in BBB destruction (Romanic et al., 1998), the prominent inhibitory effects of HSYA on MMP-2 and MMP-9 mentioned in cardiovascular diseases may also contribute to BBB improvement.

It is worth mentioning that CI plays a causal role in facilitating neuronal death (Martin and Wang, 2010). A metabonomic study revealed that HSYA could attenuate excitatory amino acid-induced neurotoxicity, at least partially, through inhibiting the NF-κB pathway in the cerebral tissues of the middle cerebral artery occlusion (MCAO) model rats (Liu et al., 2013). Other protective mechanisms of HSYA against excitotoxic neuronal death include the inhibition of NR2B-containing N-methyl-d-aspartate receptors (NMDARs) expression and the Bcl-2 family regulation in cortical cultures, and the inhibition of the N-methyl-d-aspartate-induced and NMDARs-mediated intracellular Ca^{2+} increase in hippocampal cultures (Yang et al., 2010; Wang X. T. et al., 2016). In oxygen-glucose deprivation (OGD)-induced BV2 microglia, the neuroprotective action of HSYA involves the decreased expressions of pro-inflammatory cytokines, including interleukin (IL) -1β, TNF-α, inducible nitric oxide synthase, cyclooxygenase-2, and monocyte chemotactic protein-1, as well as the reserved phosphorylation of p38 and nuclear translocation of p65 (Li et al., 2013). Peroxynitrite-mediated protein tyrosine nitration and nitrosative stress represent the crucial pathogenic mechanisms of CI, while the anti-nitrative pathway might contribute to the neuroprotective efficacy of HSYA. Specifically, Sun et al. discovered that HSYA blocked authentic peroxynitrite-induced tyrosine nitration in primary cortical neurons by the reduction of inducible nitric oxide synthase expression and NO content, suggestive of its peroxynitrite scavenging abilities (Sun et al., 2013). They further reported that HSYA prevented peroxisome proliferator-activated receptor γ nitrative modification in primary neurons and resumed eroxisome proliferator-activated receptor γ activity stimulated by either 15-deoxy-delta prostaglandin J2 or rosiglitazone (Sun L. et al., 2018).

Effects on Cerebral Ischemia/Reperfusion (CI/R) Injury

A growing body of research has evidenced that oxidative stress is implicated in the pathogenesis of CI/R injury. Wei et al. showed that HSYA might oppose CI/R injury of MCAO rats through attenuating the elevation of malondialdehyde (MDA) level and decreasing superoxide dismutase (SOD) activity in the ipsilateral hemisphere and serum (Wei et al., 2005). HSYA could also reduce CI/R-induced protein oxidation and nitration, attenuate BBB destruction, and importantly inhibit the up-regulation of 12/15-lipoxygenase, which is implicated in the oxidative stress of CI/R (Sun et al., 2012). In an *in vitro* assay, HSYA was shown to block OGD/reoxygenation (OGD/R)-induced PC12 cells apoptosis through the suppression of intracellular oxidative stress (Fan et al., 2011).

An inflammatory reaction is a recognized player in CI/R damage. Through suppressing TLR4 pathway-mediated signaling responses, HSYA could up-regulate brain-derived neurotrophic factor (BDNF) in MCAO mice at post-ischemia/reperfusion (Lv et al., 2015), but also exert neurotrophic and anti-inflammatory functions in LPS-activated co-existence systems for microglia and neurons (Lv et al., 2016). In the microglia of the ischemic cortex after acute CI/R, HSYA exerted anti-inflammatory effects by activating the TLR9 signaling pathway and suppressing the NF-κB pathway (Gong et al., 2018). Further studies demonstrated that HSYA significantly inhibited NF-κB p65 nuclear translation and p65 binding activity, both mRNA and protein levels of intercellular adhesion molecule 1, and the infiltration of neutrophils (Sun et al., 2010).

Cognitive impairment is becoming a serious mental deficit that severely affects the life quality of patients following CI/R (Jokinen et al., 2006). HSYA has the capacity to improve neurological deficit scores and increase the surviving hippocampal CA1 pyramidal cells in focal CI/R rats (Sun et al., 2010). HSYA injected *via* the common carotid artery significantly rescued the neurological and cognitive functional deficits of MCAO rats against CI/R injury. Meanwhile, HSYA could markedly down-regulate JAK2-mediated signaling, while promoting the expression of the suppressor of cytokine signaling protein 3 (SOCS3) (Yu et al., 2018; Yu et al., 2020). Furthermore, the neuroprotective effect of HSYA against CI/R injury might be conferred through activating the Akt-dependent autophagy pathway (Qi et al., 2014). In both OGD/R-induced primary mouse neurons and PC12 cells, HSYA inhibited phenylalanine biosynthesis to enhance mitochondrial function and biogenesis for neuroprotection (Chen S. N. et al., 2019).

PI3K-mediated signaling pathways are also involved in the protective effects of HSYA against apoptosis and autophagy during CI/R. Chen et al. reported that HSYA critically reduced CI/R-mediated apoptosis through the PI3K/Akt/glycogen synthase kinase 3β signaling pathway (Chen et al., 2013).

HSYA protected the cerebral microvascular endothelial cells against OGD/R-induced injury by inhibiting autophagy *via* the Class I PI3K/Akt/mTOR signaling pathway (Yang et al., 2018). Proteomic analysis showed that mTOR, Eftud2, Rab11, Ppp2r5e, and HIF-1 signaling pathways were the key hub proteins and pathways in HSYA against CI/R injury (Xu et al., 2019). Therefore, the PI3K/Akt/mTOR signaling pathway needs to be further studied to clarify the mechanisms of actions of HSYA against CI/R injury.

Similar to MI/R, CI/R also results in mtPTP opening. Mechanically, HSYA could inhibit mtPTP opening by inhibiting Ca^{2+}-induced ROS generation and H_2O_2-induced swelling of mitochondria isolated from rat brains, improving mitochondrial energy metabolism and enhancing ATP levels and the respiratory control ratio in the ischemia brain (Tian et al., 2008).

Effects on Vascular Dementia (VaD)

VaD is characterized by pathological damage and a decline in intelligence resulting from hypoxic-ischemic or hemorrhagic brain injury (Zhao X. X. et al., 2018). As mentioned before, HSYA has the capacity to improve cognitive impairment. In VaD rats, Zhang et al. revealed that HSYA promoted angiogenesis and increased synaptic plasticity *via* up-regulating the hippocampal expressions of VEGF-A, NMDAR type-1, BDNF, and GluN2B (a subunit of NMDAR), thus improving spatial learning and memory (Zhang et al., 2014; Xing et al., 2016). Although no drug is approved, the above findings may shed light on the therapeutic potential of HSYA for managing the progress of VaD.

Effects on Traumatic Brain Injury (TBI)

TBI refers to the injury of cerebral tissue structure/function caused by various kinds of mechanical violence in the outside world (Xiu et al., 2018). HSYA has the potential to be utilized as a neuroprotective agent in cases of TBI. Firstly, TBI enables HSYA to distribute in the cerebral tissues of rats (Bie et al., 2010). Then, the antioxidant effect of HSYA in the brain of the TBI rats could explain the TBI improvement *via* increasing SOD, catalase and glutathione levels, while reducing MDA and oxidized glutathione (GSSG) levels (Bie et al., 2010; Wang Y. et al., 2016). Lastly, HSYA could increase mitochondrial ATPase (i.e., Na^+, K^+-ATPase, Ca^{2+}-ATPase, and Mg^{2+}-ATPase) and tissue plasminogen activator activities, while decreasing plasma plasminogen activator inhibitor-1 activity and MMP-9 expression in the hippocampus of TBI rats (Wang Y. et al., 2016).

In summary, the above findings buttress the assertion that HSYA exerts cardio-cerebrovascular protective activities through complex pathways and exhibits a definite curative effect in the application of CCVDs (**Table 1** and **Figure 2**).

CONCLUSIONS AND PROSPECTS

It is becoming clear that the important mechanisms by which HSYA exerts extensive biological activities in CCVDs are through its antioxidant, anti-inflammatory, and neuroprotective effects. And there is no doubt that HSYA is a promising lead drug candidate in designing new multi-targeted therapeutic agents against CCVDs. The other analogues of HSYA, safflor yellow A (Duan et al., 2013) and safflor yellow B (Wang et al., 2009; Wang et al., 2013), showed similar protective effects against CCVDs. Further structural modification of HSYA should be extensively made and coupled with quantitative structure-activity relationship studies to develop more selective and safe drugs.

The oral bioavailability of HSYA is extremely low (~1.2%) (Ekin, 2005) and oral administration of HSYA accounts for about 0.9% of all *in vivo* experiments from **Table 1**. However, among many administration routes, oral administration is of great significance in drug formation because of its convenience and safety. Thus, the microemulsion, nanoemulsion, and nanoparticles of HSYA have been developed to overcome the low oral bioavailability (Li et al., 2010; Qi et al., 2011; Lv et al., 2012; Shi et al., 2018; Zhao B. X. et al., 2018). Considering the weak ability of HSYA to penetrate the BBB in physiologic condition (He et al., 2008), Borneolum Syntheticum and Acori Talarinowii Rhizoma were used to enhance its BBB permeability (Wu et al., 2011). Nevertheless, more attention should be payed to the overall efficacy and safety evaluation of HSYA before and after improving oral bioavailability and BBB permeability.

In the past decade, cohesive evidence showed that gut microbiota may serve as a therapeutic target of natural compounds derived from TCM (Yue et al., 2019a; Yue et al., 2019b). Oral HSYA is mostly retained in the intestinal tract as its prototype, which inevitably interacts with gut microbiota. Obesity is considered to be one of the most important risk factors of CCVDs. Our research group reported that HSYA mediated its anti-obesity effects by reversing gut microbiota dysbiosis in obese mice, followed by increasing short-chain fatty acid (SCFA)-producing bacteria (Liu J. et al., 2018). For SCFAs, it has a well-established role in maintaining host immune function after MI (Tang et al., 2019). The potential for microbial synthesis of SCFAs, including propionate and butyrate, was low in patients with atherosclerotic cardiovascular disease (Jie et al., 2017). Recently, the microbiota-gut-brain axis has shown to influence BBB permeability and the pathological process of TBI (Braniste et al., 2014; Ma et al., 2017). Thus, it is feasible to unveil the underlying mechanisms of oral HSYA on CCVDs from the new perspective of gut microbiota modulation.

Carthami Flos is a common part of preparations used in TCM and other traditional medicinal systems. It is necessary to strengthen the compatibility research of HSYA and other TCM-derived components. As an example, HSYA and Danshensu synergistically enhanced the antioxidant defense system and anti-apoptotic effects on MI/R injury through the Akt/Nrf2/HO-1 signaling pathway (Hu et al., 2016). Their combination further achieved enhanced neuroprotective effects on CI/R injury by alleviating pro-inflammatory and oxidative stress reactions *via* the TLR4/NF-κB and Nrf2/HO-1 pathways (Xu et al., 2017). On the other hand, combinations with existing western medications may also provide new therapy options for CCVDs patients. For example, HSYA as an add-on therapy to acetylglutamine could synergistically modulate the neuronal

TABLE 1 | Summary of pharmacological effects and mechanisms of HSYA on cardio-cerebrovascular diseases.

Disease	Species/Strains	Effective dose/ concentration	Route	Mechanism of action	Reference
Myocardial ischemia (MI)	Acute MI model in mongrel dogs	14, 28 mg·kg^{-1}	i.v.	Inhibits endothelin release, increases myocardial blood flow, and improves the cardiac oxygen metabolism	(Li et al., 2006)
	Acute MI model in male Wistar rats	5, 10, 20 mg·kg^{-1}	i.v.	Increases the activity of serum NO synthase, the content of NO and 6-keto-prostaglandin F1α, and decreases the levels of creatine kinase-MB, lactate dehydrogenase, thromboxane B2 and Ang II	(Wang et al., 2007)
	Acute MI model in male C57 mice	25 mg·kg^{-1}	i.p.	Promotes the migration and tube formation of HUVECs, enhances the expressions of nucleolin, VEGF-A and MMP-9	(Zou et al., 2018)
	MI model in male C57BL/6 mice	15, 30, 60 mg·kg^{-1}	i.v.	Promotes endothelial progenitor cells function through the HO-1/VEGF-A/stromal cell-derived factor-1α signaling cascade	(Wei et al., 2017)
	MI model in male SD rats	2, 5 mg·kg^{-1}	i.p.	Activates the Nrf2/NAD(P)H:quinone oxidoreductase 1/HO-1 signaling pathway	(Ni et al., 2018)
	Ang II-induced H9c2 cells	80 μmol·L^{-1}	/	Increases the cell viability, however, reduces protein synthesis rate, mitigates cell surface area and decreases the expression of brain natriuretic factor and β-myosin heavy chain	(Ni et al., 2018)
Myocardial ischemia/ reperfusion (MI/ R) injury	MI/R model in male SD rats	5 mg·kg^{-1}	i.p.	Decreases JAK2/signal transducer and activator of transcription 1 activity, enhances antioxidant capacity and decreases apoptosis	(Zhou et al., 2019)
	Hyperlipidemia combined with MI/R model in male Wistar rats	8, 16, 32 mg·kg^{-1}	i.p.	Suppresses the over-expression of TLR4	(Han et al., 2016)
	H/R and LPS-induced neonatal rat ventricular myocytes	1, 3, 10 μmol·L^{-1}	/	Decreases excessive secretion of inflammatory cytokines, down-regulates over-expression of TLR4 and NF-κB	(Han et al., 2016)
	H/R-induced H9c2 cells	6.25, 12.5, 25 μmol·L^{-1}	/	Improves cardiomyocyte viability, maintains mitochondrial membrane potential, reduces apoptotic cardiomyocytes, decreases Caspase-3 activity, and inhibits NLRP3 inflammasome activation	(Ye et al., 2020)
	H/R-induced H9c2 cells	1.25, 5, 20 μmol·L^{-1}	/	Activates the hexokinase II proteins, restores mitochondrial energy, reduces ROS generation	(Min and Wei, 2017)
	H/R-induced H9c2 cells	20 μmol·L^{-1}	/	Inactivates the JAK2/signal transducer and activator of the transcription 1 pathway	(Zhou et al., 2019)
	H/R-induced H9c2 cells	5, 20, 80 μmol·L^{-1}	/	Up-regulates HO-1 expression through the PI3K/Akt/Nrf2 signaling pathway	(Liu et al., 2012)
	MI/R model in hearts isolated from male SD rats	50, 100, 200 μmol·L^{-1}	/	Enhances NO production by eNOS activation	(Liu et al., 2008)
	H/R-induced cardiomyocytes isolated from SD rat hearts	100, 200 μmol·L^{-1}	/	Modulates the reduction of viability and the loss of rod-shaped cells, and interacts with the mtPTP	(Huber et al., 2018)
Hypertension	Male spontaneous hypertension rat and normotensive Wistar-Kyoto rats	0.1-3 mg·kg^{-1}	i.v.	Reduces blood pressure and heart rate, activates BK_{Ca} and K_{ATP} channels	(Nie et al., 2012)
	Male spontaneous hypertension rat and normotensive Wistar-Kyoto rats	0.6-2.4 mg·kg^{-1}	i.v.	Activates BK_{Ca} channels, inhibits Ca-L channels, and reduces intracellular free Ca^{2+} level	(Wang et al., 2020)
	Ang II-induced vascular adventitial fibroblasts in male SD rats	0.5, 1 ml/kg	i.p.	Reduces matrix metallopeptidase-1, transforming growth factor-β1, α-smooth muscle actin and NF-κB p65 expression	(Yuan et al., 2014)
	AT1 receptor-induced hypertension in male Wistar rats	10 mg·kg^{-1}	i.g.	Reverses the vascular structure and function and improves plasma biochemical parameters	(Jin et al., 2013)
	Pressure overload-induced cardiac hypertrophy in male Wistar rats	20, 40 mg·kg^{-1}	i.g.	Increases the Bcl-2/Bax ratio, blocks the levels of MMP-2 and MMP-9 in serum	(Wang et al., 2014)
	PE-induced pulmonary artery rings of Wistar rats	0.01-10 μmol·L^{-1}	/	Activates the K_V channel in placental VSMCs and relaxes rat pulmonary artery	(Bai et al., 2012)
	KCl-precontracted thoracic aorta rings of male Wistar rats	EC_{50} = 18.2 μmol·L^{-1}/17.4 μmol·L^{-1}	/	Induces relaxation in endothelium-intact/endothelium-precontracted aortas precontracted by KCl	(Zhang et al., 2011)
	PE-precontracted thoracic aorta rings of male Wistar rats	EC_{50} = 18.7 μmol·L^{-1}/17.9 μmol·L^{-1}	/	Induces relaxation in endothelium-intact/endothelium-precontracted aortas precontracted by phenylephrine	(Zhang et al., 2011)
	$CaCl_2$-precontracted thoracic aorta rings of male Wistar rats	20 μmol·L^{-1}	/	Reduces Ca^{2+} influx and inhibits inositol 1,4,5-triphosphate receptor	(Zhang et al., 2011)
	U46619-induced mesenteric arteries of male Wistar rats	1-100 μmol·L^{-1}	/	Reverses the constriction, promotes Ca^{2+} influx, eNOS phosphorylation and NO production	(Yang et al., 2020)

(Continued)

TABLE 1 | Continued

Disease	Species/Strains	Effective dose/ concentration	Route	Mechanism of action	Reference
Atherosclerosis (AS)	Human THP-1 monocytes	400-800 μmol·L^{-1}	/	Induces an autophagic response *via* the PI3K/Akt/mTOR signaling pathway and inhibits inflammation by ROS	(Jiang et al., 2017)
	Human coronary artery endothelial cells injury model	200-1600 μmol·L^{-1}	/	Up-regulates the eNOS gene and protein expression, increases NO release, inhibits lactate dehydrogenase release and down-regulates LDL receptor 1 expression	(Miao et al., 2019)
	TNF-α-stimulated primary mouse kidney arterial endothelial cells and RAW264.7 macrophage cells	120 μmol·L^{-1}	/	Inhibits the TNF-α receptor type 1-mediated classical NF-κB pathway	(Wang H. F. et al., 2016)
	Ox-LDL-induced foamy macrophages	65.30 μmol·L^{-1}	/	Up-regulates the abnormal metabolism of C12:0, C14:0, C18:1	(Wei et al., 2018)
	Ox-LDL-induced human endothelial cells	1, 5, 25 μmol·L^{-1}	/	Inhibits cell apoptosis by voltage-dependent anion-selective channel protein 2	(Ye et al., 2017)
	Ox-LDL-induced VSMCs	10 μmol·L^{-1}	/	Increases mitogen-activated protein kinase phospholipase-1 expression and the proportion of cells in G0/G1 phase, reduces p-extracellular signal-regulated protein kinase 1/2 activity, and suppresses cell cycle	(Sheng et al., 2012)
	PAF-induced WRP suspension of male New Zealand white rabbits	250-1470 μmol·L^{-1}	/	Inhibits PAF binding to WRP receptors	(Zang et al., 2002)
	PAF-induced WRP suspension of male New Zealand white rabbits	IC_{50} = 990 μmol·L^{-1}	/	Reduces PAF-mediated WRP aggregation	(Zang et al., 2002)
	PAF-induced polymorphonuclear leukocytes suspension of male New Zealand white rabbits	IC_{50} = 700 μmol·L^{-1}	/	Reduces PAF-mediated polymorphonuclear leukocytes aggregation	(Zang et al., 2002)
Vascular Injury Diseases	Hypoxia-induced canine aortic endothelial cell	10, 100, 1000 μmol·L^{-1}	/	Promotes vein endothelial cells proliferation *via* promoting vascular endothelial growth factor and its receptor secretion	(Song et al., 2005)
	Hypoxia-induced HUVECs	1, 10, 100 μmol·L^{-1}	/	Up-regulates the Bcl-2/Bax ratio and promotes HIF-1α protein accumulation	(Ji et al., 2008)
	Hypoxia-induced HUVECs	1, 10, 100 μmol·L^{-1}	/	Inhibits cell apoptosis and cell cycle G1 arrest induced by hypoxia	(Ji et al., 2009)
Vascular Remodeling Diseases	Platelet-derived growth factor -induced VSMCs	1-60 μmol·L^{-1}	/	Reduces the expression of proliferating cell nuclear antigen, blocks signal transduction of mitogen-activated protein kinase/extracellular regulated protein kinases	(Zhao et al., 2015)
	Platelet-derived growth factor -induced VSMCs	20 μmol·L^{-1}	/	Suppresses Akt signaling activation	(Song et al., 2014)
	LPS-induced VSMCs	0.1-100 μmol·L^{-1}	/	Inhibits TLR4/Ras-related C3 botulinum toxin substrate 1/Akt pathway	(Yang et al., 2015)
Cerebral ischemia (CI)	MCAO model in male SD rats	10, 50 mg·kg^{-1}	i.v.	Corrects the impaired metabolic pathways, suppresses pro-inflammatory cytokine expression and p65 translocation and binding activity	(Liu et al., 2013)
	MCAO model in male SD rats	10, 20 mg·kg^{-1}	i.v.	Reduces lipid peroxides, inhibits Ca^{2+} overload, scavenges free radicals	(Tian et al., 2004)
	MCAO model in male Wistar-Kyoto rats	3, 6 mg·kg^{-1}	i.v.	Inhibits thrombosis formation and platelet aggregation, regulates prostacyclin/thromboxane and blood rheological changes	(Zhu et al., 2005)
	MCAO model in C57BL/6J mice	1, 2, 4 mg·kg^{-1}	i.p.	Attenuates BBB dysfunction *via* the tight junction pathway, and attenuates the expression of occludin, claudin-5, and zonula occludens-1	(Lv and Fu, 2018)
	Nitrified bovine serum albumin and primary cultured male SD rats cortical neurons exposed to peroxynitrite	10, 100, 1000 μmol·L^{-1}	/	Blocks authentic peroxynitrite-induced tyrosine nitration	(Sun et al., 2013)
	Peroxynitrite donor SIN-1-induced primary cultured male SD rats cortical neurons	10, 100, 1000 μmol·L^{-1}	/	Blocks peroxisome proliferator-activated receptor γ nitration	(Sun et al., 2018)
	Primary cultured SD rats cortical neurons exposed to NMDA	10 μmol·L^{-1}	/	Inhibits the expression NR2B-containing NMDA receptors and regulates Bcl-2 family	(Yang et al., 2010)
		10-80 μmol·L^{-1}	/		(Tian et al., 2008)

(Continued)

TABLE 1 | Continued

Disease	Species/Strains	Effective dose/ concentration	Route	Mechanism of action	Reference
	Ca^{2+} and H_2O_2-induced brain mitochondrial suspension of male SD rats			Inhibits Ca^{2+}-induced generation of ROS, and mtPTP opening by a free radical scavenging action	
	NMDA-induced and NMDA receptors-mediated C57BL/6 mice hippocampal neurons	IC_{50} = 17.60 μmol·L^{-1}	/	Protects hippocampal neurons from excitotoxic damage through the inhibition of NMDA receptors	(Wang X. T. et al., 2016)
	OGD-induced BV2 microglia	40-1280 μmol·L^{-1}	/	Suppresses inflammatory responses by inhibiting the NF-κB signaling pathway and phosphorylation of p38	(Li et al., 2013)
	Phenylephrine-precontracted Coronary artery and basilar artery of beagle dogs	5 mg into the 20 ml Krebs'-Henseleit buffer every 5 min for 4–5 times	/	Attenuates the contractile responsiveness	(Sun et al., 2018)
Cerebral ischemia/ reperfusion (CI/ R) injury	MCAO/R model in male SD rats	1, 5, 10 mg·kg^{-1}	i.v.	Reduces protein oxidation and nitration, inhibits the up-regulation of 12/15-lipoxygenase, and attenuates BBB breakdown	(Sun et al., 2012)
	MCAO/R model in male SD rats	8, 16 mg·kg^{-1}	CCAI	Protects cognitive function and synaptic plasticity	(Yu et al., 2018)
	MCAO/R model in male SD rats	8, 16 mg·kg^{-1}	CCAI	Downregulates the expression of JAK2-mediated signaling, while promotes the expression of SOCS3	(Yu et al., 2020)
	MCAO/R model in male SD rats	6 mg·kg^{-1}	i.p.	Regulates Eftud2, Rab11, Ppp2r5e, and HIF-1 signaling pathway	(Xu et al., 2019)
	Acute MCAO/R model in male SD rats	6 mg·kg^{-1}	i.p.	Activates TLR9 in the microglia of ischemic cortex and suppresses the NF-κB pathway	(Gong et al., 2018)
	Acute I/R stroke model in male SD rats	2 mg·kg^{-1}	i.v.	Activates the Akt autophagy pathway in penumbra tissue	(Qi et al., 2014)
	MCAO/R model in male Wistar rats	2, 4, 8 mg·kg^{-1}	i.v.	Attenuates the elevation of MDA content, the decrease in SOD activity, and the total antioxidative capability	(Wei et al., 2005)
	MCAO/R model in male Wistar rats	2, 4, 8 mg·kg^{-1}	i.v.	Suppresses thrombin generation and thrombin-induced inflammatory responses by reducing Ang II content	(Sun et al., 2010)
	MCAO/R model in male Wistar rats	4, 8 mg·kg^{-1}	i.v.	Reduces apoptosis *via* PI3K/Akt/glycogen synthase kinase 3β signaling pathway	(Chen et al., 2013)
	MCAO/R model in C57BL/6J mice	2 mg·kg^{-1}	i.v.	Inhibits TLR4 pathway-mediated signaling responses	(Lv et al., 2015)
	OGD/R induced human brain microvascular endothelial cells	10-80 μmol·L^{-1}	/	Inhibits autophagy *via* the Class I PI3K/Akt/mTOR signaling pathway	(Yang et al., 2018)
	LPS-stimulated non-contact transwell co-culture system comprised microglia and neurons	50, 100 μmol·L^{-1}	/	Exerts neurotrophic and anti-inflammatory functions in response to LPS stimulation by inhibiting TLR4 pathway-mediated signaling	(Lv et al., 2016)
	OGD-induced PC12 cells	10, 100 μmol·L^{-1}	/	Suppresses the intracellular oxidative stress and mitochondria-dependent caspase cascade	(Fan et al., 2011)
	OGD/Reduced primary neurons and PC12 cells	1, 10 μmol·L^{-1}	/	Reduces phenylalanine level, promotes mitochondrial function and biogenesis for neuroprotection	(Chen S. N. et al., 2019)
Vascular dementia (VaD)	VaD model in male SD rats	6 mg·kg^{-1}	i.v.	Promotes angiogenesis and increases synaptic plasticity	(Zhang et al., 2014)
	VaD model in male SD rats	6 mg·kg^{-1}	i.v.	Increases in the expression levels of BDNF and GluN2B	(Xing et al., 2016)
Traumatic brain injury (TBI)	TBI model in male SD rats	2, 4 mg·kg^{-1}	i.v.	Increases mitochondrial ATPase and tissue plasminogen activator activities, decreases plasma plasminogen activator inhibitor-1 activity and MMP-9 expression	(Bie et al., 2010)
	TBI model in male SD rats	10, 30 mg·kg^{-1}	i.g.	Enhances SOD and catalase activities, glutathione level and the glutathione/GSSG ratio while reduces MDA and GSSG levels	(Wang Y. et al., 2016)

/, no information in the original paper; i.v., intravenous injection; i.g., intragastrical administration; i.p., intraperitoneal injection; CCAI, common carotid artery injection.

apoptosis and inflammation process during CI/R (Deng et al., 2018). However, it is important to note that HSYA is able to inhibit cytochrome P450 (CYP) enzymes' (i.e. CYP1A2 and CYP2C11) activities but induces CYP3A1 activity (Xu et al., 2014). Hence, more detailed and advanced research should be done in the future to develop new compound formulas with HSYA, which may bring about important benefits for CCVDs patients and TCM modernization.

The randomized controlled clinical trial (RCT) is an essential step in confirming the efficacy and safety of drugs. In contrast with the large numbers of preclinical experiments, only a few completed RCTs of HSYA were reported, which were mainly reflected in evaluating the efficacy and safety of HSYA injection in the treatment of acute ischemic stroke with blood stasis syndrome (Qin et al., 2016; Hu et al., 2020), followed by a currently ongoing phase III RCT (No. CTR20150839, http://www.chinadrugtrials.org.cn/).

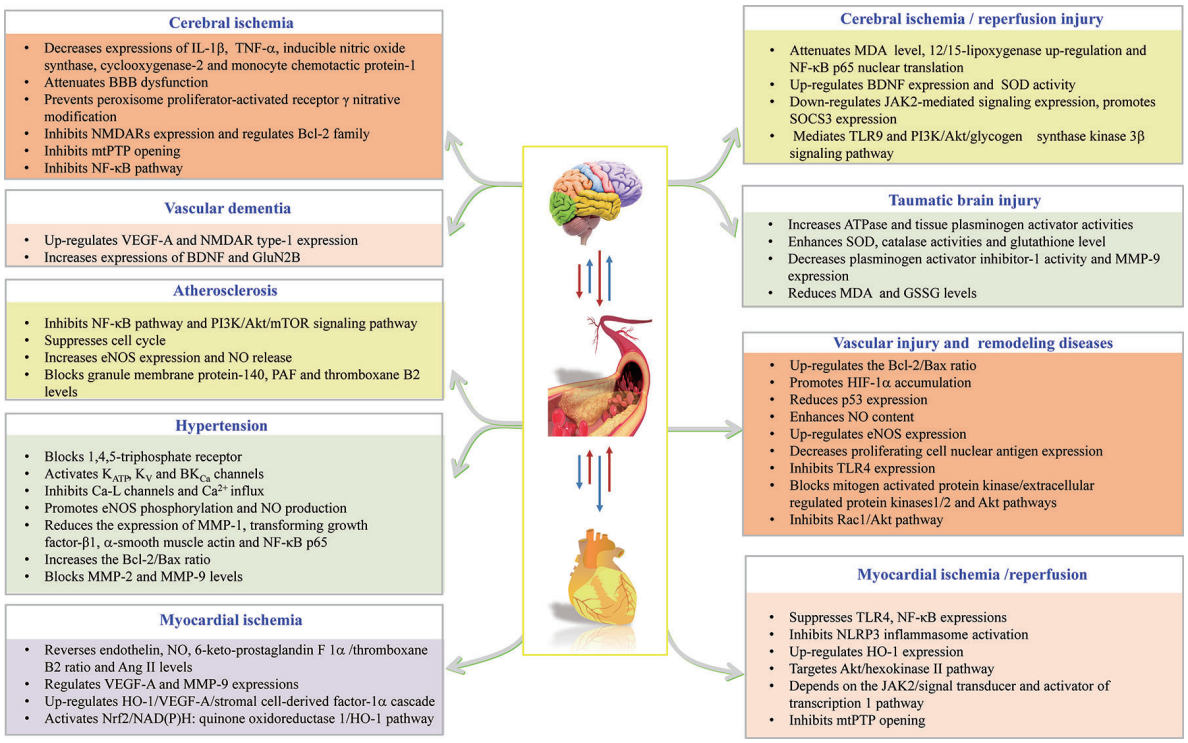

FIGURE 2 | HSYA acts on the functional targets and signaling pathways of cardio-cerebrovascular diseases.

However, none of the existing RCTs were of high methodological quality, and the conclusions need to be further verified by large sample, multicenter, and double-blind RCTs (as compared to traditional treatment regimens). In addition, clinical evidence supporting the application of HSYA for the management of CCVDs other than acute ischemic stroke with blood stasis syndrome is still lacking.

AUTHOR CONTRIBUTIONS

S-JY and Y-PT conceived and designed the review. XB searched the literature and drafted the manuscript. W-XW and R-JF examined the literature and made the figures. HG edited the manuscript. S-JY, Y-YC, and Y-PT made a critical revision of the review. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (81903786, 81773882), the Natural Science Foundation of Shaanxi Province (2019JQ-054), the Young Talent Support Program from the Association for Science and Technology of Colleges in Shaanxi Province (20190306), grants from the Key Research and Development Program of Shaanxi (2019ZDLSF04-05), Shaanxi Administration of Traditional Chinese Medicine (2019-ZZ-JC018), and Subject Innovation Team of Shaanxi University of Chinese Medicine (2019-YL10).

REFERENCES

Bacchetti, T., Morresi, C., Bellachioma, L., and Ferretti, G. (2020). Antioxidant and pro-oxidant properties of *Carthamus tinctorius*, hydroxy safflor yellow A, and safflor yellow A. *Antioxidants* 9, 119. doi: 10.3390/antiox9020119

Bai, Y. H., Lu, P., Han, C. H., Yu, C. Y., Chen, M. G., He, F., et al. (2012). Hydroxysafflor yellow A (HSYA) from flowers of *Carthamus tinctorius* L. and its vasodilatation effects on pulmonary artery. *Molecules* 17, 14918–14927. doi: 10.3390/molecules171214918

Bhosale, G., and Duchen, M. R. (2019). Investigating the mitochondrial permeability transition pore in disease phenotypes and drug screening. *Curr. Protoc. Pharmacol.* 85, e59. doi: 10.1002/cpph.59

Bie, X. D., Han, J., and Dai, H. B. (2010). Effects of hydroxysafflor yellow A on the experimental traumatic brain injury in rats. *J. Asian Nat. Prod. Res.* 12, 239–247. doi: 10.1080/10286020903510636

Braniste, V., AlAsmakh, M., Kowal, C., Anuar, F., Abbaspour, A., Toth, M., et al. (2014). The gut microbiota influences blood-brain barrier permeability in mice. *Sci. Transl. Med.* 6, 263ra158. doi: 10.1126/scitranslmed.3009759

Chen, L., Xiang, Y. X., Kong, L. J., Zhang, X. M., Sun, B. Z., and Wei, X. B. (2013). Hydroxysafflor yellow A protects against cerebral ischemia-reperfusion injury by antiapoptotic effect through PI3K/Akt/GSK3β pathway in rat. *Neurochem. Res.* 38, 2268–2275. doi: 10.1007/s11064-013-1135-8

Chen, S. N., Sun, M., Zhao, X. H., Yang, Z. F., Liu, W. X., Cao, J. Y., et al. (2019). Neuroprotection of hydroxysafflor yellow A in experimental cerebral ischemia/

reperfusion injury via metabolic inhibition of phenylalanine and mitochondrial biogenesis. *Mol. Med. Rep.* 19, 3009–3020. doi: 10.3892/mmr.2019.9959

Chen, Z. X., Xu, Q. Q., Shan, C. S., Shi, Y. H., Wang, Y., Chang, R. C. C., et al. (2019). Borneol for regulating the permeability of the blood-brain barrier in experimental ischemic stroke: preclinical evidence and possible mechanism. *Oxid. Med. Cell Longev.* 2019:2936737. doi: 10.1155/2019/2936737

Deng, L., Wan, H. T., Zhou, H. F., Yu, L., and He, Y. (2018). Protective effect of hydroxysafflor yellow A alone or in combination with acetylglutamine on cerebral ischemia reperfusion injury in rat: A PET study using 18F-fuorodeoxyglucose. *Eur. J. Pharmacol.* 825, 119–132. doi: 10.1016/j.ejphar.2018.02.011

Duan, J. L., Wang, J. W., Guan, Y., Yin, Y., Wei, G., Cui, J., et al. (2013). Safflor yellow A protects neonatal rat cardiomyocytes against anoxia/reoxygenation injury in vitro. *Acta Pharmacol. Sin.* 34, 487–495. doi: 10.1038/aps.2012.185

Ekin, Z. (2005). Resurgence of safflower (*Carthamus tinctorius* L.) utilization: a global view. *J. Agron.* 4, 83–87. doi: 10.3923/ja.2005.83.87

Fan, L. H., Dang, X. Q., Shi, Z. B., Zhang, C., and Wang, K. Z. (2011). Hydroxysafflor yellow A protects PC12 cells against the apoptosis induced by oxygen and glucose deprivation. *Cell. Mol. Neurobiol.* 31, 1187–1194. doi: 10.1007/s10571-011-9720-3

Fan, J. X., Qin, X. M., and Li, Z. Y. (2019). Molecular docking and multivariate analysis studies of active compounds in the safflower injection. *J. Liq. Chromatogr. R. T.* 42, 673–680. doi: 10.1080/10826076.2019.1665540

Gong, Z., Pan, J. R., Li, X. P., Wang, H. X., He, L., and Peng, Y. (2018). Hydroxysafflor yellow A reprograms TLR9 signalling pathway in ischaemic cortex after cerebral ischaemia and reperfusion. *CNS Neurol. Disord-DR.* 17, 370–382. doi: 10.2174/1871527317666180502110205

Guo, X. J., Zheng, M., Pan, R. Y., Zang, B. X., Gao, J. W., Ma, H. Y., et al. (2019). Hydroxysafflor yellow A (HSYA) targets the platelet-activating factor (PAF) receptor and inhibits human bronchial smooth muscle activation induced by PAF. *Food Funct.* 10, 4661–4673. doi: 10.1039/c9fo00896a

Han, D., Wei, J., Zhang, R., Ma, W., Shen, C., Feng, Y. D., et al. (2016). Hydroxysafflor yellow A alleviates myocardial ischemia/reperfusion in hyperlipidemic animals through the suppression of TLR4 signaling. *Sci. Rep.* 6:35319. doi: 10.1038/srep35319

He, P. P., Fu, F. H., Wang, T., Li, C. K., Xin, W. Y., and Zhang, X. M. (2008). Effect of cerebral ischemia/reperfusion injury on hydroxysafflor yellow A penetrating across the blood-brain barrier. *Sci. Pharm.* 76, 713–724. doi: 10.3797/scipharm.0811-01

Hu, T. X., Wei, G., Xi, M. M., Yan, J. J., Wu, X. X., Wang, Y. H., et al. (2016). Synergistic cardioprotective effects of Danshensu and hydroxysafflor yellow A against myocardial ischemia-reperfusion injury are mediated through the Akt/Nrf2/HO-1 pathway. *Int. J. Mol. Med.* 38, 83–94. doi: 10.3892/ijmm.2016.2584

Hu, Z. C., Xie, Z. J., Tang, Q., Li, X. B., Fu, X., Feng, Z. H., et al. (2018). Hydroxysafflor yellow A (HSYA) targets the NF-κB and MAPK pathways and ameliorates the development of osteoarthritis. *Food Funct.* 9, 4443–4456. doi: 10.1039/C8FO00732B

Hu, M. Z., Zhou, Z. Y., Zhou, Z. Y., Lu, H., Gao, M., Liu, L. M., et al. (2020). Effect and safety of hydroxysafflor yellow A for injection in patients with acute ischemic stroke of blood stasis syndrome: A phase II, multicenter, randomized, double-blind, multiple-dose, active-controlled clinical trial. *Chin. J. Integr. Med.* 26, 420–427. doi: 10.1007/s11655-020-3094-7

Huang, L., Li, L., Hu, E. C., Chen, G., Meng, X. M., Xiong, C. M., et al. (2018). Potential biomarkers and targets in reversibility of pulmonary arterial hypertension secondary to congenital heart disease: an explorative study. *Pulm. Circ.* 8:204589321875598. doi: 10.1177/2045893218755987

Huber, G. A., Priest, S. M., and Geisbuhler, T. P. (2018). Cardioprotective effect of hydroxysafflor yellow A via the cardiac permeability transition pore. *Planta Med.* 84, 507–518. doi: 10.1055/s-0043-122501

Ivey, M. E., Osman, N., and Little, P. J. (2008). Endothelin-1 signalling in vascular smooth muscle: Pathways controlling cellular functions associated with atherosclerosis. *Atherosclerosis* 199, 237–247. doi: 10.1016/j.atherosclerosis.2008.03.006

Ji, D. B., Zhu, M. C., Zhu, B., Zhu, Y. Z., Li, C. L., Ye, J., et al. (2008). Hydroxysafflor yellow A enhances survival of vascular endothelial cells under hypoxia via upregulation of the HIF-1α-VEGF pathway and regulation of Bcl-2/Bax. *J. Cardiovasc. Pharm.* 52, 191–202. doi: 10.1097/FJC.0b013e318181fb02

Ji, D. B., Zhang, L. Y., Li, C. L., Ye, J., and Zhu, H. B. (2009). Effect of hydroxysafflor yellow A on human umbilical vein endothelial cells under hypoxia. *Vasc. Pharmacol.* 50, 137–145. doi: 10.1016/j.vph.2008.11.009

Jiang, Y. Q., Kou, J. Y., Han, X. B., Li, X. S., Zhong, Z. Y., Liu, Z. N., et al. (2017). ROS-dependent activation of autophagy through the PI3K/Akt/mTOR pathway is induced by hydroxysafflor yellow A-sonodynamic therapy in THP-1 macrophages. *Oxid. Med. Cell. Longev.* 2017, 8519169. doi: 10.1155/2017/8519169

Jie, Z. Y., Xia, H. H., Zhong, S. L., Feng, Q., Li, S. H., Liang, S. S., et al. (2017). The gut microbiome in atherosclerotic cardiovascular disease. *Nat. Commun.* 8, 845. doi: 10.1038/s41467-017-00900-1

Jin, Z., Zhang, W. H., Chai, W. R., Zheng, Y. Q., and Zhi, J. M. (2013). Antibodies against AT1 receptors are associated with vascular endothelial and smooth muscle function impairment: Protective effects of hydroxysafflor yellow A. *PloS One* 8, e67020. doi: 10.1371/journal.pone.0067020

Jokinen, H., Kalska, H., Mäntylä, R., Pohjasvaara, T., Ylikoski, R., Hietanen, M., et al. (2006). Cognitive profile of subcortical ischaemic vascular disease. *J. Neurol. Neurosurg. Psychiatry* 77, 28–33. doi: 10.1136/jnnp.2005.069120

Kazantsev, A. G., and Outeiro, T. F. (2010). Drug discovery for CNS disorders: from bench to bedside. *CNS Neurol. Disord-DR.* 9, 668. doi: 10.2174/187152710793237395

Kotla, S., Singh, N. K., and Rao, G. N. (2017). ROS via BTK-p300-STAT1-PPARγ signaling activation mediates cholesterol crystals-induced CD36 expression and foam cell formation. *Redox Biol.* 11, 350–364. doi: 10.1016/j.redox.2016.12.005

Li, X. Z., Liu, J. X., Shang, X. H., and Fu, J. H. (2006). Protective effects of hydroxysafflor yellow A on acute myocardial ischemia in dogs. *Chin. Pharm. Bull.* 22, 533–537.

Li, J. R., Sun, M. J., Ping, Q. N., Chen, X. J., Qi, J. P., and Han, D. E. (2010). Metabolism, excretion and bioavailability of hydroxysafflor yellow A after oral administration of its lipid-based formulation and aqueous solution in rats. *Chin. J. Nat. Med.* 8, 233–240. doi: 10.3724/SP.J.1009.2010.00223

Li, J., Zhang, S. Y., Lu, M. R., Chen, Z. B., Chen, C., Han, L. J., et al. (2013). Hydroxysafflor yellow A suppresses inflammatory responses of BV2 microglia after oxygen-glucose deprivation. *Neurosci. Lett.* 535, 51–56. doi: 10.1016/j.neulet.2012.12.056

Liu, Y. N., Zhou, Z. M., and Chen, P. (2008). Evidence that hydroxysafflor yellow A protects the heart against ischaemia-reperfusion injury by inhibiting mitochondrial permeability transition pore opening. *Clin. Exp. Pharmacol. P.* 35, 211–216. doi: 10.1111/j.1440-1681.2007.04814.x

Liu, S. X., Zhang, Y., Wang, Y. F., Li, X. C., Xiang, M. X., Bian, C., et al. (2012). Upregulation of heme oxygenase-1 expression by hydroxysafflor yellow A conferring protection from anoxia/reoxygenation-induced apoptosis in H9c2 cardiomyocytes. *Int. J. Cardiol.* 160, 95–101. doi: 10.1016/j.ijcard.2011.03.033

Liu, Y. Y., Lian, Z. Q., Zhu, H. B., Wang, Y. H., Yu, S. S., Chen, T. T., et al. (2013). A systematic, integrated study on the neuroprotective effects of hydroxysafflor yellow A revealed by ^{1}H NMR-based metabonomics and the NF-κB pathway. *Evid. Based. Compl. Alt.* 2013, 147362. doi: 10.1155/2013/147362

Liu, Y. Q., Tian, X. F., Cui, M. Z., and Zhao, S. Z. (2014). Safflower yellow inhibits angiotensin II-induced adventitial fibroblast proliferation and migration. *J. Pharmacol. Sci.* 126, 107–114. doi: 10.1254/jphs.14055FP

Liu, J., Yue, S. J., Yang, Z. R., Feng, W. W., Meng, X. T., Wang, A. T., et al. (2018). Oral hydroxysafflor yellow A reduces obesity in mice by modulating the gut microbiota and serum metabolism. *Pharmacol. Res.* 134, 40–50. doi: 10.1016/j.phrs.2018.05.012

Liu, Z., Xu, Y. Q., and Ji, X. M. (2018). Progress in clinical diagnosis and treatment of cardio-cerebrovascular diseases. *Chin. J. Geriatr. Heart Brain Vessel Dis.* 20, 1219–1220. doi: 10.3969/j.issn.1009-0126.2018.11.025

Lv, Y. N., and Fu, L. S. (2018). The potential mechanism for hydroxysafflor yellow A attenuating blood-brain barrier dysfunction via tight junction signaling pathways excavated by an integrated serial affinity chromatography and shotgun proteomics analysis approach. *Neurochem. Int.* 112, 38–48. doi: 10.1016/j.neuint.2017.10.012

Lv, L. Z., Tong, C. Q., Lv, Q., Tao, X. J., Li, L. M., Fang, Q. X., et al. (2012). Enhanced absorption of hydroxysafflor yellow A using a self-double-emulsifying drug delivery system: In vitro and in vivo studies. *Int. J. Nanomed.* 7, 4099–4107. doi: 10.2147/IJN.S33398

Lv, Y. N., Qian, Y. S., Fu, L. S., Chen, X. Y., Zhong, H. L., and Wei, X. H. (2015). Hydroxysafflor yellow A exerts neuroprotective effects in cerebral ischemia

reperfusion-injured mice by suppressing the innate immune TLR4-inducing pathway. *Eur. J. Pharmacol.* 769, 324–332. doi: 10.1016/j.ejphar.2015.11.036

Lv, Y. N., Qian, Y. S., Ouyang, A. J., and Fu, L. S. (2016). Hydroxysafflor yellow A attenuates neuron damage by suppressing the lipopolysaccharide-induced TLR4 pathway in activated microglial cells. *Cell. Mol. Neurobiol.* 36, 1241–1256. doi: 10.1007/s10571-015-0322-3

Ma, G. N., Yu, F. L., Zhang, H., Li, Z. P., and Mei, X. G. (2014). Pharmacokinetics and biopharmaceutics of hydroxysafflor yellow A: research advances. *J. Int. Pharm. Res.* 41, 195–200. doi: 10.13220/j.cnki.jipr.2014.02.012

Ma, E. L., Smith, A., Desai, N., Cheung, L., Hanscom, M., Stoica, B. A., et al. (2017). Bidirectional brain-gut interactions and chronic pathological changes after traumatic brain injury in mice. *Brain Behav. Immun.* 66, 56–69. doi: 10.1016/j.bbi.2017.06.018

Martin, H. G. S., and Wang, Y. T. (2010). Blocking the deadly effects of the NMDA receptor in stroke. *Cell* 140, 174–176. doi: 10.1016/j.cell.2010.01.014

McSweeney, S. R., Warabi, E., and Siow, R. C. M. (2016). Nrf2 as an endothelial mechanosensitive transcription factor. *Hypertension* 67, 20–29. doi: 10.1161/HYPERTENSIONAHA.115.06146

Miao, T. J., Qian, L., Yu, F., Hu, L. G., Han, J. Q., and An, Y. (2019). Protective effects of hydroxysafflor yellow A on high oxidized low-density lipoprotein induced human coronary artery endothelial cells injuries. *Cancer Cell Res.* 22, 581–589.

Mills, K. T., Bundy, J. D., Kelly, T. N., Reed, J. E., Kearney, P. M., Reynolds, K., et al. (2016). Global disparities of hypertension prevalence and control: a systematic analysis of population-based studies from 90 countries. *Circulation* 134, 441–450. doi: 10.1161/CIRCULATIONAHA.115.018912

Min, J., and Wei, C. (2017). Hydroxysafflor yellow A cardioprotection in ischemia-reperfusion (I/R) injury mainly via Akt/hexokinase II independent of ERK/GSK-3β pathway. *Biomed. Pharmacother.* 87, 419–426. doi: 10.1016/j.biopha.2016.12.113

Minno, A. D., Stornaiuolo, M., and Novellino, E. (2019). Molecular scavengers, oxidative stress and cardiovascular disease. *J. Clin. Med.* 8, 1895. doi: 10.3390/jcm8111895

Moskowitz, M. A., Lo, E. H., and Iadecola, C. (2010). The science of stroke: mechanisms in search of treatments. *Neuron* 67, 181–198. doi: 10.1016/j.neuron.2010.07.002

Ni, B., Zhou, D. L., Jing, Y. Y., and Liu, S. X. (2018). Hydroxysafflor yellow A protects against angiotensin II-induced hypertrophy. *Mol. Med. Rep.* 18, 3649–3656. doi: 10.3892/mmr.2018.9399

Nie, P. H., Zhang, L., Zhang, W. H., Rong, W. F., and Zhi, J. M. (2012). The effects of hydroxysafflor yellow A on blood pressure and cardiac function. *J. Ethnopharmacol.* 139, 746–750. doi: 10.1016/j.jep.2011.11.054

Pearson, A. C., Pasierski, T., and Labovitz, A. J. (1991). Left ventricular hypertrophy: Diagnosis, prognosis and management. *Am. Heart J.* 121, 148–157. doi: 10.1016/0002-8703(91)90968-N

Petrovski, G., Ayna, G., Májai, G., Hodrea, J., Benkő, S., Mádi, A., et al. (2011). Phagocytosis of cells dying through autophagy induces inflammasome activation and IL-1β release in human macrophages. *Autophagy* 7, 321–330. doi: 10.4161/auto.7.3.14583

Qi, J. P., Zhuang, J., Wu, W., Lu, Y., Song, Y. M., Zhang, Z. T., et al. (2011). Enhanced effect and mechanism of water-in-oil microemulsion as an oral delivery system of hydroxysafflor yellow A. *Int. J. Nanomed.* 6, 985–991. doi: 10.2147/IJN.S18821

Qi, Z. F., Yan, F., Shi, W. J., Zhang, C. C., Dong, W., Zhao, Y. M., et al. (2014). AKT-related autophagy contributes to the neuroprotective efficacy of hydroxysafflor yellow A against ischemic stroke in rats. *Transl. Stroke Res.* 5, 501–509. doi: 10.1007/s12975-014-0346-x

Qin, S. C., Liang, W. X., Wen, X. W., and Deng, M. Z. (2016). Clinical efficacy of hydroxysafflor yellow A injection for acute ischemic stroke with blood stasis syndrome. *Chin. J. Exp. Tradit. Med. Form.* 9, 40–41. doi: 10.13422/j.cnki.syfjx.2016210157

Ribeiro, R., Badiwala, M. V., Ramzy, D., Tumiati, L. C., and Rao, V. (2019). Recipient hypertonic saline infusion prevents cardiac allograft dysfunction. *J. Thorac. Cardiov. Surg.* 157, 615–625. doi: 10.1016/j.jtcvs.2018.07.018

Romanic, A. M., White, R. F., Arleth, A. J., Ohlstein, E. H., and Barone, F. C. (1998). Matrix metalloproteinase expression increases after cerebral focal ischemia in rats: inhibition of matrix metalloproteinase-9 reduces infarct size. *Stroke* 29, 1020–1030. doi: 10.1161/01.str.29.5.1020

Scarabelli, T. M., Stephanou, A., and Pasini, E. (2002). Different signaling pathways induce apoptosis in endothelial cells and cardiac myocytes during ischemia/reperfusion injury. *Circ. Res.* 90, 745–748. doi: 10.1161/01.RES.0000015224.07870.9A

Sheng, L., Bi, S. J., Cheng, C., and Zhang, J. B. (2012). Hydroxysafflor yellow A suppresses oxidized low density lipoprotein induced proliferation of vascular smooth muscle cells. *Bangl. J. Pharmacol.* 7, 87–93. doi: 10.3329/bjp.v7i2.10499

Shi, M. X., Guo, Q., Huang, Y. R., Lu, W. J., and Zhang, J. Q. (2018). Pharmacokinetics of hydroxysafflor yellow A water-in-oil nanoemulsion. *Chin. Pharm. J.* 53, 2108–2112. doi: 10.11669/cpj.2018.24.007

Song, Y., Zhang, L., Qu, K., Li, C. L., and Zhu, H. B. (2005). Hydroxysafflor yellow A promotes vascular endothelial cell proliferation via VEGF/VEGF receptor. *J. Chin. Pharm. Sci.* 14, 181–185.

Song, Y. M., Long, L. L., Zhang, N., and Liu, Y. H. (2014). Inhibitory effects of hydroxysafflor yellow A on PDGF-BB-induced proliferation and migration of vascular smooth muscle cells via mediating Akt signaling. *Mol. Med. Rep.* 10, 1555–1560. doi: 10.3892/mmr.2014.2336

Sun, X., Wei, X. B., Qu, S. F., Zhao, Y. X., and Zhang, X. M. (2010). Hydroxysafflor yellow A suppresses thrombin generation and inflammatory responses following focal cerebral ischemia-reperfusion in rats. *Bioorg. Med. Chem. Lett.* 20, 4120–4124. doi: 10.1016/j.bmcl.2010.05.076

Sun, L., Yang, L., Xu, Y. W., Liang, H., Han, J., Zhao, R. J., et al. (2012). Neuroprotection of hydroxysafflor yellow A in the transient focal ischemia: inhibition of protein oxidation/nitration, 12/15-lipoxygenase and blood-brain barrier disruption. *Brain Res.* 1473, 227–235. doi: 10.1016/j.brainres.2012.07.047

Sun, L., Yang, L., Fu, Y., Han, J., Xu, Y. W., Liang, H., et al. (2013). Capacity of HSYA to inhibit nitrotyrosine formation induced by focal ischemic brain injury. *Nitric. Oxide* 35, 144–151. doi: 10.1016/j.niox.2013.10.002

Sun, L., Xu, Y. W., Han, J., Xiao, C., Cao, S. S., Liang, H., et al. (2018). Hydroxysafflor yellow A shows protection against PPARγ inactivation in nitrosative neurons. *Oxid. Med. Cell Longev.* 2018:9101740. doi: 10.1155/2018/9101740

Sun, Y., Xu, D. P., Qin, Z., Wang, P. Y., Hu, B. H., Yu, J. G., et al. (2018). Protective cerebrovascular effects of hydroxysafflor yellow A (HSYA) on ischemic stroke. *Eur. J. Pharmacol.* 818, 604–609. doi: 10.1016/j.ejphar.2017.11.033

Tang, T. W. H., Chen, H. C., Chen, C. Y., Yen, C. Y. T., Lin, C. J., Prajnamitra, R. P., et al. (2019). Loss of gut microbiota alters immune system composition and cripples postinfarction cardiac repair. *Circulation* 139, 647–659. doi: 10.1161/CIRCULATIONAHA.118.035235

Thiagarajan, H., Thiyagamoorthy, U., Shanmugham, I., Nandagopal, G. D., and Kaliyaperumal, A. (2017). Angiogenic growth factors in myocardial infarction: A critical appraisal. *Heart Fail. Rev.* 22, 665–683. doi: 10.1007/s10741-017-9630-7

Tian, J. W., Fu, F. H., Jiang, W. L., Wang, C. Y., Sun, F., and Zhang, T. P. (2004). Protective effect of hydroxysaflor yellow A against rat cortex mitochondrial injuries induced by cerebral ischemia. *Acta Pharm. Sin.* 39, 774–777. doi: 10.16438/j.0513-4870.2004.10.002

Tian, J. W., Li, G. S., Liu, Z. F., and Fu, F. H. (2008). Hydroxysafflor yellow A inhibits rat brain mitochondrial permeability transition pores by a free radical scavenging action. *Pharmacology* 82, 121–126. doi: 10.1159/000141653

Upadhyay, R. K. (2014). Drug delivery systems, CNS protection, and the blood brain barrier. *Biomed. Res. Int.* 2014, 869269. doi: 10.1155/2014/869269

Wang, T., Fu, F. H., Han, B., Zhu, M., Li, G. S., and Liu, K. (2007). Protection and mechanisms of hydroxysafflor yellow A on experimental myocardial infarction in rats. *Chin. Tradit. Herb. Drugs* 38, 1853–1856.

Wang, Z. J., Ke, Y. N., Chen, W. L., and Yu, C. A. (2008). Analysis of the platelet activated function of hyperlipemia rats. *Chin. J. Compar. Med.* 18, 10–13.

Wang, C. Y., Ma, H. M., Zhang, S. P., Wang, Y. F., Liu, J. T., and Xiao, X. H. (2009). Safflor yellow B suppresses pheochromocytoma cell (PC12) injury induced by oxidative stress via antioxidant systemand Bcl-2/Bax pathway. *N. S. Arch. Pharmacol.* 380, 135–142. doi: 10.1007/s00210-009-0424-x

Wang, C. Y., He, Y. H., Yang, M., Sun, H. L., Zhang, S. P., and Wang, C. H. (2013). Safflor yellow B suppresses angiotensin II-mediated human umbilical vein cell injury via regulation of Bcl-2/$p22^{phox}$ expression. *Toxicol. Appl. Pharm.* 273, 59–67. doi: 10.1016/j.taap.2013.08.018

Wang, J. P., Zhang, Q., Mei, X. H., and Zhang, X. Z. (2014). Hydroxysafflor yellow A attenuates left ventricular remodeling after pressure overload-induced cardiac hypertrophy in rats. *Pharm. Biol.* 52, 31–35. doi: 10.3109/13880209.2013.805791

Wang, H. F., Liu, J. L., Yang, Y. J., Cao, Q. W., Huo, X. P., Ma, S. H., et al. (2016). Hydroxy-safflower yellow A inhibits the TNFR1-mediated classical NF-κB

pathway by inducing shedding of TNFR1. *Phytother. Res.* 30, 790–796. doi: 10.1002/ptr.5579

Wang, X. T., Ma, Z. Y., Fu, Z. X., Gao, S., Yang, L., Jin, Y., et al. (2016). Hydroxysafflor yellow A protects neurons from excitotoxic death through inhibition of NMDARs. *ASN Neuro* 8, 1759091416642345. doi: 10.1177/1759091416642345

Wang, Y., Zhang, C. H., Peng, W. J., Xia, Z. A., Gan, P. P., Huang, W., et al. (2016). Hydroxysafflor yellow A exerts antioxidant effects in a rat model of traumatic brain injury. *Mol. Med. Rep.* 14, 3690–3696. doi: 10.3892/mmr.2016.5720

Wang, N., He, D. M., Zhou, Y. Q., Wen, J., Liu, X. Q., Li, P. Y., et al. (2020). Hydroxysafflor yellow A actives BKCa channels and inhibits L-type Ca channels to induce vascular relaxation. *Eur. J. Pharmacol.* 870, 172873. doi: 10.1016/j.ejphar.2019.172873

Wei, X. B., Liu, H. Q., Sun, X., Fu, F. H., Zhang, X. M., Wang, J., et al. (2005). Hydroxysafflor yellow A protects rat brains against ischemia-reperfusion injury by antioxidant action. *Neurosci. Lett.* 386, 58–62. doi: 10.1016/j.neulet.2005.05.069

Wei, G., Yin, Y., Duan, J. L., Guo, C., Zhu, Y. R., Wang, Y. H., et al. (2017). Hydroxysafflor yellow A promotes neovascularization and cardiac function recovery through HO-1/VEGF-A/SDF-1α cascade. *Biomed. Pharmacother.* 88, 409–420. doi: 10.1016/j.biopha.2017.01.074

Wei, Z. Y., Xu, W. J., Dong, J. J., Liu, J., Jia, Z. X., Chen, Y. J., et al. (2018). Hydroxysafflor yellow A repairing the metabolic disturbances of early atherosclerosis based on fatty acid profiling. *Acta Pharm. Sin.* 53, 1680–1688. doi: 10.16438/j.0513-4870.2018-0409

Wu, X., Ouyang, L. N., Xiang, D. W., and Xiang, D. X. (2011). Enhancing effect of Borneolum Syntheticum and Acori Talarinowii Rhizoma on penetrating blood-brain barrier of hydroxysafflor yellow A. *Chin. Tradit. Herb. Drugs* 42, 734–737.

Xie, Y. Y., Zhang, J. P., Fang, Z. H., Li, Y. F., and Wang, M. Y. (2018). Research progress of different sources exosomes and myocardial ischemia reperfusion injury. *J. Clin. Cardiol.* 34, 841–845. doi: 10.13201/j.issn.1001-1439.2018.09.002

Xing, M. Y., Sun, Q. N., Wang, Y. Y., Cheng, Y., and Zhang, N. (2016). Hydroxysafflor yellow A increases BDNF and NMDARs in the hippocampus in a vascular dementia rat model. *Brain Res.* 1642, 419–425. doi: 10.1016/j.brainres.2016.04.030

Xiu, G. H., Sun, J., Li, X. L., Yin, Y. Y., Li, B. Q., and Lin, B. (2018). Effects of hyperbaric oxygen treatment on neurological functional and expression of CyclinD1 in rats with traumatic brain injury. *J. Clin. Neurosurg.* 15, 46–54. doi: 10.3969/j.issn.1672-7770.2018.01.011

Xu, R. A., Xu, Z. S., and Ge, R. S. (2014). Effects of hydroxysafflor yellow A on the activity and mRNA expression of four CYP isozymes in rats. *J. Ethnopharmacol.* 151, 1141–1146. doi: 10.1016/j.jep.2013.12.025

Xu, H., Liu, W. X., Liu, T. L., Su, N., Guo, C., Feng, X. N., et al. (2017). Synergistic neuroprotective effects of Danshensu and hydroxysafflor yellow A on cerebral ischemia-reperfusion injury in rats. *Oncotarget* 8, 115434–115443. doi: 10.18632/oncotarget.23272

Xu, H., Liu, T. L., Wang, W. J., Su, N., Yang, L. D., Yang, Z. F., et al. (2019). Proteomic analysis of hydroxysafflor yellow A against cerebral ischemia-reperfusion injury in rats. *Rejuv. Res.* 22, 503–512. doi: 10.1089/rej.2018.2145

Xuan, J. W., Huang, M., Lu, Y. J., and Tao, L. B. (2018). Economic evaluation of safflower yellow injection for the treatment of patients with stable angina pectoris in China: A cost-effectiveness analysis. *J. Altern. Complem. Med.* 24, 564–569. doi: 10.1089/acm.2017.0284

Yang, Q., Yang, Z. F., Liu, S. B., Zhang, X. N., Hou, Y., Li, X. Q., et al. (2010). Neuroprotective effects of hydroxysafflor yellow A against excitotoxic neuronal death partially through down-regulation of NR2B-containing NMDA receptors. *Neurochem. Res.* 35, 1353–1360. doi: 10.1007/s11064-010-0191-6

Yang, G. S., Zhou, X. Y., Chen, T., Deng, Y. D., Yu, D., Pan, S. Y., et al. (2015). Hydroxysafflor yellow A inhibits lipopolysaccharide-induced proliferation and migration of vascular smooth muscle cells via Toll-like receptor-4 pathway. *Int. J. Clin. Exp. Med.* 8, 5295–5302.

Yang, G., Wang, N., Seto, S. W., Chang, D., and Liang, H. Z. (2018). Hydroxysafflor yellow A protects brain microvascular endothelial cells against oxygen glucose deprivation/reoxygenation injury: Involvement of inhibiting autophagy via class I PI3K/Akt/mTOR signaling pathway. *Brain Res. Bull.* 140, 243–257. doi: 10.1016/j.brainresbull.2018.05.011

Yang, J. F., Wang, R., Cheng, X. H., Qu, H. C., Qi, J., Li, D., et al. (2020). The vascular dilatation induced by hydroxysafflor yellow A (HSYA) on rat mesenteric artery through TRPV4-dependent calcium influx in endothelial cells. *J. Ethnopharmacol.* 256, 112790. doi: 10.1016/j.jep.2020.112790

Ye, F., Wang, J. H., Meng, W., Qian, J. R., and Jin, M. (2017). Proteomic investigation of effects of hydroxysafflor yellow A in oxidized low-density lipoprotein-induced endothelial injury. *Sci. Rep.* 7, 17981. doi: 10.1038/s41598-017-18069-4

Ye, J. X., Wang, M., Wang, R. Y., Liu, H. T., Qi, Y. D., Fu, J. H., et al. (2020). Hydroxysafflor yellow A inhibits hypoxia/reoxygenation-induced cardiomyocyte injury via regulating the AMPK/NLRP3 inflammasome pathway. *Int. Immunopharmacol.* 82, 106316. doi: 10.1016/j.intimp.2020.106316

Yu, L., Duan, Y. H., Zhao, Z., He, W. D., Xia, M., Zhang, Q. J., et al. (2018). Hydroxysafflor yellow A (HSYA) improves learning and memory in cerebral ischemia reperfusion-injured rats via recovering synaptic plasticity in the hippocampus. *Front. Cell. Neurosci.* 12, 371. doi: 10.3389/fncel.2018.00371

Yu, L., Liu, Z. L., He, W. D., Chen, H. F., Lai, Z. L., Duan, Y. H., et al. (2020). Hydroxysafflor Yellow A Confers Neuroprotection from focal cerebral ischemia by modulating the crosstalk between JAK2/STAT3 and SOCS3 signaling pathways. *Cell. Mol. Neurobiol.* 40, 1271–1281. doi: 10.1007/s10571-020-00812-7

Yuan, W. D., Yang, D. X., Sun, X. H., Liu, W., Wang, L., Li, X. Y., et al. (2014). Effects of hydroxysafflor yellow A on proliferation and collagen synthesis of rat vascular adventitial fibroblasts induced by angiotensin II. *Int. J. Clin. Exp. Patho.* 7, 5772–5781.

Yue, S. J., Tang, Y. P., Li, S. J., and Duan, J. A. (2013). Chemical and biological properties of quinochalcone *C*-glycosides from the florets of *Carthamus Tinctorius*. *Molecules* 18, 15220–15254. doi: 10.3390/molecules181215220

Yue, S. J., Tang, Y. P., Wang, L. Y., Tang, H., Li, S. J., Liu, P., et al. (2014). Separation and evaluation of antioxidant constituents from *Carthamus tinctorius*. *China J. Chin. Mater. Med.* 39, 3295–3300.

Yue, S. J., Qu, C., Zhang, P. X., Tang, Y. P., Jin, Y., Jiang, J. S., et al. (2016). Carthorquinosides A and B, quinochalcone *C*–glycosides with diverse dimeric skeletons from *Carthamus tinctorius*. *J. Nat. Prod.* 79, 2644–2651. doi: 10.1021/acs.jnatprod.6b00561

Yue, S. J., Wang, W. X., Yu, J. G., Chen, Y. Y., Shi, X. Q., Yan, D., et al. (2019a). Gut microbiota modulation with traditional Chinese medicine: A system biology-driven approach. *Pharmacol. Res.* 148, 104453. doi: 10.1016/j.phrs.2019.104453

Yue, S. J., Liu, J., Wang, A. T., Meng, X. T., Yang, Z. R., Peng, C., et al. (2019b). Berberine alleviates insulin resistance by reducing peripheral branched-chain amino acids. *Am. J. Physiol. Endoc. M.* 316, E73–E85. doi: 10.1152/ajpendo.00256.2018

Zang, B. X., Jin, M., Si, N., Zhang, Y., Wu, W., and Piao, Y. Z. (2002). Antagonistic effect of hydroxysafflor yellow A on the platelet activating factor receptor. *Acta Pharm. Sin.* 37, 696–699. doi: 10.16438/j.0513-4870.2002.09.006

Zhang, L., Nie, P. H., Zhang, G. H., Rong, W. F., and Zhi, J. M. (2011). Endothelium-independent vasodilation effect of hydroxysafflor yellow A in thoracic aorta of wistar rats. *J. Med. Plant Res.* 5, 2187–2191. doi: 10.1002/cmdc.201100121

Zhang, N., Xing, M. Y., Wang, Y. Y., Liang, H., Yang, Z., Shi, F. D., et al. (2014). Hydroxysafflor yellow A improves learning and memory in a rat model of vascular dementia by increasing VEGF and NR1 in the hippocampus. *Neurosci. Bull.* 30, 417–424. doi: 10.1007/s12264-013-1375-2

Zhang, J. X., Qi, M. J., Shi, M. Z., Chen, J. J., Zhang, X. Q., Yang, J., et al. (2019). Effects of Danhong injection, a traditional Chinese medicine, on nine cytochrome P450 isoforms in vitro. *Biomed. Chromatogr.* 33, e4454. doi: 10.1002/bmc.4454

Zhao, J. S., Fang, M. X., Guo, Q. S., Li, Y. F., Xu, B. Y., Lai, S. H., et al. (2015). Hydroxysafflor yellow A inhibited rat vascular smooth muscle cells proliferation induced by PDGF. *Chin. J. Cell Biol.* 37, 827–831. doi: 10.11844/cjcb.2015.06.0032

Zhao, B. X., Gu, S. F., Du, Y., Shen, M. J., Liu, X. R., and Shen, Y. Q. (2018). Solid lipid nanoparticles as carriers for oral delivery of hydroxysafflor yellow A. *Int. J. Pharmacol.* 535, 164–171. doi: 10.1016/j.ijpharm.2017.10.040

Zhao, X. X., Li, H. S., and Wang, H. Y. (2018). Research progress on risk factors of vascular dementia. *Adv. Cardiovasc. Dis.* 39, 328–331. doi: 10.16806/j.cnki.issn.1004-3934.2018.03.007

Zhou, D. L., Ding, T. T., Ni, B., Jing, Y. Y., and Liu, S. X. (2019). Hydroxysafflor yellow A mitigated myocardial ischemia/reperfusion injury by inhibiting the

activation of the JAK2/STAT1 pathway. *Int. J. Mol. Med.* 44, 405–416. doi: 10.3892/ijmm.2019.4230

Zhu, H. B., Zhang, L., Wang, Z. H., Tian, J. W., Fu, F. H., Liu, K., et al. (2005). Therapeutic effects of hydroxysafflor yellow A on focal cerebral ischemic injury in rats and its primary mechanisms. *J. Asian Nat. Prod. Res.* 7, 607–613. doi: 10.1080/10286020310001625120

Zou, J., Wang, N. A., Liu, M. T., Bai, Y. P., Wang, H., Liu, K., et al. (2018). Nucleolin mediated pro-angiogenic role of hydroxysafflor yellow A in ischaemic cardiac dysfunction: Post-transcriptional regulation of VEGF-A and MMP-9. *J. Cell. Mol. Med.* 22, 2692–2705. doi: 10.1111/jcmm.13552

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.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 26 November 2020
doi: 10.3389/fphar.2020.574638

The Pharmacokinetics, Tissue Distribution, Metabolism, and Excretion of Pinostrobin in Rats: Ultra-High-Performance Liquid Chromatography Coupled With Linear Trap Quadrupole Orbitrap Mass Spectrometry Studies

Xiaoya Sun[1,2], *Xiaojun Liu*[1] *and Suiqing Chen*[1,2]*

[1]School of Pharmacy, Henan University of Chinese Medicine, Zhengzhou, China, [2]Collaborative Innovation Center for Respiratory Disease Diagnosis and Treatment & Chinese Medicine Development of Henan Province, Henan University of Chinese Medicine, Zhengzhou, China

OPEN ACCESS

Edited by:
Salvatore Salomone,
University of Catania, Italy

Reviewed by:
Pan Deng,
University of Kentucky, United States
E. Sun,
Jiangsu Provincial Hospital of Traditional Chinese Medicine, China

***Correspondence:**
Suiqing Chen
suiqingchen0371@163.com

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: 20 June 2020
Accepted: 21 October 2020
Published: 26 November 2020

Citation:
Sun X, Liu X and Chen S (2020) The Pharmacokinetics, Tissue Distribution, Metabolism, and Excretion of Pinostrobin in Rats: Ultra-High-Performance Liquid Chromatography Coupled With Linear Trap Quadrupole Orbitrap Mass Spectrometry Studies. Front. Pharmacol. 11:574638. doi: 10.3389/fphar.2020.574638

Pinostrobin is a natural flavonoid found in various plants, well known for its wide range of pharmacological activities. However, there are few reports regarding the pharmacokinetics, tissue distribution, metabolism, and excretion of pinostrobin in rats after oral administration as a single compound. Therefore, we established a method using ultra-high-performance liquid chromatography coupled with linear trap quadrupole orbitrap mass spectrometry (UPLC-LTQ orbitrap-MS/MS) to determine pinostrobin and its metabolites in rat plasma, urine, feces, bile, and tissue homogenates. Pharmacokinetic parameters were measured. The large apparent volume of distribution implied that pinostrobin preferentially bound to tissues and preferably remained within the body. Based on previous pharmacological studies of its antiulcer, anti-*HP*, anti-inflammatory, and antioxidant activities, pinostrobin is mostly distributed in the gastrointestinal tract, indicating its potential as an effective component of traditional Chinese medicines for the treatment of peptic ulcers. Furthermore, 30 flavonoid metabolites were screened using UPLC-LTQ orbitrap-MS/MS. The metabolism pathways (mainly hydroxylation, demethylation, glucuronidation, and sulfation) of pinostrobin in rats have also been proposed. A small amount of pinostrobin in its parent form is excreted through the urine, feces, and bile, indicating that it is mainly metabolized *in vivo*. In this study, we systemically investigated the pharmacokinetics, tissue distribution, metabolism, and excretion of pinostrobin in rats. Our results provide a significant basis for the clinical development and application of pinostrobin as well as traditional Chinese medicines containing pinostrobin.

Keywords: pinostrobin, UPLC-LTQ orbitrap-MS/MS, pharmacokinetics, tissue distribution, metabolites, excretion

INTRODUCTION

Peptic ulcers are a common disease encountered in the clinic. Epidemiology estimated the morbidity rate of peptic ulcer and its related disorders to be as high as 10%, especially in the past 10 years, and its incidence continues to increase (Lanas and Chan, 2017; Kuna et al., 2019). For the past few years, *Lindera reflexa* Hemsl., *Boesenbergia rotunda* (L.) Mansfield, *Dysphania graveolens*, and *Teloxys graveolens* have been used to treat gastrointestinal disorders, including peptic ulcers. Flavonoids, especially pinostrobin, are the main active components of these plants that contribute to the effective treatment of peptic ulcer (Meckes et al., 1998; Sun et al., 2016; Déciga-Campos et al., 2017; Kanchanapiboon et al., 2020).

Pinostrobin, a widely studied dietary bioflavonoid, was discovered in the heartwood of pine (*Pinus strobus* L.) more than 7 decades ago (Erdtman, 1944). In addition to Pinaceae, pinostrobin has also been found in more than 10 families such as Lauraceae, Zingiberaceae, Fabaceae, and Polygonaceae (Patel and Bhutani, 2014; Chen et al., 2015; Gómez-Betancur et al., 2015; Dzoyem et al., 2017; Vasas et al., 2020). Pinostrobin content has been found to control the quality of Linderae Reflexae Radix. It is found more abundantly in roots than other secondary metabolites, such as pinocembrin and pinosylvin (Wang et al., 2016). Pinostrobin is known to have various pharmacological activities (Patel et al., 2016), including antiulcer, anti-*Helicobacter pylori* (Bhamarapravati et al., 2006), antioxidant, anti-inflammatory, anticancer (Jaudan et al., 2018; Jones and Gehler, 2020; Sopanaporn et al., 2020), antidiarrheal, antiviral (Wu et al., 2011), antimicrobial (Hernández Tasco et al., 2020), anti-Alzheimer's, antiprotozoal, antinociceptive (Déciga-Campos et al., 2017), antimutagenic, antiplatelet (Zhang et al., 2017), antiproliferative (Siekmann et al., 2013; Jadaun et al., 2017), antileukemic (Smolarz et al., 2005), antiosteoporotic (Gu et al., 2017), and antiparasitic properties (Vechi et al., 2020). Moreover, pinostrobin can protect the gastric mucosa by reducing the ulcer area and mucosal content and reducing or eliminating submucosal edema and leukocyte infiltration. Pinostrobin has a significant, dose-dependent protective effect on ethanol-induced gastric mucosal injury by scavenging free radicals produced by ethanol through the activation of cellular antioxidant defenses (Abdelwahab et al., 2011). Meanwhile, it exhibits nontoxic and nongenotoxic effects (Charoensin et al., 2010). However, to the best of our knowledge, the pharmacokinetic processes of pinostrobin have not yet been explained clearly.

Pharmacokinetic profiling, including absorption, distribution, metabolism, and excretion (ADME) processes, is vital to understanding the *in vivo* behavior and mechanism of action of compounds (Zeng et al., 2019). To date, there have been only a few studies regarding the determination of pinostrobin in rat plasma using high-performance liquid chromatography-ultraviolet (HPLC-UV) or HPLC-MS/MS after intravenous and intragastric administration, and only some of the pharmacokinetic parameters of pinostrobin have been reported (Hua et al., 2011; Sayre et al., 2013; Sayre et al., 2015). Hence, it is meaningful to explore the ADME processes of pinostrobin in rats after oral administration as a single compound.

In this study, we systemically investigated the pharmacokinetics and tissue distribution, metabolism, and excretion properties of pinostrobin in rats after a single oral administration. The pharmacokinetic parameters were consistent with those reported in previous studies. Pinostrobin was mostly distributed in the gastrointestinal tract. The amount of pinostrobin excreted via urine, feces, and bile in the parent form was less than 1.567%. Thirty flavonoid metabolites were identified or partially identified in biosamples collected after dosing. In addition, we proposed metabolic pathways in rats. This study provides helpful information for the clinical study of pinostrobin and traditional Chinese medicines containing pinostrobin.

MATERIALS AND METHODS

Chemicals and Materials

Pinostrobin was isolated from the roots of *L. reflexa* Hemsl. in our laboratory. Its structure (**Figure 1**) was unequivocally elucidated using spectroscopic methods (IR, MS, ^{1}H-NMR, and ^{13}C-NMR) (Chen et al., 2015; Sun et al., 2016). The absolute stereochemistry of pinostrobin was found to be 2S by comparing the ECD spectra (**Supplementary Figure S1**) with previously published data (Gaffield, 1970; Su et al., 2003). Its purity was determined to be above 98% by normalizing the peak area using an HPLC-diode array detector. Isoliquiritigenin (HPLC ≥ 98%) obtained from Shanghai Yuanye Bio-Technology Co., Ltd., was used as the internal standard (IS). Ultrapure water was prepared using a Milli-Q water purification system (Millipore, Milford, MA, United States). Heparin sodium was purchased from Beijing Dingguo Changsheng Bio-Technology Co., Ltd. Other reagents used were of HPLC grade. Acetonitrile, methanol, and formic acid were purchased from Fisher Scientific (Fairlawn, NJ, United States).

Animals

Male Sprague-Dawley rats, weighing 200 g ± 20 g (Certificate No. SCXK 2015-0004), were obtained from the Henan Experimental Animal Center (Zhengzhou, China). All animals were maintained on a 12/12 h light/dark cycle in an environmentally controlled breeding room (temperature 22°C ± 2°C; relative humidity 55% ± 5%). They were acclimated for one week before the initiation of dosing, with free access to food and water, and then fasted overnight (12 h) with water *ad libitum* prior to experimentation. The animal experiments were conducted in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and approved by the Experimental Animal Ethics Committee of the Henan University of Chinese Medicine. A single 48.51 mg/kg dose of pinostrobin was separately administered orally to each group (pharmacokinetics, tissue distribution, and excretion groups). The samples from pharmacokinetics, tissue distribution, and excretion groups were used for metabolism studies.

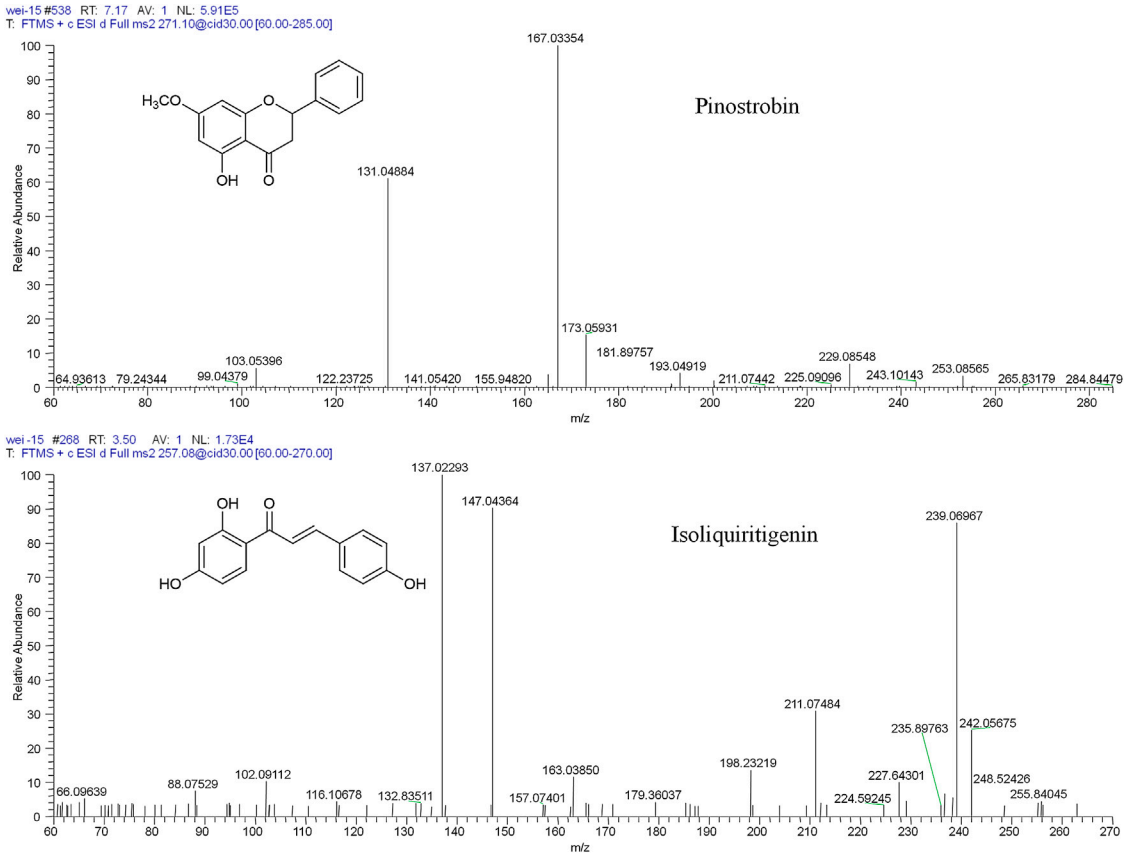

FIGURE 1 | The chemical structures and mass spectra of pinostrobin and isoliquiritigenin (IS).

Pharmacokinetic Study

Pinostrobin solution was prepared in 0.9% sterile saline containing 2% polysorbate 80 (v/v). Pinostrobin was administered orally to five male rats. The volume of the dosing solution administered was adjusted according to the body weights recorded before dose administration. At 0 (prior to dosing), 0.133, 0.167, 0.33, 0.50, 1, 1.33, 1.67, 2, 4, 6, and 12 h after dosing, blood samples (~400 μL) were collected from each animal via the fosse orbital vein using a heparinized 1.5 mL polythene tube. Plasma samples were obtained by immediate centrifugation (4,000 rpm for 10 min, 4°C) and stored at −80°C until analysis.

Tissue Distribution Study

To investigate the tissue distribution of pinostrobin, rats were randomly divided into 11 groups of five rats each. Before and after oral administration of pinostrobin, the heart, liver, spleen, lung, kidney, stomach, small intestine, and large intestine samples were collected at 0 (prior to dosing), 0.083, 0.33, 0.75, 1, 2, 3, 6, 8, 12, and 24 h after dosing. Tissue samples were washed with normal saline and dried with filter paper. The chyme in the gastrointestinal tract (including the stomach, small intestine, and large intestine) was removed before washing. Subsequently, the samples were accurately weighed to obtain their wet weight and stored at −80°C until analysis.

Excretion Study

Rats were randomly divided into two groups (n = 5). One group was used in the urinary and fecal excretion study. Five rats were orally administered pinostrobin and housed individually in stainless-steel metabolic cages, which allowed the separate collection of urine and feces. Urine and fecal samples were collected before administration and at time intervals of 0–2, 2–4, 4–6, 6–8, 8–12, 12–24, 24–36, and 36–54 h after dosing. The specimens were stored at −80°C after measuring the urine volume and the dry weight of feces for each time interval.

The other group was used in the biliary excretion study. Five rats were anesthetized and a cannula was implanted into the bile duct to collect bile. After the oral administration of pinostrobin, bile samples were collected at 0–3, 3–5, 5–7, 7–9, 9–12, 12–16, 16–20, and 20–24 h after dosing and stored at −80°C after recording the volume for each time interval.

Instrument and Analytical Conditions

An ultra-high-performance liquid chromatography-linear trap quadrupole orbitrap mass spectrometry (UPLC-LTQ orbitrap-MS/MS) method was applied using a Dionex Ultimate 3000 UPLC tandem LTQ Orbitrap XL Hybrid mass spectrometer equipped with an electrospray ionization source (Thermo Fisher Scientific, Bremen, Germany). Chromatographic

separation was performed on a Hypersil GOLD C18 column (2.1 mm × 50 mm, 1.9 μm) in tandem with a guard column and a UPLC filter cartridge (2.1 mm × 0.2 μm; Thermo Fisher Scientific) at a column temperature of 30°C. The mobile phase comprised 0.1% formic acid (v/v) aqueous solution (A) and acetonitrile (B). The elution gradient was set as follows: 0.0–15.0 min, 28.0–100.0% B; 15.0–15.1 min, 100.0–28.0% B; 15.1–18.0 min, 28% B. The flow rate was 0.3 mL/min, and the injection volume was 5 μL, with the autosampler conditioned at 7°C.

Mass spectrometry was performed in the positive ion mode. The ion source parameters were set as follows: capillary temperature, 350°C; ion spray voltage, 4.2 kV; capillary voltage, 49 V; tube lens voltage, 105 V; and the sheath (N_2) and auxiliary gas (He) flow rates of 40 and 10 arbitrary units, respectively. For pharmacokinetics, tissue distribution, and excretion studies, selected ion monitoring (SIM) was used with the ion *m*/*z* 271.09607 indicating pinostrobin and *m*/*z* 257.08026 indicating isoliquiritigenin (**Figure 1**). For the metabolism study, full scanning was used with a scan range of *m*/*z* 80–1,500. The normalized collision energy for collision-induced dissociation (CID) was adjusted to 30% of the maximum. The isolation width of the precursor ions was *m*/*z* 2.0, and the default values were used for other CID parameters. Accurate masses were calibrated according to the manufacturer's guidelines using a standard mixture of caffeine, MRFA, and Ultramark 1621. Data acquisition and processing were performed using Xcalibur 3.0 software (Thermo Fisher Scientific) and Mass Frontier 7.0 software (Thermo Fisher Scientific, Waltham, MA, United States) (Sun et al., 2016).

Preparation of Standard and Quality Control Samples

The stock solutions of pinostrobin (1.0 mg/mL) and IS (0.4 mg/mL) were separately prepared in methanol and stored at 4°C. Standard solutions of pinostrobin at desired concentrations were prepared by serial dilution of the stock solution with methanol every 2 weeks. The IS solution was diluted with methanol to 400 ng/mL. All solutions were kept at 4°C before use. Calibration standards were prepared by spiking 20 μL of the appropriate standard solution, 50 μL IS solution, and 180 μL of blank rat plasma, urine, fecal, bile, or tissue homogenate samples. Calibration standards were prepared at concentrations of 4, 10, 50, 100, 500, 1,000, and 2,000 ng/mL for plasma, urine, fecal, and bile; 8, 20, 100, 200, 1,000, 2,000, and 4,000 ng/g for the heart, liver, spleen, lung, and kidney; 8, 20, 100, 200, 1,000, 2,000, 4,000, and 10,000 ng/g for the stomach, small intestine, and large intestine. The quality control (QC) samples included lower limit of quantification (LLOQ), low, middle, and high QCs. They were prepared in the same manner using blank biological samples of 4, 10, 400, and 2,000 ng/mL for plasma, urine, fecal, and bile samples; 8, 20, 800, and 4,000 ng/g for the heart, liver, spleen, lung, and kidney; and 8, 20, 1,000, and 10,000 ng/g for the stomach, small intestine, and large intestine, respectively. Calibration and QC samples were stored at 4°C until analysis.

Sample Pretreatment

For the pharmacokinetic study, an aliquot of 200 μL of rat plasma samples (blank plasma, calibration standards, QC samples, and pharmacokinetic plasma samples) was vortexed with IS (50 μL, 400 ng/mL) and then with a solution of methanol–acetonitrile (1.0 mL, 5 : 95, v/v) for extraction. The mixture was vortexed for 5 min and then centrifuged at 10,000 rpm for 10 min at 4°C. The supernatant was transferred into a clean tube and evaporated to dryness under vacuum. The dried residue was reconstituted with 100 μL of methanol and vortexed at 2,000 rpm for 5 min. After centrifugation at 13,600 rpm for 10 min at 4°C, 5 μL of the supernatant was injected into the UPLC-LTQ orbitrap-MS/MS system for analysis. To study tissue distribution, each weighed tissue sample was homogenized in ice-cold physiological saline solution (1 : 2, w/v). The subsequent steps were identical to those described above for the treatment of plasma samples. For the excretion study, the fecal samples were pulverized with a mortar and pestle and homogenized in ice-cold physiological saline solution (1 : 2, w/v). Bile, urine, and homogenized fecal samples were processed using a similar method as the plasma samples.

Method Validation

The method was validated for specificity, linearity, sensitivity, accuracy and precision, extraction recovery, and matrix effect according to the FDA guidance for industry on bioanalytical method validation. Specificity was assessed by comparing the chromatograms of the standard-spiked samples with the biosamples from six different sources. Calibration curves were plotted as the peak area ratio (drug/IS) versus the pinostrobin nominal concentration using weighted least-squares linear regression analysis. The lower limit of quantification was defined as the lowest concentration on the calibration curve and evaluated by analyzing the samples prepared in six replicates. Precision and accuracy were assessed with six replicates at four QC levels in three separate runs for three consecutive days using calibration curves that were established daily. Precision was expressed as the relative standard deviation (RSD, %), and accuracy was expressed as the relative error (RE %). Matrix effects are usually due to the influence of coeluting compounds on the actual analyte ionization process (Cappiello et al., 2008). Blank biosamples were used to evaluate the matrix effect at three QC levels ($n = 6$). The matrix effect was calculated by comparing the peak area ratio of the analyte relative to the IS in the analyte-spiked postextracted sample with that acquired using a neat solution (Zeng et al., 2019). Extraction recoveries of pinostrobin through the protein precipitation procedure were also evaluated at three QC concentrations ($n = 6$) and determined by comparing the peak areas obtained from blank biological matrices spiked with analyte before and after extraction.

Metabolite Profiling

Metabolites were profiled using samples from the pharmacokinetics, tissue distribution, and excretion studies. Metabolite identification was performed using MetWorks software (Version 1.3 SP4, Thermo Fisher Scientific, San Jose, CA, United States) based on retention times, chemical

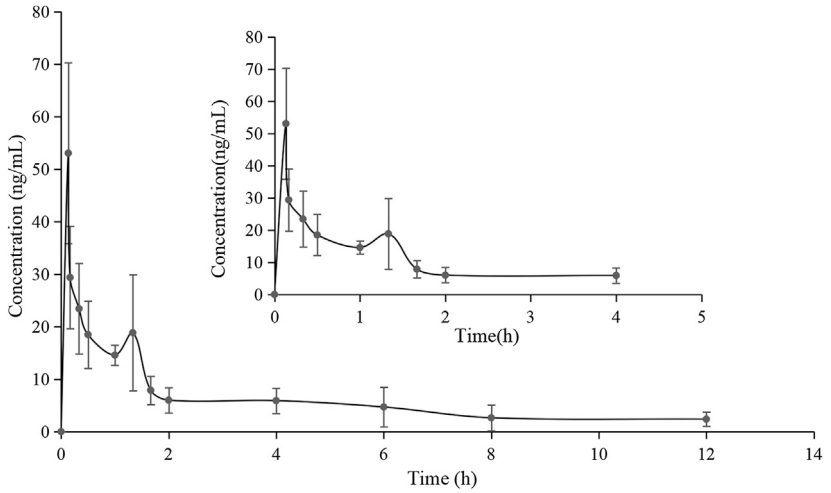

FIGURE 2 | Mean plasma concentration-time curve of pinostrobin after oral administration of 48.51 mg/kg pinostrobin (n = 5, mean ± SD).

composition, and fragmentation patterns on UPLC-LTQ orbitrap-MS/MS and compared with the available standards and literature to describe the metabolic profiles of pinostrobin in plasma, urine, feces, bile, and tissues.

Data Analysis

The pharmacokinetics were calculated using Kinetica software (Version 5.1, Thermo Fisher Scientific, San Jose, CA, United States) with a noncompartmental statistical model. The peak concentration (C_{max}) and time to reach C_{max} (T_{max}) were obtained from actual data. The results are expressed as the mean ± standard deviation (SD).

TABLE 1 | Noncompartmental pharmacokinetic parameters of pinostrobin in rats after oral administration of 48.51 mg/kg pinostrobin (n = 5, mean ± SD).

Pharmacokinetic parameters	Unit	Value
$t_{1/2}$	h	4.047 ± 1.843
T_{max}	h	0.133 ± 0
V_z	L/kg	627.480 ± 111.057
CL	L/h/kg	126.278 ± 51.962
C_{max}	ng/mL	53.034 ± 15.407
MRT_{0-t}	h	3.437 ± 0.896
$MRT_{0-\infty}$	h	5.906 ± 2.056
AUC_{0-24h}	ng/mL h	721.659 ± 197.849
$AUC_{0-\infty}$	ng/mL h	881.114 ± 289.587

RESULTS

Method Validation

Owing to the high selectivity and specificity of the SIM mode, no significant endogenous interference was observed at the retention times of pinostrobin (7.17 min) and IS (3.50 min). The calibration curves for all analytes showed good linearity ($r^2 > 0.9923$) over the concentration ranges. Intra- and interday precisions for LLOQ were less than 20%, while those for low, middle, and high QC were within 15%. Intra- and interday accuracies for LLOQ were within −13.0% to 15.3%, while those for low, middle, and high QC were within −13.2% to 10.4%. Extraction recoveries ranged from 81.8% to 107.1%, with an RSD% less than 14.7%. Matrix effects ranged from 81.0% to 106.6%, with an RSD% less than 12.1%. These results (**Supplementary Figures S2–S13** and **Supplementary Tables S1–S3**) indicated that the developed method was reproducible, accurate, and reliable for quantitative analysis of pinostrobin.

Pharmacokinetics Study

The validated method for the quantitation of pinostrobin in rat plasma was applied for the pharmacokinetic study in rats after oral administration of pinostrobin at a dose of 48.51 mg/kg. The mean plasma concentration-time curves are shown in **Figure 2**. The corresponding pharmacokinetic parameters calculated with noncompartmental analysis are listed as the mean ± SD and are shown in **Table 1**. The results showed that T_{max} was 0.133 h after oral administration in rats, and C_{max} of pinostrobin was 53.034 ng/mL ± 15.407 ng/mL. The apparent elimination half-life ($t_{1/2}$) was 4.047 h ± 1.843 h, indicating that pinostrobin had a moderate $t_{1/2}$. The mean residence time ($MRT_{0-\infty}$) was 5.906 h ± 2.056 h. The AUC_{0-12h} and $AUC_{0-\infty}$ were 721.659 ± 197.849 and 881.114 ± 289.587, respectively. The large apparent volume of distribution (V_z) was 627.480 L/kg ± 111.057 L/kg.

Tissue Distribution

In the distribution study, the concentrations of pinostrobin were determined in multiple tissues within 24 h after a single oral administration. Tissue distribution profiles of pinostrobin in rats at different time points are shown in **Figure 3**. The C_{max}, T_{max}, and AUC for pinostrobin in tissues are shown in **Supplementary Table S4**. The results indicated that pinostrobin was widely distributed in the gastrointestinal tract and major organs. The highest concentration was observed in the stomach at 2 h,

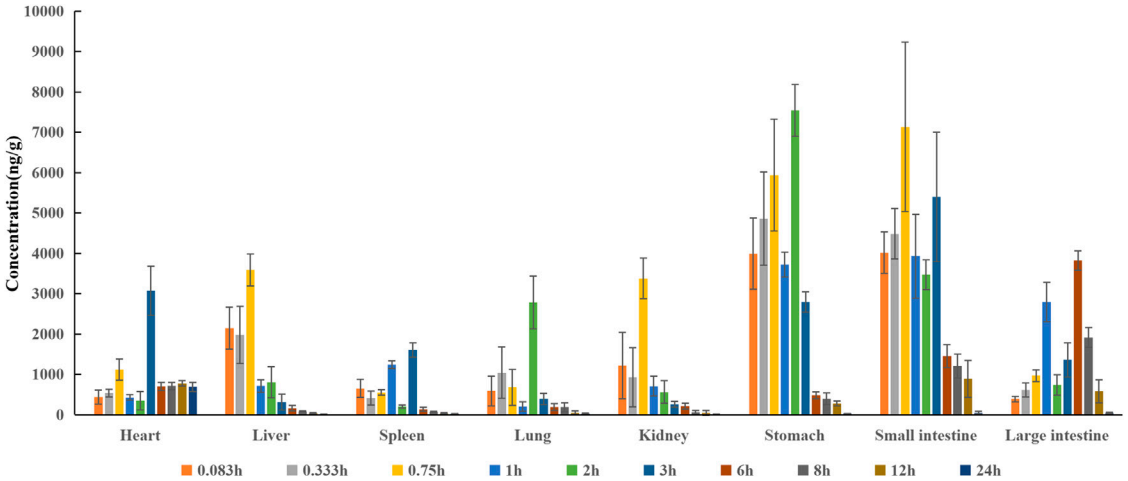

FIGURE 3 | Tissue distribution profile of pinostrobin at different time points after a single oral administration of 48.51 mg/kg pinostrobin (n = 5, mean ± SD).

followed by the small intestine, large intestine, liver, kidney, heart, lung, and spleen. The accumulation of pinostrobin in tissues was in the order of small intestine > large intestine > stomach > heart > lung > liver > spleen > kidney, suggesting that the gastrointestinal tract may be the main target organ of pinostrobin. Combined with the previous pharmacological activities of antiulcer, anti-*HP*, anti-inflammatory, and antioxidation of pinostrobin, this suggests that it may be an effective component of traditional Chinese medicines for the treatment of peptic ulcer.

Metabolite Profiling of Pinostrobin

Thirty flavonoid metabolites were identified or partially identified in all biosamples. Their chromatograms are shown in **Supplementary Figures S14–S17**. The maximum peak area ratio of metabolites relative to the internal standard (R_{max}) was used to estimate the content of these metabolites quantitatively. Their retention times, semiquantitative results, and fragment information are shown in **Table 2**. The proposed metabolic pathways of the above metabolites are illustrated in **Figure 4**. For these flavonoid metabolites, the subsequent loss of H_2O, CO, and Retro Diels–Alder (RDA) reaction was observed in MS/MS fragmentation (**Table 2**), which is in line with reported results (Sun et al., 2016). A mass spectrum peak gave rise to $[M + H]^+$ ions at m/z 271.09607 ($C_{16}H_{15}O_4$) at a retention time of 7.14 min with the same fragmentation pathways as the pinostrobin standard, identified as the parent drug, pinostrobin. MS/MS fragmentation showed the ions at m/z 253.0854 (−18 Da, loss of H_2O), $^{2.4}A^+$ ions at m/z 229.0854, $^{1.3}A^+$ ions at m/z 167.0333, $^{1.4}B^+$ ions at m/z 131.0487, $^{5}A^+$ ions at m/z 193.0488, and m/z 103.0539 (−28 Da, product of the fragment ions at m/z 131.0487 through the loss of CO) (Sun et al., 2016).

Metabolites M25 and M26 eluted at 2.94 min and 3.09 min, respectively, with the same $[M + H]^+$ ions at m/z 447.12780 ($C_{22}H_{23}O_{10}$), 176 Da higher than pinostrobin. The prominent fragment ion at m/z 271.0961 ($C_{16}H_{15}O_4$) was obtained by losing a terminal glucose group ($C_6H_8O_6$). The other product ions were similar to those of pinostrobin, indicating glucuronidation of pinostrobin. Metabolites M24 and M28 eluted at 2.82 min and 4.08 min, respectively, with the same $[M + H]^+$ ions at m/z 351.0532 ($C_{16}H_{15}O_7S$), 80 Da higher than pinostrobin. The neutral loss of a sulfate group (SO_3) yielded the major fragment ion at m/z 271.0962 ($C_{16}H_{15}O_4$), and the other ions at m/z 253.0852 and $^{2.4}A^+$ ions at m/z 229.0861, $^{1.3}A^+$ ions at m/z 167.0335, $^{1.4}B^+$ ions at m/z 131.0487, and $^{5}A^+$ ions at m/z 193.0492 and m/z 103.0540 were also observed, indicating sulfation of pinostrobin. Metabolites M1 and M20 had a retention time of 0.76 min and 2.57 min, respectively, with the same $[M + H]^+$ ions at m/z 328.1181 ($C_{18}H_{18}O_5N$), 57 Da higher than pinostrobin. The major fragment ion at m/z 271.0969 ($C_{16}H_{15}O_4$) was produced by losing a glycine group (C_2H_3NO), and the other fragment ions at 167.0334 and 131.0483 were similar to those of pinostrobin, indicating glycine conjugation of pinostrobin. Metabolites M15 and M29 eluted at 2.18 and 4.55 min, respectively, with the same $[M + H]^+$ ions at m/z 432.1105 ($C_{21}H_{22}O_7NS$), 161 Da higher than pinostrobin. The fragment ion at m/z 271.0969 ($C_{16}H_{15}O_4$) was due to the loss of a glycine group ($C_5H_7NO_3S$). The other fragment ions at 167.0333 and 131.0486 were similar to those of pinostrobin, suggesting glycine conjugation of pinostrobin. Metabolite M14 showed $[M + H]^+$ ions at m/z 378.1000 ($C_{18}H_{20}O_6NS$), with a retention time of 2.07 min. The fragment ion at m/z 271.0956 ($C_{16}H_{15}O_4$) was produced through the loss of a taurine group ($C_2H_5NO_2S$, −107 Da), and mass signals at 167.0334 and 131.0486 were produced by $^{1.3}A^+$ and $^{1.4}B^+$ ions, which were from the ions at m/z 271.0956, indicating taurine conjugation of pinostrobin.

Metabolite M23 eluted at 2.77 min, with the $[M + H]^+$ ions at m/z 243.1008 ($C_{15}H_{15}O_3$), 28 Da (CO) lower than pinostrobin. The signals at m/z 123.0439, 107.0489, and 137.0589 demonstrated the existence of $^{0.3}A^+$, $^{0.2}B^+$, and $^{0.2}A^+$ fragment ions, suggesting decarbonylation on the C ring. Metabolites M4 and M18 showed the same $[M + H]^+$ ions at m/z 419.1327

TABLE 2 | The metabolites in rats after oral administration of 48.51 mg/kg pinostrobin.

Metabolites	RT (min)	$[M + H]^+$	Chemical formula	ppm	Mass shift	Formula change	Metabolic pathway	MS^n m/z	Source (T_{max}, R_{max})[a]
Parent drug	7.14	271.09607	$C_{16}H_{15}O_4$	−1.532				253.08542[M + H-H_2O]$^+$, 229.08537[M + H-C_2H_2O]$^+$, 193.04883[M + H-C_6H_6]$^+$, 173.05914 [M + H-C_2H_2O-C_3H_4O]$^+$, 167.03336[M + H-C_8H_8]$^+$, 131.04871[M + H-$C_7H_8O_3$]$^+$, 103.05387[M + H-$C_7H_8O_3$-CO]$^+$	Plasma, urine, feces, bile, heart, liver, spleen, lung, kidney, stomach, small intestine, and large intestine
M1	0.76	328.11810	$C_{18}H_{18}O_5N$	0.460	57.0215	[M-OH + $C_2H_4NO_2$]	Glycine conjugation	271.09689[M + H-C_2H_3NO]$^+$, 167.03342[M + H-C_2H_3NO-C_8H_8]$^+$, 131.04826[M + H-C_2H_3NO-$C_7H_8O_3$]$^+$, 103.05365[M + H-C_2H_3NO-$C_7H_8O_3$-CO]$^+$	Urine (8–12, 3.563), liver (0.083, 0.471), kidney (0.75, 0.075), and plasma (2, 0.011)
M2	0.84	433.11163	$C_{21}H_{21}O_{10}$	−2.986	162.0164	[M-CH_2 + $C_6H_8O_6$]	Demethylation and glucuronidation	415.09778[M + H-H_2O]$^+$, 257.08044[M + H-$C_6H_8O_6$]$^+$, 131.04860[M + H-$C_6H_8O_6$-$C_6H_6O_3$]$^+$	Urine (6–8, 2.671), bile (12–16, 0.539), small intestine (8, 0.090), lung (8, 0.062), and liver (0.33, 0.039)
M3	0.98	463.12338	$C_{22}H_{23}O_{11}$	−0.233	192.0270	[M + $C_6H_8O_7$]	Hydroxylation andglucuronidation	287.09042[M + H-$C_6H_8O_6$]$^+$, 269.07990[M + H-$C_6H_8O_6$-H_2O]$^+$, 255.06409[M + H-$C_6H_8O_6$-CH_4O]$^+$, 183.02962[M + H-$C_6H_8O_6$-C_8H_8]$^+$, 131.04849[M + H-$C_6H_8O_6$-$C_7H_8O_4$]$^+$	Bile (16–20, 0.447); liver (0.33, 0.014)
M4	1.00	419.13272	$C_{21}H_{23}O_9$	−2.240	148.0372	[M + $C_5H_8O_5$]	Decarboxylation and glucuronidation	401.12317[M + H-H_2O]$^+$, 383.11270[M + H-$_2H_2O$]$^+$, 243.10162[M + H-$C_6H_8O_6$]$^+$, 137.05962[M + H-$C_6H_8O_6$-C_7H_6O]$^+$, 133.06473 [M + H-$C_6H_8O_6$-$C_6H_6O_2$]$^+$, 123.04401[M + H-$C_6H_8O_6$-C_8H_8O]$^+$, 107.04910[M + H-$C_6H_8O_6$-$C_8H_8O_2$]$^+$	Urine (12–24, 4.916), plasma (12, 0.040), kidney (12, 0.015), and lung (12, 0.011)
M5	1.04	273.11218	$C_{16}H_{17}O_4$	0.163	2.0157	[M + H_2]	Hydrogenation	243.10094[M + H-CH_2O]$^+$, 105.03316[M + H-$C_{11}H_{12}O_3$]$^+$, 91.05390[M + H-$C_9H_{10}O_4$]$^+$	Urine (8–12, 1.236), kidney (12, 0.077), and bile (0–3, 0.031)
M6	1.13	301.07001	$C_{16}H_{13}O_6$	−2.174	29.9742	[M-H_2 + O_2]	Hydroxylation and ketone	271.05948[M + H-CH_2O]$^+$, 131.04871[M + H-CH_2O-$C_7H_8O_3$]$^+$	Urine (24–36, 3.059), stomach (1, 0.016), heart (0.083, 0.014), plasma (1.67, 0.010), lung (24, 0.009), and spleen (24, 0.009)
M7	1.17	313.10632	$C_{18}H_{17}O_5$	−2.332	42.0106	[M + $COCH_2$]	Acetylation	295.09641[M + H-H_2O]$^+$, 277.08640[M + H-$2H_2O$]$^+$, 267.10165[M + H-H_2O-CO]$^+$, 253.08600 [M + H-H_2O-C_2H_2O]$^+$, 249.09061[M + H-H_2O-CO-H_2O]$^+$, 130.03491[M + H-$C_9H_{11}O_4$]$^+$	Feces (4–6, 12.513), large intestine (0.75, 0.423), urine (4–6, 0.352), stomach (0.75, 0.238), small intestine (0.083, 0.091), spleen (0.75, 0.031), and heart (12, 0.008)
M8	1.53	255.06461	$C_{15}H_{11}O_4$	−2.256	−16.0308	[M-CH_2-H_2]	Demethylation and dehydrogenation	237.05391[M + H-H_2O]$^+$, 227.06953[M + H-CO]$^+$, 209.05901[M + H-H2O-CO]$^+$, 199.07465[M + H-2CO]$^+$, 181.06406[M + H-2CO-H_2O]$^+$, 157.06416 [M + H-2CO-C_2H_2O]$^+$, 145.02780[M + H-$C_6H_6O_2$]$^+$, 137.02280[M + H-C_8H_6O]$^+$	Urine (8–12, 108.404), small intestine (0.33, 8.065), stomach (0.33, 4.531), feces (2–4, 3.038), plasma (4, 0.504), kidney (12, 0.389), bile (0–3, 0.362), heart (0.33, 0.313), lung (0.33, 0.277), spleen (0.083, 0.257), liver (0.33, 0.115), and large intestine (0.33, 0.110)
M9	1.55	463.12256	$C_{22}H_{23}O_{11}$	−2.004	192.0270	[M + $C_6H_8O_7$]	Hydroxylation and glucuronidation	287.09058[M + H-$C_6H_8O_6$]$^+$, 269.08008[M + H-$C_6H_8O_6$-H_2O]$^+$, 255.06412[M + H-$C_6H_8O_6$-CH_4O]$^+$, 183.02753[M + H-$C_6H_8O_6$-C_8H_8]$^+$, 131.04784[M + H-$C_6H_8O_6$-$C_7H_8O_4$]$^+$	Small intestine (8, 0.335), liver (0.33, 0.219), and kidney (0.75, 0.132)

(Continued on following page)

TABLE 2 | (*Continued*) The metabolites in rats after oral administration of 48.51 mg/kg pinostrobin.

Metabolites	RT (min)	$[M + H]^+$	Chemical formula	ppm	Mass shift	Formula change	Metabolic pathway	MS^n m/z	Source (T_{max}, R_{max})[a]
M10	1.63	285.07507	$C_{16}H_{13}O_5$	−2.385	13.9793	[M-H_2 + O]	Hydroxylation and dehydrogenation	270.05151[M + H-CH_3]$^+$, 267.06436[M + H-H_2O]$^+$, 257.08008[M + H-CO]$^+$, 229.08514[M + H-2CO]$^+$, 225.05389[M + H-CO-CH_4O]$^+$, 197.05893[M + H-CO-CH_4O-CO]$^+$, 167.03322[M + H-C_8H_6O]$^+$, 145.02779[M + H-$C_7H_8O_3$]$^+$, 137.02269[M + H-$C_9H_8O_2$]$^+$	Urine (24–36, 107.577), small intestine (0.33, 2.346), feces (8–12, 2.019), stomach (0.33, 1.195), plasma (4, 0.200), kidney (12, 0.115), lung (0.33, 0.090), heart (0.33, 0.077), bile (12–16, 0.067), spleen (0.33, 0.060), and large intestine (0.083, 0.030)
M11	1.93	257.08035	$C_{15}H_{13}O_4$	−1.771	−14.0157	[M-CH_2]	Demethylation	215.06979[M + H-C_2H_2O]$^+$, 179.03320[M + H-C_6H_6]$^+$, 173.05905[M + H-2C_2H_2O]$^+$, 153.01764[M + H-C_8H_8]$^+$, 131.04865[M + H-$C_6H_6O_3$]$^+$	Urine (0–2, 6.840), feces (36–54, 1.937), plasma (6, 0.674), bile (12–16, 0.221), stomach (12, 0.186), kidney (12, 0.153), liver (6, 0.121), heart (2, 0.121), spleen (24, 0.112), large intestine (0.75, 0.109), lung (1, 0.108), and small intestine (24, 0.103)
M12	1.95	285.07562	$C_{16}H_{13}O_5$	−0.456	13.9793	[M-H_2 + O]	Hydroxylation anddehydrogenation	257.07993[M + H-CO]$^+$, 167.03354[M + H-C_8H_6O]$^+$	Urine (8–12, 30.236); bile (0–3, 0.305)
M13	1.99	433.11176	$C_{21}H_{21}O_{10}$	−2.686	162.0164	[M-CH_2 + $C_6H_8O_6$]	Demethylation and glucuronidation	257.07996[M + H-$C_6H_8O_6$]$^+$, 215.06911[M + H-$C_6H_8O_6$-C_2H_2O]$^+$, 153.01761[M + H-$C_6H_8O_6$-C_8H_8]$^+$, 131.04823[M + H-$C_6H_8O_6$-$C_6H_6O_3$]$^+$, 103.05369[M + H-$C_6H_8O_6$-$C_6H_6O_3$-CO]$^+$	Bile (16–20, 0.647), small intestine (8, 0.204), and urine (12–24, 0.114)
M14	2.07	378.10001	$C_{18}H_{20}O_6NS$	−1.519	107.0041	[M + $C_2H_5NO_2S$]	Taurine conjugation	271.09558[M + H-$C_2H_5NO_2S$]$^+$, 167.03336[M + H-$C_2H_5NO_2S$-C_8H_8]$^+$, 131.04861[M + H-$C_2H_5NO_2S$-$C_7H_8O_3$]$^+$	Urine (0–2, 0.502), heart (3–6, 0.146), spleen (0.75, 0.071), plasma (4, 0.053), small intestine (3, 0.027), and stomach (0.33, 0.008)
M15	2.18	432.11047	$C_{21}H_{22}O_7NS$	−1.572	161.0147	[M + $C_5H_7NO_3S$]	*N*-Acetylcysteine conjugation	399.13025[M + H-HS]$^+$, 381.11737[M + H-HS-H_2O]$^+$, 271.09607[M + H-$C_5H_7NO_3S$]$^+$, 173.05907[M + H-$C_5H_7NO_3S$-C_2H_2O-C_3H_4O]$^+$, 167.03336[M + H-$C_5H_7NO_3S$-C_8H_8]$^+$, 131.04878[M + H-$C_5H_7NO_3S$-$C_7H_8O_3$]$^+$, 103.05384[M + H-$C_5H_7NO_3S$-$C_7H_8O_3$-CO]$^+$	Urine (0–2, 0.792)
M16	2.24	287.09088	$C_{16}H_{15}O_5$	−1.811	15.9949	[M + O]	Hydroxylation	269.08017[M + H-H_2O]$^+$, 255.06470[M + H-CH_4O]$^+$, 245.08011[M + H-C_2H_2O]$^+$, 183.02817[M + H-C_8H_8]$^+$, 173.05904[M + H-C_2H_2-$C_3H_4O_2$]$^+$, 131.04868[M + H-$C_7H_8O_4$]$^+$, 103.05385[M + H-$C_7H_8O_4$-CO]$^+$	Urine (0–2, 59.881), bile (0–3, 21.765), small intestine (3, 3.497), plasma (6, 3.053), kidney (0.75, 1.180), liver (3, 0.321), lung (0.083, 0.082), heart (0.33, 0.072), and stomach (0.33, 0.038)
M17	2.35	463.12259	$C_{22}H_{23}O_{11}$	−1.939	192.0270	[M + $C_6H_8O_7$]	Hydroxylation and glucuronidation	445.11160[M + H-H_2O]$^+$, 427.10352[M + H-2H_2O]$^+$, 287.09064[M + H-$C_6H_8O_6$]$^+$, 269.08014[M + H-$C_6H_8O_6$-H_2O]$^+$, 255.06468[M + H-$C_6H_8O_6$-CH_4O]$^+$, 245.08002[M + H-$C_6H_8O_6$-C_2H_2O]$^+$, 183.02814[M + H-$C_6H_8O_6$-C_8H_8]$^+$, 173.05910[M + H-$C_6H_8O_6$-C_2H_2-$C_3H_4O_2$]$^+$, 131.04861[M + H-$C_6H_8O_6$-$C_7H_8O_4$]$^+$, 103.05376[M + H-$C_6H_8O_6$-$C_7H_8O_4$-CO]$^+$	Urine (0–2, 109.471), bile (0–3, 31.084), plasma (6, 6.648), small intestine (3, 5.256), kidney (0.75, 1.794), liver (3, 0.457), heart (0.33, 0.135), stomach (8, 0.116), lung (0.083, 0.112), and spleen (6, 0.016)
M18	2.36	419.13333	$C_{21}H_{23}O_9$	−0.784	148.0372	[M + $C_5H_8O_5$]	Decarboxylation and glucuronidation	243.10045[M + H-$C_6H_8O_6$]$^+$, 137.05922[M + H-$C_6H_8O_6$-C_7H_6O]$^+$, 123.04356[M + H-$C_6H_8O_6$-C_8H_8O]$^+$, 107.04873[M + H-$C_6H_8O_6$-$C_8H_8O_2$]$^+$	Urine (2–4, 0.546)

(Continued on following page)

TABLE 2 | (*Continued*) The metabolites in rats after oral administration of 48.51 mg/kg pinostrobin.

Metabolites	RT (min)	$[M + H]^+$	Chemical formula	ppm	Mass shift	Formula change	Metabolic pathway	MS^n m/z	Source (T_{max}, R_{max})[a]
M19	2.45	433.11200	$C_{21}H_{21}O_{10}$	−2.132	162.0164	[M-CH_2 + $C_6H_8O_6$]	Demethylation and glucuronidation	415.10141[M + H-H_2O]$^+$, 397.09091[M + H-2H_2O]$^+$, 257.08026[M + H-$C_6H_8O_6$]$^+$, 239.06920 [M + H-$C_6H_8O_6$-H_2O]$^+$, 215.06953[M + H-$C_6H_8O_6$-C_2H_2O]$^+$, 179.03311[M + H-$C_6H_8O_6$-C_6H_6]$^+$, 173.05902[M + H-$C_6H_8O_6$-2C_2H_2O]$^+$, 153.01761 [M + H-$C_6H_8O_6$-C_8H_8]$^+$, 145.06438[M + H-$C_6H_8O_6$-2C_2H_2O-CO]$^+$, 131.04861[M + H-$C_6H_8O_6$-$C_6H_6O_3$]$^+$, 103.05379[M + H-$C_6H_8O_6$-$C_6H_6O_3$-CO]$^+$	Urine (0–2, 203.526), plasma (4, 25.663), bile (0–3, 18.171), kidney (0.75, 4.648), small intestine (3, 3.140), heart (0.75, 0.420), liver (3, 0.241), lung (0.083, 0.186), stomach (8, 0.158), and spleen (2, 0.009)
M20	2.57	328.11697	$C_{18}H_{18}O_5N$	−2.984	57.0215	[M-OH + $C_2H_4NO_2$]	Glycine conjugation	271.09583[M + H-C_2H_3NO]$^+$, 167.03299[M + H-C_2H_3NO-C_8H_8]$^+$, 131.04858[M + H-C_2H_3NO-$C_7H_8O_3$]$^+$	Bile (3–5, 0.509), urine (8–12, 0.461), small intestine (0.75, 0.298), plasma (4, 0.165), spleen (0.75, 0.103), liver (3, 0.102), stomach (24, 0.069), kidney (0.75, 0.041), large intestine (8, 0.033), heart (3, 0.027), and lung (8, 0.020)
M21	2.64	271.05951	$C_{15}H_{11}O_5$	−2.176	−0.0364	[M-CH_4 + O]	Demethylation and methylene to ketone	253.04883[M + H-H_2O]+, 243.06450[M + H-CO]+, 229.04872[M + H-C_2H_2O]+, 225.05394[M + H-CO-H_2O]$^+$, 215.06960[M + H-2CO]$^+$, 197.05894[M + H-2CO-H_2O]$^+$, 187.03836[M + H-2C_2H_2O]$^+$, 169.06418[M + H-3CO-H_2O]$^+$, 159.04343[M + H-2C_2H_2O-CO]$^+$, 153.01765[M + H-C_8H_6O]+, 149.02278[M + H-$C_7H_6O_2$]+, 145.02782[M + H-$C_8H_6O_3$]$^+$, 119.04878[M + H-$C_7H_4O_4$]$^+$	Urine (24–36, 200.177), small intestine (0.33, 5.512), stomach (0.33, 4.903), feces (2–4, 2.119), liver (0.33, 0.425), heart (0.33, 0.365), lung (0.33, 0.271), spleen (0.75, 0.221), plasma (4, 0.179), large intestine (8, 0.108), and spleen (6, 0.064)
M22	2.72	273.07455	$C_{15}H_{13}O_5$	−4.394	1.9793	[M-CH_2 + O]	Demethylation and hydroxylation	153.01706[M + H-C_8H_8O]$^+$, 121.06379[M + H-$C_7H_4O_4$]$^+$	Feces (8–12, 1.467), liver (0.33, 1.427), small intestine (6, 1.101), large intestine (8, 0.483), urine (4–6, 0.457), stomach (0.75, 0.249), kidney (0.75, 0.243), spleen (0.75, 0.220), lung (0.75, 0.134), and heart (2, 0.021)
M23	2.77	243.10080	$C_{15}H_{15}O_3$	−3.171	−27.9949	[M-CO]	Decarboxylation	137.05891[M + H-C_7H_6O]$^+$, 133.06447[M + H-$C_6H_6O_2$]$^+$, 123.04386[M + H-C_8H_8O]$^+$, 107.04888[M + H-$C_8H_8O_2$]$^+$	Urine (36–54, 136.900), feces (0–2, 0.270), large intestine (2, 0.222), stomach (6, 0.179), small intestine (12, 0.100), liver (2, 0.059), kidney (12, 0.052), spleen (6, 0.038), and lung (12, 0.020)
M24	2.82	351.05319	$C_{16}H_{15}O_7S$	−0.313	79.9568	[M + SO_3]	Sulfation	271.09619[M + H-SO_3]$^+$, 253.08521[M + H-SO_3-H_2O]$^+$, 229.08615[M + H-SO_3-C_2H_2O]$^+$, 193.04921[M + H-SO_3-C_6H_6]$^+$, 173.05945[M + H-SO_3-C_2H_2O-C_3H_4O]$^+$, 167.03351[M + H-SO_3-C_8H_8]$^+$, 131.04875[M + H-SO_3-$C_7H_8O_3$]$^+$, 103.05399[M + H-SO_3-$C_7H_8O_3$-CO]$^+$	Plasma (2, 0.119); urine (0–2, 0.064)

(Continued on following page)

TABLE 2 | (*Continued*) The metabolites in rats after oral administration of 48.51 mg/kg pinostrobin.

Metabolites	RT (min)	$[M + H]^+$	Chemical formula	ppm	Mass shift	Formula change	Metabolic pathway	MS^n m/z	Source (T_{max}, R_{max})[a]
M25	2.94	447.12799	$C_{22}H_{23}O_{10}$	−1.305	176.0321	[M + $C_6H_8O_6$]	Glucuronidation	429.11697[M + H-H_2O]$^+$, 411.10663[M + H-2H_2O]$^+$, 313.10654[M + H-$C_4H_6O_5$]$^+$, 271.09592 [M + H-$C_6H_8O_6$]$^+$, 253.08542[M + H-$C_6H_8O_6$-H_2O]$^+$, 229.08516[M + H-$C_6H_8O_6$-C_2H_2O]$^+$, 193.04890[M + H-$C_6H_8O_6$-C_6H_6]$^+$, 173.05908[M + H-$C_6H_8O_6$-C_2H_2O-C_3H_4O]$^+$, 167.03326[M + H-$C_6H_8O_6$-C_8H_8]$^+$, 131.04866[M + H-$C_6H_8O_6$-$C_7H_8O_3$]$^+$, 103.05376[M + H-$C_6H_8O_6$-$C_7H_8O_3$-CO]$^+$	Urine (0–2, 406.932), bile (0–3, 174.894), plasma (0.133, 59.330), small intestine (3, 25.059), liver (0.083, 7.034), kidney (0.75, 7.225), stomach (0.33, 1.090), heart (0.33, 0.967), lung (0.083, 0.538), and spleen (0.33, 0.039)
M26	3.09	447.12772	$C_{22}H_{23}O_{10}$	−1.908	176.0321	[M + $C_6H_8O_6$]	Glucuronidation	429.11682[M + H-H_2O]$^+$, 411.10645[M + H-2H_2O]$^+$, 313.10657[M + H-$C_4H_6O_5$]$^+$, 271.09586 [M + H-$C_6H_8O_6$]$^+$, 253.08565[M + H-$C_6H_8O_6$-H_2O]$^+$, 229.08559[M + H-$C_6H_8O_6$-C_2H_2O]$^+$, 193.04910[M + H-$C_6H_8O_6$-C_6H_6]$^+$, 173.05930[M + H-$C_6H_8O_6$-C_2H_2O-C_3H_4O]$^+$, 167.03351[M + H-$C_6H_8O_6$-C_8H_8]$^+$, 131.04883[M + H-$C_6H_8O_6$-$C_7H_8O_3$]$^+$, 103.05399[M + H-$C_6H_8O_6$-$C_7H_8O_3$-CO]$^+$	Urine (0–2, 385.560), bile (0–3, 221.411), plasma (0.133, 92.104), small intestine (3, 35.601), liver (0.083, 14.678), kidney (0.75, 10.093), stomach (0.33, 1.536), heart (0.33, 1.508), lung (0.083, 0.767), and spleen (0.33, 0.086)
M27	3.14	367.04578	$C_{16}H_{15}O_8S$	−6.633	95.9517	[M + O_4S]	Hydroxylation and sulfation	287.09122[M + H-SO_3]$^+$, 269.08093[M + H-SO_3-H_2O]$^+$, 255.06427[M + H-SO_3-CH_4O]$^+$, 131.04861 [M + H-SO_3-$C_7H_8O_4$]$^+$	Plasma (0.167, 0.157), urine (0–2, 0.134), liver (0.33, 0.023), kidney (0.33, 0.022), and lung (0.083, 0.006)
M28	4.08	351.05292	$C_{16}H_{15}O_7S$	−1.082	79.9568	[M + SO_3]	Sulfation	271.09613[M + H-SO_3]+, 253.08557[M + H-SO_3-H_2O]$^+$, 229.08524[M + H-SO_3-C_2H_2O]$^+$, 193.04897[M + H-SO_3-C_6H_6]$^+$, 173.05910[M + H-SO_3-C_2H_2O-C_3H_4O]$^+$, 167.03342[M + H-SO_3-C_8H_8]$^+$, 131.04875[M + H-SO_3-$C_7H_8O_3$]$^+$, 103.05390[M + H-SO_3-$C_7H_8O_3$-CO]$^+$	Plasma (0.167, 1.786); urine (0–2, 0.554)
M29	4.55	432.11029	$C_{21}H_{22}O_7NS$	1.988	161.0147	[M + $C_5H_7NO_3S$]	*N*-Acetylcysteine conjugation	271.09592[M + H-$C_5H_7NO_3S$]$^+$, 173.05887[M + H-$C_5H_7NO_3S$-C_2H_2O-C_3H_4O]$^+$, 167.03326[M + H-$C_5H_7NO_3S$-C_8H_8]$^+$, 131.04874[M + H-$C_5H_7NO_3S$-$C_7H_8O_3$]$^+$	Urine (0–2, 0.289)
M30	5.01	257.07980	$C_{15}H_{13}O_4$	−4.028	−14.0157	[M-CH_2]	Demethylation	239.07014[M + H-H_2O]+, 215.06995[M + H-C_2H_2O]$^+$, 179.03355[M + H-C_6H_6]$^+$, 173.05936 [M + H-2C_2H_2O]$^+$, 153.01790[M + H-C_8H_8]$^+$, 145.06441[M + H-2C_2H_2O-CO]$^+$, 131.04884[M + H-$C_6H_6O_3$]+,103.05395[M + H-$C_6H_6O_3$-CO]$^+$	Urine (4–6, 17.160), small intestine (0.75, 15.426), liver (0.33, 11.587), large intestine (8, 9.070), kidney (0.75, 6.286), feces (12, 4.799), stomach (0.75, 2.849), spleen (0.75, 1.897), lung (0.75, 1.772), and heart (0.33, 0.858)

T_{max} is the time or time interval to reach the maximum peak area ratio of metabolites relative to the internal standard in biological samples. R_{max} is the maximum peak area ratio of metabolites relative to the internal standard.
[a]Sorted by R_{max} from large to small.

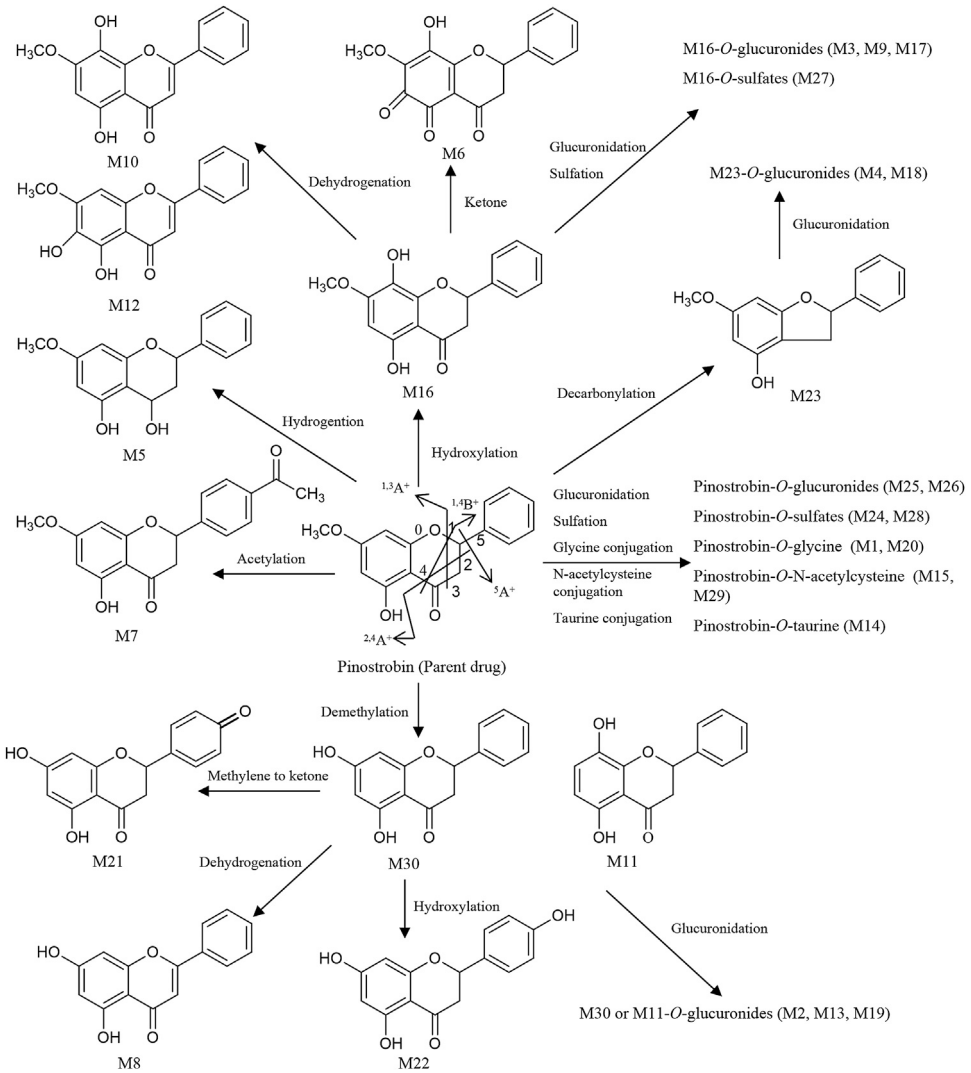

FIGURE 4 | Proposed metabolic pathways of the metabolites in rats after oral administration of 48.51 mg/kg pinostrobin.

($C_{21}H_{23}O_9$), with a retention time of 1.00 and 2.36 min, respectively, which was 176 Da higher than M23. The prominent fragment ion at *m/z* 243.1016 ($C_{15}H_{15}O_3$) was obtained through the loss of a glucose group ($C_6H_8O_6$). The other product ions were similar to that of M23, indicating decarbonylation and glucuronidation of pinostrobin.

Metabolite M16, with a retention time of 2.24 min, generated $[M + H]^+$ ions at *m/z* 287.0908 ($C_{16}H_{15}O_5$), which was 16 Da higher than that of pinostrobin. The product ion at *m/z* 269.0802 was obtained through the loss of H_2O. The RDA reaction of M16 generated $^{2.4}A^+$ ions at *m/z* 245.0801, $^{1.3}A^+$ ions at *m/z* 183.0282, and $^{1.4}B^+$ ions at *m/z* 131.0489. The fragment ions at *m/z* 131.0489 could be fragmented to produce the ion at *m/z* 103.0539 through the loss of CO. The fragment ion at *m/z* 183.0282 and 131.0488 showed hydroxylation on the A ring. Metabolites M10 and M12 eluted at 1.63 min and 1.95 min, respectively, with the same $[M + H]^+$ ions at *m/z* 285.0756 ($C_{16}H_{13}O_5$), 2 Da lower than M16. The fragment ions at *m/z* 167.0333 and 145.0277 demonstrated the existence of $^{1.3}A^+$ and $^{0.4}B^+$ fragment ions, suggesting hydroxylation of M16 on the C ring, which indicates the hydroxylation and dehydrogenation of pinostrobin. Metabolite M6 eluted at 1.13 min, with the $[M + H]^+$ ions at *m/z* 301.0700 ($C_{16}H_{13}O_6$), 14 Da (−2H + O) lower than that of pinostrobin. The signals at *m/z* 271.05948 and 131.04871 suggested hydroxylation and a ketone on the A ring of pinostrobin. Metabolites M3, M9, and M17 showed the same $[M + H]^+$ ions at *m/z* 463.1225 ($C_{22}H_{23}O_{11}$), with retention times of 0.98, 1.55, and 2.35 min, respectively, which was 176 Da higher than M16. The prominent fragment ion at *m/z* 287.0906 ($C_{16}H_{15}O_5$) was obtained through the loss of a glucose group ($C_6H_8O_6$). The other product ions were similar to that of M16, indicating hydroxylation and glucuronidation of pinostrobin. Metabolite M27 had a retention time of 3.14 min, with the $[M + H]^+$ ions at *m/z* 367.0458 ($C_{16}H_{15}O_8S$), 80 Da higher than M16. The neutral loss of a

sulfate group (SO_3) yielded the major fragment ion at *m/z* 287.0912 ($C_{16}H_{15}O_5$). The other ions at *m/z* 269.0809 and 131.0486 were also observed, indicating hydroxylation and sulfation of pinostrobin.

The metabolite M30 (pinocembrin), with a retention time of 5.01 min, generated $[M + H]^+$ ions at *m/z* 257.0798 ($C_{15}H_{13}O_4$), indicating demethylation of pinostrobin. The fragment ion at *m/z* 239.0701 ($C_7H_5O_4$) was obtained through the loss of H_2O. The RDA reaction of M30 generated $^{2.4}A^+$ ions at *m/z* 215.0699 ($C_{13}H_{11}O_3$), $^{1.3}A^+$ ions at *m/z* 153.0179 ($C_7H_5O_4$), and $^{1.4}B^+$ ions at *m/z* 131.0488 (C_9H_7O). The fragment ions at *m/z* 131.0488 could be fragmented to produce the ion at *m/z* 103.0540 (C_8H_7) through the loss of CO. The fragment ion at *m/z* 153.0179 and 131.0488 indicated demethylation on the A ring. The metabolite M11 generated the same $[M + H]^+$ ions at 257.0804 ($C_{15}H_{13}O_4$) and fragment ions at *m/z* 215.0698, 153.0176, and 131.0487 following an RDA reaction, which was identified as the isomer of M30. Metabolites M2, M13, and M19 showed the same $[M + H]^+$ ions at *m/z* 463.1225 ($C_{22}H_{23}O_{11}$), with a retention time of 0.84 min, 1.99 min, and 2.45 min, respectively, which was 176 Da higher than M30 or M11. The prominent fragment ion at *m/z* 257.0803 ($C_{15}H_{13}O_4$) was obtained through the loss of a glucose group ($C_6H_8O_6$). The other product ions were similar to those of M30 and M11, indicating demethylation and glucuronidation of pinostrobin. Metabolite M21 eluted at 2.64 min, with the $[M + H]^+$ ions at *m/z* 271.0595 ($C_{15}H_{11}O_5$), 14 Da (−2H + O) lower than M30. The fragment ions at *m/z* 153.0177, 119.0487, and 145.0278 demonstrated the existence of $^{1.3}A^+$, $^{1.3}B^+$, and $^{1.4}B^+$ fragment ions, suggesting conversion of the methylene to a ketone of M30 on the B ring, indicating the demethylation and methylene conversion to the ketone of pinostrobin. Metabolite M8 eluted at 1.53 min, with the $[M + H]^+$ ions at *m/z* 255.0646 ($C_{15}H_{11}O_4$), 2 Da lower than M30. The mass signals at *m/z* 145.0277 and 137.0228 demonstrated the existence of $^{0.4}B^+$ and $^{0.3}B^+$ fragment ions, suggesting dehydrogenation of M30 on the C ring, indicating the demethylation and dehydrogenation of pinostrobin. Metabolite M22 with a retention time of 2.72 min generated $[M + H]^+$ ions at *m/z* 273.0746 ($C_{15}H_{13}O_5$), 16 Da higher than M30. The characteristic fragment ions $^{1.3}A^+$ ions at *m/z* 153.0171 and $^{1.3}B^+$ ions at *m/z* 121.0638 showed hydroxylation of M30 on the B ring, suggesting the demethylation and hydroxylation of pinostrobin.

Metabolite M7 had a retention time of 1.17 min and showed $[M + H]^+$ ions at *m/z* 313.1063 ($C_{18}H_{17}O_5$), 42 Da ($COCH_2$) higher than pinostrobin. The fragment ion at *m/z* 130.03491 showed acetylation on the B ring. Metabolite M5 eluted at 1.04 min, with the $[M + H]^+$ ions at *m/z* 273.1121 ($C_{16}H_{17}O_4$), 2 Da higher than that of pinostrobin. The signals at *m/z* 91.0539 demonstrated the $^{1.2}B^+$ fragment ions, suggesting hydrogenation on the C ring.

Meanwhile, intestinal bacteria play an essential role in the metabolism of flavonoids, generating ring fission products (Liu et al., 2012; Zeng et al., 2018). In this study, 15 catabolites were detected (**Table 3**). Their chromatograms are shown in **Supplementary Figures S18, S19**. Pinostrobin was catabolized into C13 ($C_{10}H_9O_4$) by fission of the fifth bond as well as C4 ($C_6H_7O_3$) and C14 ($C_9H_9O_2$) by the fission of the first and fourth bonds. C4 was converted to C10 by demethylation. Meanwhile, C14 was hydroxylated, hydrogenated, and glycine-conjugated, yielding C9 ($C_9H_{11}O_3$), C12 ($C_9H_{11}O_2$), and C11 ($C_{11}H_{12}O_3N$). Then, C12 was glycine-conjugated and deethylated, generating C7 ($C_{10}H_{12}O_3N$) and C15 ($C_7H_7O_2$). C15 was glycine-conjugated to yield C6 ($C_9H_{10}O_3N$). The dehydration of C6 produced C5. In addition, C5 could be hydrogenated, carbonylated, and glucuronidated, yielding C8 ($C_9H_{10}O_2N$), C2 ($C_{10}H_8O_3N$), and C1 ($C_{15}H_{16}O_8N$). C6 was further converted to C3 by glucuronidation. Proposed pathways of the catabolic metabolites of pinostrobin are illustrated in **Figure 5**.

Excretion Study

The accumulative excretion ratio of pinostrobin is shown in **Figure 6**. The 54 h accumulative excretion ratios in urine and feces were 1.51% and 0.033%, respectively. The 24 h accumulative excretion ratio in bile was 0.024%. The excretion peak of pinostrobin in urine samples was observed 2–12 h after oral administration. After 24 h, a small amount of pinostrobin was detected in the urine. In feces, pinostrobin slowly reached its highest levels until 36 h and then gradually decreased. Similar to the urine excretion data, pinostrobin was rapidly excreted from the bile in the parent form from 3 to 9 h.

DISCUSSION

The pharmacokinetic results are consistent with those of previous studies (Hua et al., 2011; Sayre et al., 2013; Sayre et al., 2015). The significantly large V_z implied that pinostrobin preferentially binds to tissues and preferably remains within the body. Furthermore, pinostrobin showed double peaks in the concentration-time curve, similar to the pharmacokinetics studies of other flavonoid compounds that have been reported, which may be related to enterohepatic circulation or gastric emptying-regulated absorption (Xiong et al., 2015; Deng et al., 2017; Zeng et al., 2019). The accumulative excretion ratios in the urine remained below 1.51%, indicating that very little of the orally administered doses are excreted in the urine. A small amount of pinostrobin was detected in the urine after 24 h, which is consistent with previous reports using HPLC (Sayre et al., 2015). Pinostrobin slowly reached its highest levels in the feces until 36 h and was rapidly excreted from the bile from 3 to 9 h. Similar to urine excretion, the accumulative excretion ratios in the feces and bile are also very low. Therefore, it is likely that pinostrobin is mostly metabolized *in vivo* and plays a role in different organs.

Pinostrobin has been shown to have a significant protective effect against ethanol-induced gastric mucosal injury by scavenging free radicals produced by ethanol by activating cellular antioxidant defense. It can protect the gastric mucosa by reducing ulcer area and mucosal content and reducing or eliminating submucosal edema and leukocyte infiltration (Abdelwahab et al., 2011). Thus, the gastrointestinal tract (including the stomach, small intestine, and large intestine) is considered the primary target organ of pinostrobin in relieving gastrointestinal diseases. Pinostrobin accumulation was higher in the stomach, small intestine, and large intestine than in other tissues. Therefore, pinostrobin may effectively reduce peptic ulcers. In contrast, pinostrobin and its metabolites were poorly distributed in the tissues of the spleen,

TABLE 3 | The catabolic metabolites in rats after oral administration of 48.51 mg/kg pinostrobin.

Metabolites	RT (min)	$[M + H]^+$	Chemical formula	ppm	MS^n m/z	Source (T_{max}, R_{max})[a]
C1	0.70	338.08725	$C_{15}H_{16}O_8N$	−1.902	320.07602[M + H-H_2O]$^+$, 162.05435[M + H-C_6H8O_6]$^+$, 144.04382[M + H-$C_6H_8O_6$-H_2O]$^+$, 116.04903[M + H-C_6H8O_6-H_2O-CO]$^+$	Urine (36–54, 23.329); feces (4–6, 0.076)
C2	0.69	190.04930	$C_{10}H_8O_3N$	−2.997	172.03946[M + H-H_2O]$^+$, 162.05510[M + H-CO]$^+$, 144.04449[M + H-H_2O-CO]$^+$, 116.04951[M + H-H_2O-2CO]$^+$	Urine (36–54, 3.524); bile (20–24, 0.426); stomach (6, 0.024)
C3	0.70	340.10287	$C_{15}H_{18}O_8N$	0.521	322.09247[M + H-H_2O]$^+$, 304.08197[M + H-2H_2O]$^+$, 260.09192 [M + H-2H_2O-CO_2]$^+$, 206.08089[M + H-$C_4H_6O_5$]$^+$, 174.05484[M + H-$C_5H_{10}O_6$]$^+$, 164.07057174.05484[M + H-$C_6H_8O_6$]$^+$, 146.06000[M + H-$C_6H_8O_6$-H_2O]$^+$, 136.07559[M + H-$C_6H_8O_6$-CO]$^+$, 122.05999[M + H-$C_6H_8O_6$-C_2H2O]$^+$	Urine (24–36, 20.271); bile (20–24, 0.344)
C4	0.71	127.03858	$C_6H_7O_3$	−3.075	83.04864[M + H-CO_2]$^+$	Urine (24–36, 2.557); liver (6, 0.0.154)
C5	0.79	162.05447	$C_9H_8O_2N$	−2.993	144.04382[M + H-H_2O]$^+$, 134.05949[M + H-CO]$^+$, 105.03305[M + H-C_2H_3ON]$^+$	Urine (36–54, 101.501); feces (12–24, 8.460); bile (20–24, 0.277); stomach (6, 0.054); liver (12, 0.0.038)
C6	0.81	180.06503	$C_9H_{10}O_3N$	−2.720	162.05437[M + H-H_2O]$^+$, 134.05952[M + H-H_2O-CO]$^+$, 105.03307[M + H-H_2O-CO-CH_3N]$^+$	Urine (8–12, 109.725); bile (20–24, 2.565); liver (12, 0.076); stomach (3, 0.004)
C7	0.88	194.08064	$C_{10}H_{12}O_3N$	−2.730	176.07001[M + H-H_2O]$^+$, 148.07516[M + H-H_2O-CO]$^+$, 120.08035[M + H-H_2O-2CO]$^+$, 91.05382[M + H-H_2O-2CO-CH_3N]$^+$, 76.03896[M + H-C_8H_6O]$^+$	Urine (36–54, 222.178); bile (20–24, 6.207); liver (12, 0.305); stomach (6, 0.091); feces (0–2, 0.049)
C8	0.92	164.07057	$C_9H_{10}O_2N$	−0.214	146.05969[M + H-H_2O]$^+$, 136.07573[M + H-CO]$^+$, 122.05989[M + H-C_2H_2O]$^+$	Urine (24–36, 18.666); feces (2–4, 6.623); bile (16–20, 0.502); large intestine (1, 0.113); stomach (6, 0.102); liver (6, 0.049)
C9	1.10	167.07028	$C_9H_{11}O_3$	0.055	149.05870[M + H-H_2O]$^+$, 121.06412[M + H-CO]$^+$	Liver (0.083, 0.576); urine (0–2, 0.309); feces (4–6, 0.204)
C10	1.13	141.05414	$C_7H_9O_3$	−3.408	105.03313[M + H-2H_2O]$^+$, 97.06436[M + H-CO_2]$^+$	Urine (8–12, 0.025)
C11	1.24	206.08061	$C_{11}H_{12}O_3N$	−2.716	131.04919[M + H-$C_2H_5O_2N$]$^+$, 103.05421[M + H-$C_2H_5O_2N$-CO]$^+$	Urine (24–36, 9.382); bile (16–20, 3.133); feces (0–2, 0.116); stomach (0.33, 0.015)
C12	1.38	151.07503	$C_9H_{11}O_2$	−2.159	133.06436[M + H-H_2O]$^+$, 123.07993[M + H-CO]$^+$, 105.06962[M + H-H_2O-CO]$^+$	Urine (12–24, 0.416); feces (0–2, 0.058)
C13	1.63	193.04962	$C_{10}H_9O_4$	0.439	123.04360[M + H-$C_3H_2O_2$]$^+$	Urine (24–36, 0.028)
C14	2.73	149.05949	$C_9H_9O_2$	−1.450	131.04880[M + H-H_2O]$^+$, 103.05347[M + H-CO]$^+$	Urine (24–36, 0.089); feces (24–36, 0.044)
C15	2.78	123.04377	$C_7H_7O_2$	−2.324	105.03307[M + H-H_2O]$^+$, 95.04909[M + H-CO]$^+$	Urine (36–54, 0.889); feces (24–36, 0.368); large intestine (2, 0.160); stomach (6, 0.091); liver (6, 0.044)

T_{max} is the time or time interval to reach in biological samples. R_{max} is the maximum peak area ratio of catabolic metabolites relative to the internal standard.
[a]Sorted by R_{max} from large to small.

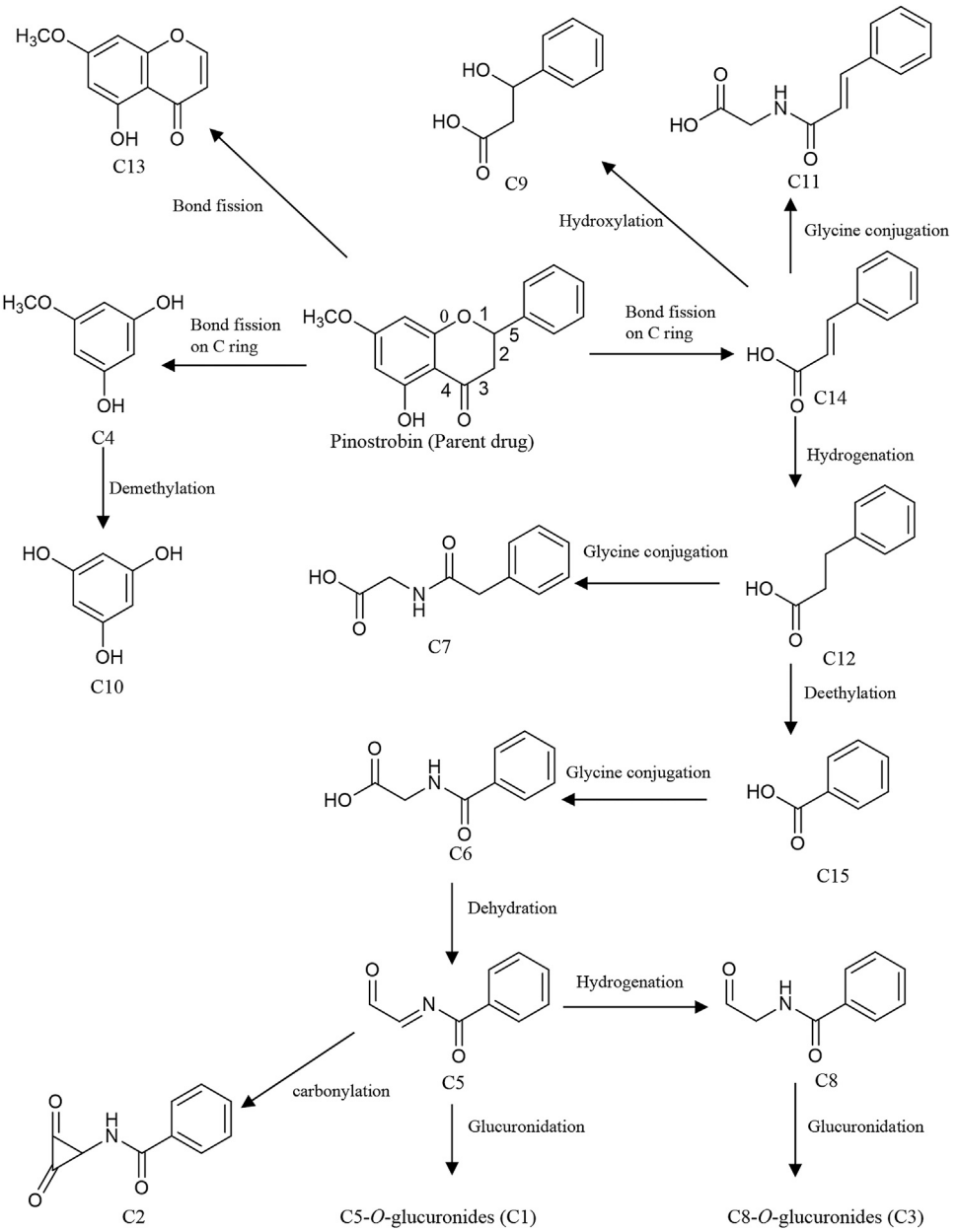

FIGURE 5 | Proposed pathways of the catabolic metabolites by intestinal bacteria in rats after oral administration of 48.51 mg/kg pinostrobin.

lung, heart, and kidney. Furthermore, very little pinostrobin was found in any examined tissues within 24 h, indicating no long-term accumulation in the tissues.

Owing to the high sensitivity and resolution of UPLC-LTQ orbitrap-MS/MS, more flavonoid metabolites and catabolic metabolites were detected in the biosamples. Pinostrobin underwent hydrogenation, hydroxylation, demethylation, decarbonylation, and acetylation to yield several other aglycones, including M5, M16, M11, M30, M23, and M7. Moreover, pinostrobin and the generated aglycones extensively undergo dehydrogenation, hydroxylation, and ketone conversion catalyzed by phase I metabolic enzymes and glucuronidation, sulfation, glycine conjugation, N-acetylcysteine conjugation, and taurine conjugation catalyzed by phase II metabolic enzymes in the gastrointestinal tract, liver, and other tissues. It is noteworthy that glucuronidation, decarbonylation, hydroxylation, demethylation, dehydrogenation, and glycine conjugation play essential roles in the metabolism of pinostrobin in rats.

CONCLUSION

In this study, a series of rapid, reliable, and sensitive UPLC-LTQ orbitrap-MS/MS methods were established, validated, and applied to

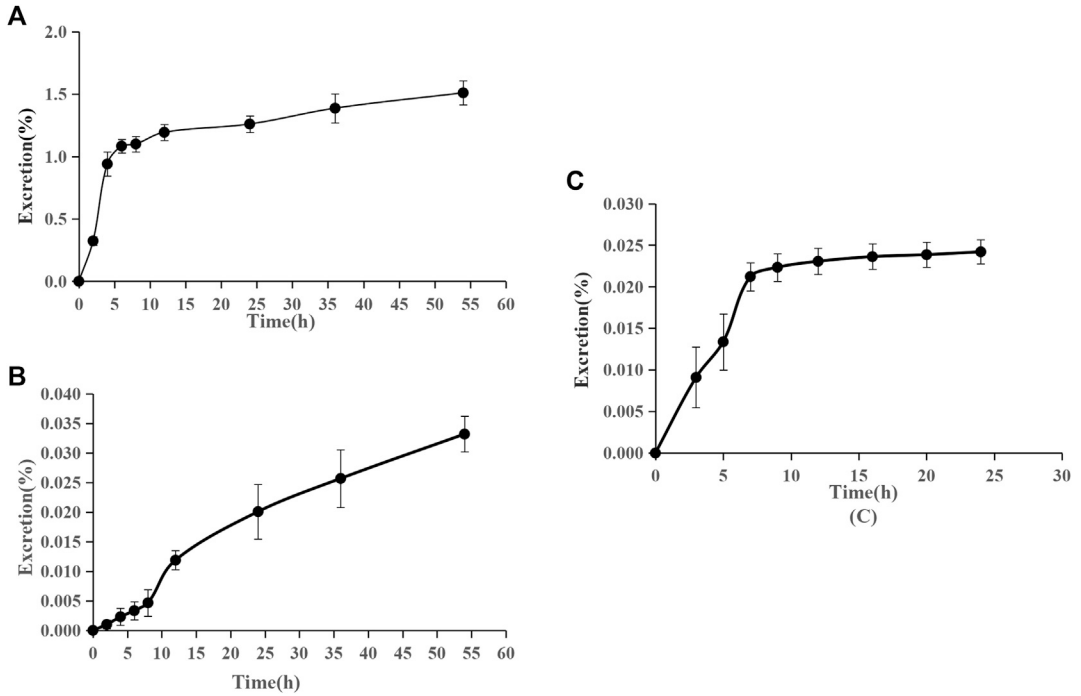

FIGURE 6 | Accumulative excretion ratio of pinostrobin in urine **(A)**, feces **(B)**, and bile **(C)** after oral administration of 48.51 mg/kg pinostrobin (n = 5, mean ± SD).

quantify pinostrobin in the plasma, urine, feces, bile, and various tissue samples collected from rats (Fu et al., 2018). After a single oral administration of pinostrobin, the pharmacokinetics showed a large apparent V_z, indicating that pinostrobin preferentially binds to tissues to exert a therapeutic effect. Pinostrobin was mostly distributed in the gastrointestinal tract, suggesting that it may be an effective component of traditional Chinese medicines to treat peptic ulcers. The excretion study showed that the amount of pinostrobin excreted through the urine, feces, and bile in the parent form was less than 1.567%, indicating that it is mainly metabolized *in vivo*. Using UPLC-LTQ orbitrap-MS/MS, 30 flavonoid metabolites were identified or partially identified in biosamples collected after dosing. In addition, we proposed the metabolism pathways of pinostrobin in rats. This study systemically investigated the pharmacokinetics, tissue distribution, metabolism, and excretion of pinostrobin in rats. These results would be helpful for the interpretation of the pharmacokinetics and pharmacodynamics of pinostrobin as well as traditional Chinese medicines containing pinostrobin. However, further investigations are necessary to better understand the pharmacokinetics of pinostrobin in rats with peptic ulcer and the mechanism of its intervention on peptic ulcer.

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.

ETHICS STATEMENT

The animal study was reviewed and approved by the Experimental Animal Ethics Committee of Henan University of Chinese Medicine.

AUTHOR CONTRIBUTIONS

XS and SC conceived and designed the study. XS and XL analyzed the research data and wrote the manuscript. XS and XL performed the experiments.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 81773859), the National Science and Technology Major Project of the Ministry of Science and Technology of China (Grant No. 2012ZX09103201-024), and the Doctoral Research Fund Project of Henan University of Chinese Medicine (Grant No. BSJJ2018-07).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2020.574638/full#supplementary-material.

REFERENCES

Abdelwahab, S. I., Mohan, S., Abdulla, M. A., Sukari, M. A., Abdul, A. B., Taha, M. M., et al. (2011). The methanolic extract of *Boesenbergia rotunda* (L.) Mansf. and its major compound pinostrobin induces anti-ulcerogenic property *in vivo*: possible involvement of indirect antioxidant action. *J. Ethnopharmacol.* 137, 963–970. doi:10.1016/j.jep.2011.07.010

Bhamarapravati, S., Juthapruth, S., Mahachai, W., and Mahady, G. (2006). Antibacterial activity of *Boesenbergia rotunda* (L.) Mansf. and *Myristica fragrans* Houtt. against *Helicobacter pylori*. *Songklanakarin J. Sci. Technol.* 28, 157–163.

Cappiello, A., Famiglini, G., Palma, P., Pierini, E., Termopoli, V., and Trufelli, H. (2008). Overcoming matrix effects in liquid chromatography-mass spectrometry. *Anal. Chem.* 80, 9343–9348. doi:10.1021/ac8018312

Charoensin, S., Punvittayagul, C., Pompimon, W., Mevatee, U., and Wongpoomchai, R. (2010). Toxicological and clastogenic evaluation of pinocembrin and pinostrobin isolated from *Boesenbergia pandurata* in Wistar rats. *Thai. J. Toxicol.* 25, 29–40.

Chen, S., Wang, L., Zhang, W., and Wei, Y. (2015). Secondary metabolites from the root of *Lindera reflexa* Hemsl. *Fitoterapia* 105, 222–227. doi:10.1016/j.fitote.2015.07.005

Déciga-Campos, M., Mata, R., and Rivero-Cruz, I. (2017). Antinociceptive pharmacological profile of *Dysphania graveolens* in mouse. *Biomed. Pharmacother.* 89, 933–938. doi:10.1016/j.biopha.2017.02.096.

Deng, C., Gao, C., Tian, X., Chao, B., Wang, F., Zhang, Y., et al. (2017). Pharmacokinetics, tissue distribution and excretion of luteolin and its major metabolites in rats: metabolites predominate in blood, tissues and are mainly excreted via bile. *J. Funct. Foods* 35, 332–340. doi:10.1016/j.jff.2017.05.056

Dzoyem, J. P., Nkuete, A. H. L., Ngameni, B., and Eloff, J. N. (2017). Anti-inflammatory and anticholinesterase activity of six flavonoids isolated from Polygonum and Dorstenia species. *Arch. Pharm. Res. (Seoul)* 40, 1129–1134. doi:10.1007/s12272-015-0612-9

Erdtman, H. (1944). *Sven. Kem. Tidskr.* 56, 2.

Fu, Y., Sun, X., Wang, L., and Chen, S. (2018). Pharmacokinetics and tissue distribution study of pinosylvin in rats by ultra-high-performance liquid chromatography coupled with linear trap quadrupole orbitrap mass spectrometry. *Evid. Based Complement. Alternat. Med.*, 2018, 1–14. doi:10.1155/2018/4181084

Gaffield, W. (1970). Circular dichroism, optical rotatory dispersion and absolute configuration of flavanones, 3-hydroxyflavanones and their glycosides. *Tetrahedron* 26, 4093–4108. doi:10.1016/S0040-4020(01)93050-9

Gómez-Betancur, I., Cortés, N., Benjumea, D., Osorio, E., León, F., and Cutler, S. J. (2015). Antinociceptive activity of extracts and secondary metabolites from wild growing and micropropagated plants of *Renealmia alpinia*. *J. Ethnopharmacol.* 165, 191–197. doi:10.1016/j.jep.2015.02.012

Gu, C., Fu, L., Yuan, X., and Liu, Z. (2017). Promoting effect of pinostrobin on the proliferation, differentiation, and mineralization of murine pre-osteoblastic MC3T3-E1 cells. *Molecules* 22, 1735. doi:10.3390/molecules22101735

Hernández Tasco, A. J., Ramírez Rueda, R. Y., Alvarez, C. J., Sartori, F. T., Sacilotto, A. C. B. C., Ito, I. Y., et al. (2020). Antibacterial and antifungal properties of crude extracts and isolated compounds from *Lychnophora markgravii*. *Nat. Prod. Res.* 34, 863–867. doi:10.1080/14786419.2018.1503263

Hua, X., Fu, Y., Zu, Y., Zhang, L., Wang, W., and Luo, M. (2011). Determination of pinostrobin in rat plasma by LC-MS/MS: application to pharmacokinetics. *J. Pharmaceut. Biomed. Anal.* 56, 841–845. doi:10.1016/j.jpba.2011.07.038

Jadaun, A., Subbarao, N., and Dixit, A. (2017). Allosteric inhibition of topoisomerase I by pinostrobin: molecular docking, spectroscopic and topoisomerase I activity studies. *J. Photochem. Photobiol. B* 167, 299–308. doi:10.1016/j.jphotobiol.2017.01.010.

Jaudan, A., Sharma, S., Malek, S., and Dixit, A. (2018). Induction of apoptosis by pinostrobin in human cervical cancer cells: possible mechanism of action. *PLoS One* 13, 1–23. doi:10.1371/journal.pone.0191523

Jones, A. A., and Gehler, S. (2020). Acacetin and pinostrobin inhibit malignant breast epithelial cell adhesion and focal adhesion formation to attenuate cell migration. *Integr. Cancer Ther.* 19, 1–12. doi:10.1177/1534735420918945

Kanchanapiboon, J., Kongsa, U., Pattamadilok, D., Kamponchaidet, S., Wachisunthon, D., Poonsatha, S., et al. (2020). *Boesenbergia rotunda* extract inhibits *Candida albicans* biofilm formation by pinostrobin and pinocembrin. *J. Ethnopharmacol.* 261, 113193. doi:10.1016/j.jep.2020.113193

Kuna, L., Jakab, J., Smolic, R., Raguz-Lucic, N., Vcev, A., and Smolic, M. (2019). Peptic ulcer disease: a brief review of conventional therapy and herbal treatment options. *J. Clin. Med.* 8, 179. doi:10.3390/jcm8020179

Lanas, A., and Chan, F. K. L. (2017). Peptic ulcer disease. *Lancet* 390, 613–624. doi:10.1016/S0140-6736(16)32404-7

Liu, M., Zou, W., Yang, C., Peng, W., and Su, W. (2012). Metabolism and excretion studies of oral administered naringin, a putative antitussive, in rats and dogs. *Biopharm. Drug Dispos.* 33, 123–134. doi:10.1002/bdd.1775

Meckes, M., Paz, D., Acosta, J., and Mata, R. (1998). The effects of chrysin and pinostrobin, two flavonoids isolated from *Teloxys graveolens* leaves, on isolated guinea-pig ileum. *Phytomedicine* 5, 459–463. doi:10.1016/S0944-7113(98)80042-1

Patel, N. K., and Bhutani, K. K. (2014). Pinostrobin and Cajanus lactone isolated from *Cajanus cajan* (L.) leaves inhibits TNF-α and IL-1β production: *in vitro* and *in vivo* experimentation. *Phytomedicine* 21, 946–953. doi:10.1016/j.phymed.2014.02.011

Patel, N. K., Jaiswal, G., and Bhutani, K. K. (2016). A review on biological sources, chemistry and pharmacological activities of pinostrobin. *Nat. Prod. Res.* 30, 2017–2027. doi:10.1080/14786419.2015.11

Sayre, C. L., Alrushaid, S., Martinez, S. E., Anderson, H. D., and Davies, N. M. (2015). Pre-clinical pharmacokinetic and pharmacodynamic characterization of selected chiral flavonoids: pinocembrin and pinostrobin. *J. Pharm. Pharmaceut. Sci.* 18, 368–396. doi:10.18433/j3bk5t

Sayre, C. L., Zhang, Y., Martinez, S. E., Takemoto, J. K., and Davies, N. M. (2013). Stereospecific analytical method development and preliminary *in vivo* pharmacokinetic characterization of pinostrobin in the rat. *Biomed. Chromatogr.* 27, 548–550. doi:10.1002/bmc.2834

Siekmann, T. R., Burgazli, K. M., Bobrich, M. A., Nöll, G., and Erdogan, A. (2013). The antiproliferative effect of pinostrobin on human umbilical vein endothelial cells (HUVEC). *Eur. Rev. Med. Pharmacol. Sci.* 17, 668–672.

Smolarz, H. D., Mendyk, E., Bogucka-Kocka, A., and Kockic, J. (2005). Pinostrobin—an anti-leukemic flavonoid from *Polygonum lapathifolium* L. ssp. nodosum (Pers.) Dans. *Z. Naturforsch.* 61, 64–68. doi:10.1515/znc-2006-1-212

Sopanaporn, J., Suksawatamnuay, S., Sardikin, A., Lengwittaya, R., Chavasiri, W., Miyakawa, T., et al. (2020). Pinostrobin suppresses the Ca^{2+}-signal-dependent growth arrest in yeast by inhibiting the Swe1-mediated G2 cell-cycle regulation. *FEMS Yeast Res.* 20, 1–6. doi:10.1093/femsyr/foaa026

Su, B. N., Parka, E. J., Vigo, J. S., Graham, J. G., Cabieses, F., Fong, H. H. S., et al. (2003). Activity-guided isolation of the chemical constituents of *Muntingia calabura* using a quinone reductase induction assay. *Phytochemistry* 63, 335–341. doi:10.1016/s0031-9422(03)00112-2

Sun, X., Zhang, Y., Chen, S., and Fu, Y. (2016). Characterization and identification of the chemical constituents in the root of *Lindera reflexa* Hemsl. using ultra-high performance liquid chromatography coupled with linear trap quadrupole orbitrap mass spectrometry. *J. Pharmaceut. Biomed. Anal.* 126, 34–47. doi:10.1016/j.jpba.2016.04.023

Vasas, A., Lajter, I., Kúsz, N., Forgó, P., Jakab, G., Fazakas, C., et al. (2020). Flavonoid, stilbene and diarylheptanoid constituents of *Persicaria maculosa* Gray and cytotoxic activity of the isolated compounds. *Fitoterapia* 145, 104610. doi:10.1016/j.fitote.2020.104610

Vechi, G., Tenfen, A., Capusiri, E. S., Gimenez, A., and Cechinel-Filho, V. (2020). Antiparasitic activity of two Brazilian plants: *Eugenia mattosii* and *Marlierea eugeniopsoides*. *Nat. Prod. Res.*, 1–5. doi:10.1080/14786419.2020.1739676

Wang, L. L., Zhang, Y. B., Sun, X. Y., and Chen, S. Q. (2016). Simultaneous quantitative analysis of main components in linderae reflexae radix with one single marker. *J. Liq. Chromatogr. Relat. Technol.* 39, 422–427. doi:10.1080/10826076.2016.1169429

Wu, N., Kong, Y., Zu, Y., Fu, Y., Liu, Z., Meng, R., et al. (2011). Activity investigation of pinostrobin towards herpes simplex virus-1 as determined by atomic force microscopy. *Phytomedicine* 18, 110–118. doi:10.1016/j.phymed.2010.07.001

Xiong, F., Wang, H., Jiang, Z., Huo, M., Yan, C., Zheng, C., et al. (2015). Integrated pharmacokinetics and biodistribution of multiple flavonoid C-glycosides components in rat after oral administration of *Abrus mollis* extract and correlations with bio-effects. *J. Ethnopharmacol.* 163, 290–296. doi:10.1016/j.jep.2014.12.067

Zeng, X., Su, W., Zheng, Y., He, Y., He, Y., Rao, H., et al. (2019). Pharmacokinetics, tissue distribution, metabolism, and excretion of naringin in aged rats. *Front. Pharmacol.* 10, 34. doi:10.3389/fphar.2019.00034

Zeng, X., Su, W., Zheng, Y., Liu, H., Li, P., Zhang, W., et al. (2018). UFLC-Q-TOF-MS/MS-based screening and identification of flavonoids and derived metabolites in human urine after oral administration of Exocarpium Citri Grandis extract. *Molecules* 23, 895. doi:10.3390/molecules23040895

Zhang, Y. X., Yang, T. T., Xia, L., Zhang, W. F., Wang, J. F., and Wu, Y. P. (2017). Inhibitory effect of propolis on platelet aggregation in vitro. *J. Healthc. Eng.* 2017, 1–6. doi:10.1155/2017/3050895

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 02 December 2020
doi: 10.3389/fpls.2020.608109

mdm-miR828 Participates in the Feedback Loop to Regulate Anthocyanin Accumulation in Apple Peel

Bo Zhang[1,2†], Hui-Juan Yang[1,2†], Ya-Zhou Yang[1,2], Zhen-Zhen Zhu[1,2], Ya-Nan Li[1,2], Dong Qu[3] and Zheng-Yang Zhao[1,2]*

[1]State Key Laboratory of Crop Stress Biology for Arid Areas, College of Horticulture, Northwest A&F University, Yangling, China, [2]Apple Engineering and Technology Research Center of Shaanxi Province, Northwest A&F University, Yangling, China, [3]Shaanxi Key Laboratory Bio-resources, College of Bioscience and Engineering, Shaanxi University of Technology, Hanzhong, China

OPEN ACCESS

Edited by:
Cristina Garcia-Viguera, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
Nick Albert, The New Zealand Institute for Plant and Food Research Ltd., New Zealand
Antje Feller, University of Tübingen, Germany

***Correspondence:**
Zheng-Yang Zhao
zhaozy@nwsuaf.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: 19 September 2020
Accepted: 10 November 2020
Published: 02 December 2020

Citation:
Zhang B, Yang H-J, Yang Y-Z, Zhu Z-Z, Li Y-N, Qu D and Zhao Z-Y (2020) mdm-miR828 Participates in the Feedback Loop to Regulate Anthocyanin Accumulation in Apple Peel. Front. Plant Sci. 11:608109. doi: 10.3389/fpls.2020.608109

Anthocyanins are responsible for the red pigmentation in the peel of apple (*Malus* × *domestica* Borkh.) fruit. Relatively few studies have investigated anthocyanins at the posttranscriptional level. MicroRNAs play an important role in plant growth and development by regulating gene expression at the posttranscriptional level. In this study, mdm-miR828 showed a relatively low expression level during the rapid fruit coloration period. However, the mdm-miR828 expression level increased in the late fruit coloration stage. Overexpression of mdm-miR828 inhibited anthocyanin synthesis in apple and *Arabidopsis*. Dual-luciferase and yeast one-hybrid assays showed that MdMYB1 is capable of binding to the promoter of mdm-MIR828b to promote its expression. The results indicate that mdm-miR828 is involved in a feedback regulatory mechanism associated with anthocyanin accumulation in apple. In addition, mdm-miR828 is involved in the inhibition of anthocyanin accumulation in response to high temperature.

Keywords: apple, anthocyanin, mdm-miR828, *MdMYB1*, high temperature

INTRODUCTION

Apple (*Malus* × *domestica* Borkh.) is an economically important fruit crop that is cultivated worldwide. The peel color is an important phenotypic trait of apple fruit, which largely determines the commodity value of the fruit. Anthocyanins, a diverse group of secondary metabolites in plants is responsible for the red coloration in apple peel (Takos et al., 2006). In addition, anthocyanins may be nutritionally and medically beneficial to human health (Chaves-Silva et al., 2018). The mechanism of anthocyanin synthesis has been elucidated in many plant species, for example, petunia (*Petunia hybrida*; de Vetten et al., 1997; Quattrocchio et al., 1999; Spelt et al., 2000; Albert et al., 2011, 2014), *Arabidopsis thaliana* (Borevitz et al., 2000; Gonzalez et al., 2008; Petroni and Tonelli, 2011), grape (*Vitis vinifera*; Kobayashi et al., 2004; Hichri et al., 2010; Matus et al., 2010), nectarine (*Prunus persica*; Ravaglia et al., 2013), and strawberry (*Fragaria* × *ananassa*; Aharoni et al., 2001; Lin-Wang et al., 2010; Medina-Puche et al., 2014). Anthocyanins are synthesized *via* the flavonoid metabolism pathway (Kim et al., 2003). Many enzymes participate in this pathway, such as chalcone

synthase (CHS), flavanone 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), and flavonoid 3-*O*-glycosyltransferase (UFGT; Holton and Cornish, 1995; Shirley et al., 1995; Kim et al., 2003). These enzymes are coregulated by the MYB-bHLH-WD40 (MBW) transcription factor complex (Goff et al., 1992; Baudry et al., 2004; Lloyd et al., 2017). Three alleles *MdMYB10*, *MdMYB1*, and *MdMYBA* that regulate anthocyanin synthesis in apple have been identified. These transcription factors directly bind to the promoter of anthocyanin biosynthesis genes to promote expression of the latter (Takos et al., 2006; Ban et al., 2007; Espley et al., 2007). For example, a cold-induced bHLH transcription factor, MdbHLH3, directly binds to the promoter of the structural gene *MdUFGT* and that of *MdDFR* to activate their expression and also binds to the promoter of the regulatory gene *MdMYB1* to promote its expression (Xie et al., 2012).

MicroRNA (miRNA) is a type of endogenous noncoding single-stranded small RNA in eukaryotes that predominantly guides posttranscriptional gene-silencing activities (Cuperus et al., 2011). The miRNA genes are first transcribed into primary transcripts in the nucleus and then processed into miRNA precursors (pre-miRNAs/MIRNA; Song et al., 2010). Subsequently, these pre-miRNAs are processed to form mature miRNAs (Voinnet, 2009). In addition to miRNAs, small interfering RNAs (siRNAs) are synthesized in plants (Rajagopalan et al., 2006), of which one is *trans*-acting small interfering RNA (ta-siRNA). The miRNA cleaves *Trans-Acting SiRNA* (*TAS*) loci transcripts, and then with the aid of RNA-dependent RNA polymerase 6 (RDR6) and suppressor of gene silencing 3 (SGS3), double-stranded RNA is synthesized and is cleaved by the DCL4 enzyme to produce ta-siRNAs. Its mechanism of action is similar to that of miRNA, that is, cleavage of the target transcript to interfere with gene expression (Allen et al., 2005).

MiRNAs are involved in the regulation of anthocyanin synthesis in plants (Liu et al., 2016; Sun et al., 2017). In *Arabidopsis*, SPL9, the target gene of miR156, competes with a bHLH transcription factor to bind to the MYB75 (PAP1) transcription factor, thereby disrupting the structure of the MBW protein complex and, ultimately, negatively regulating anthocyanin synthesis (Gou et al., 2011). *Arabidopsis* overexpressing miR778 accumulates higher amounts of anthocyanins than the wild type (WT) under phosphorus deficiency (Wang et al., 2015). The miR858 induced by elongated hypocotyl5 (HY5) targets the anthocyanin inhibitor MYBL2 to promote anthocyanin accumulation in *Arabidopsis* (Wang et al., 2016b). The miR828 was the first miRNA identified to negatively regulate anthocyanin accumulation. In *Arabidopsis*, miR828 targets the positive regulator of anthocyanin synthesis MYB113 and cleaves *TAS4* to produce TAS4-siR81(−), whereas siR81(−) targets the anthocyanin-related transcription factors PAP1, PAP2, and MYB113 (Rajagopalan et al., 2006). Xia et al. (2012) conducted sequencing of sRNAs from the leaf, root, flower, and fruit of "Golden Delicious" apple and observed that mdm-miR828 cleaves *MdTAS4* to derive mdm-siR81(−), and it was predicted that mdm-siR81(−) might target *MdbHLH3*. RNA gel blot analysis showed that mdm-miR828 predominantly accumulates in flowers. However, "Golden Delicious" is an apple cultivar with yellow fruit peel. It is unclear whether mdm-miR828 and mdm-siR81(−) are involved in the coloration of fruit with red peel, and the relationship between mdm-miR828 and MdMYB1 is uncertain.

In the present study, we investigated the expression pattern of mdm-miR828 during fruit coloration in the red-peeled apple "Starkrimson Delicious." Overexpression of mdm-miR828 inhibited anthocyanin accumulation in the fruit peel, and heterologous overexpression of mdm-miR828 inhibited anthocyanin accumulation in *Arabidopsis*. MdMYB1 induced expression of mdm-miR828, which indicated that a feedback regulatory mechanism controlled the anthocyanin content in the peel. In addition, mdm-miR828 may be involved in the inhibition of anthocyanin accumulation in response to high temperature.

MATERIALS AND METHODS

Plant Materials

Plants of apple "Starkrimson Delicious" were grown at the Apple Experimental Farm of Northwest A&F University, Shaanxi Province, China. Fruit were sampled at 110, 116, 122, 128, 134, 140, and 146 days after blooming. Nine fruit were selected at each time point with three fruit considered as one biological replicate. The peel was excised with a fruit peeler, immediately frozen with liquid nitrogen, and stored at −80°C until use. Transient transformation and high-temperature treatment were applied to fruit that were wrapped in a paper bag (Hongtai, Shanxi, China) at 45 days after blooming and harvested at 120 days after blooming.

Determination of Anthocyanin Content

The total anthocyanin content of the fruit was determined using the method described by Xie et al. (2012). A sample (0.5 g) of peel was extracted in 5 ml of 1% (vol/vol) HCl-methanol and incubated in the dark at 4°C for 24 h. After centrifugation at 13,000 × *g* for 5 min, the absorbance of the upper liquid was measured at 530, 620, and 650 nm. The anthocyanin content was calculated using the formula $OD = (A_{530} - A_{620}) - 0.1(A_{650} - A_{620})$. One unit of anthocyanin content was expressed as a change of 0.1 OD (unit × 10^3 g^{-1} fresh weight). Measurement of cyanidin 3-galactoside content was performed as described previously (Liu et al., 2013). A high-performance liquid chromatograph (Waters, Milford, MA, United States) was used for all analyses. Cyanidin 3-galactoside standard (Sigma Chemical, St Louis, MO, United States) was used. Separation of cyanidin 3-galactoside was accomplished on a C18 column (5-μm internal diameter, 250 × 4.6 mm; Waters). The anthocyanin content of *Arabidopsis* was determined using the method of Wang et al. (2016b).

RNA Extraction and Quantitative Real-Time Polymerase Chain Reaction Analysis

Total RNA was extracted using TRIzol RNA Plant Plus Reagent (Tiangen, Beijing, China) and then treated with DNase I (TaKaRa, Dalian, China) to remove genomic DNA

contamination. The RNA was reverse-transcribed with the PrimeScript™ RT Reagent Kit (TaKaRa, Dalian, China). The reverse transcription of mdm-miR828 and mdm-siR81(−) used specific stem-loop primers (**Supplementary Table S1**; Chen et al., 2005), and that of the internal reference gene *MdU6* used gene-specific primers (**Supplementary Table S1**). Quantitative real-time polymerase chain reaction (RT-qPCR) analysis was conducted using SYBR® Premix Ex Taq™ II (TaKaRa, Dalian, China) on an ABI StepOnePlus™ Real-Time PCR System (Applied Biosystems, Waltham, MA, United States). Poly (A) polymerase (PAP) RT-qPCR was implemented using the miRcute Plus miRNA First-Strand cDNA Kit (Tiangen, Beijing, China) and miRcute Plus miRNA qPCR Kit (Tiangen, Beijing, China). For normalization, 5S ribosomal RNA was used. *MdActin* was used to normalize the coding genes. The $2^{-\Delta\Delta CT}$ method was used to calculate the relative expression level. Quantitative primer sequences used are listed in **Supplementary Table S1**.

RNA Ligase-Mediated 5'-RACE

To verify the cleavage of *MdTAS4* by mdm-miR828 and that of *MdbHLH3* by mdm-siR81(−), RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) experiments was performed using the SMARTer® RACE 5'/3' Kit (TaKaRa, Dalian, China). First, total RNA was reverse-transcribed using SMARTScribe™ reverse transcriptase, and the 5' adaptor was attached. Next, 5' universal primers, 3' gene-specific primers, and the cDNA were used for PCR amplification. Gel extraction, in-fusion cloning, and sequencing were then performed.

Vector Construction and Genetic Transformation

The coding sequence (CDS) of *MdbHLH3* was cloned from the cDNA derived from "Starkrimson Delicious" fruit peel. The sequence of Md-pri-miR828b (comprising the Md-MIR828b sequence and approximately 100-bp upstream and downstream) was amplified from DNA extracted from "Starkrimson Delicious" peel. The sequences were inserted into the pCAMBIA2301 vector to form *35S:MdbHLH3* and *35S:mdm-miR828*. The primer sequences used are listed in **Supplementary Table S1**. The constructed vector was transformed into *Agrobacterium tumefaciens* strain GV3101.

Agrobacterium-mediated transformation followed the methods of Yang et al. (2019) and Tian et al. (2015). *Agrobacterium* cells harboring the overexpression vector were grown, collected, and resuspended in *Agrobacterium* infiltration buffer (200 mM acetosyringone, 10 mM MES, and 10 mM $MgCl_2$) to OD = 1.2. The suspension was incubated at room temperature for 4 h. The sampled bagged covered apples were infiltrated with *Agrobacterium* carrying *35S:MdbHLH3*, *35S:mdm-miR828*, or the pCAMBIA2301 empty vector and then placed at room temperature in the dark for 24 h and transferred to an incubator maintained at 22°C illuminated with white light (200 μmol m^{-2} s^{-1}) for 3 days.

The *35S:mdm-miR828* vector was used to transform *A. thaliana* using the floral-dip method (Clough and Bent, 1998). T3 generation of 35S:miR828 transformed *Arabidopsis* was used for the experiment.

Transient Dual-Luciferase Assay

A transient dual-luciferase assay was performed using tobacco (*Nicotiana benthamiana*) leaves. The mdm-MIR828b promoter was amplified and cloned into the pGreenII 0800-LUC vector. The CDS of *MdMYB1* was amplified and cloned into the pGreenII 62-SK vector. The constructs were transformed into *Agrobacterium* strain GV3101 carrying the pSoup vector. *Agrobacterium* cells harboring the different vectors were cultured, collected, and resuspended in infiltration buffer (200 mM acetosyringone, 10 mM MES, and 10 mM $MgCl_2$). After incubation for 2 h, tobacco leaves were injected with the cell suspension. After initial incubation for 24 h in the dark, the leaves were incubated under light (200 μmol m^{-2} s^{-1}) with a 16 h/8 h (light/dark) photoperiod at 25/22°C (light/dark) for 3 days. Leaf discs were punched from tobacco leaves of similar size, and a microplate reader (Infinite M200 PRO, Tecan, Switzerland) was used to determine the firefly luciferase and *Renilla* luciferase activities using the Dual-Glo® Luciferase Assay System (Promega, Madison, WI, United States) in accordance with the manufacturer's instructions.

Yeast One-Hybrid Assay

The CDS of MdMYB1 was inserted into the pGADT7 vector to generate the recombinant vector MdMYB1-AD, and the mdm-MIR828b promoter fragment was inserted into the pHIS2.1 vector to generate the recombinant vector *proMIR828b-HIS*. Yeast (*Saccharomyces cerevisiae*) strain Y187 cells carrying the *proMIR828b-HIS* vector were cultured in −T−H screening medium supplemented with different concentrations of 3-amino-1,2,4-triazole (3-AT). The MdMYB1-AD and *proMIR828b-HIS* vectors were cotransformed into yeast Y187 cells, which were then cultured in −T−H−L selective medium containing an optimal concentration of 3-AT.

High-Temperature Treatment

Bagged covered fruit of "Starkrimson Delicious" apple were harvested at 120 days after blooming. The harvested fruit were placed in an incubator at 35°C under white light (200 μmol m^{-2} s^{-1}). The control was placed in an incubator at 23°C and under white light (200 μmol m^{-2} s^{-1}). The fruit were sampled after 12, 24, and 48 h for analysis of anthocyanin content and gene expression. Three biological replicates, consisting of three fruit per replicate, were sampled at each time point.

RESULTS

Anthocyanin Content and mdm-miR828 Expression During Apple Coloration

We determined the total anthocyanin content in the peel of "Starkrimson Delicious" fruit without covering at 110, 116, 122, 128, 134, and 146 days after blooming. No change in anthocyanin content was observed from 110 to 116 days,

whereas anthocyanins accumulated rapidly in the peel from 116 to 134 days (the rapid coloration period). The anthocyanin content stabilized from 134 to 146 days (the late fruit coloration stage; **Figures 1A,B**). The content of cyanidin 3-galactoside, which accounted for more than 94% of the total anthocyanin content, showed a similar trend (**Supplementary Figure S1**).

We used the stem-loop method to detect the expression of mdm-miR828 in the peel. Expression of mdm-miR828 decreased rapidly from 110 to 116 days and showed a relatively low expression level during the period of rapid anthocyanin synthesis (116–134 days). However, mdm-miR828 expression increased from 134 to 146 days (**Figure 1C**). The PAP RT-qPCR method yielded similar results (**Supplementary Figure S2**). In addition, we measured the expression levels of *MdTAS4*, mdm-siR81(−), and *MdbHLH3* (**Figure 1C**). The expression level of *MdTAS4* increased from 116 to 134 days and thereafter decreased. That of mdm-siR81(−) declined rapidly from 116 to 128 days, remained relatively low from 128 to 134 days, and increased after 134 days. The expression level of *MdbHLH3* increased sharply from 122 to 128 days and thereafter decreased from 134 to 146 days. The expression level of anthocyanin structural genes (*MdCHS*, *MdDFR*, *MdANS*, and *MdUFGT*) increased from 116 to 134 days and then decreased to varying degrees (**Supplementary Figure S2**).

Overexpression of mdm-miR828 Inhibits Anthocyanin Accumulation in Apples

To confirm that mdm-miR828 is involved in anthocyanin biosynthesis in the peel, we generated the construct *35S:mdm-miR828* and injected the fruit peel of "Starkrimson Delicious" with *Agrobacterium* carrying this construct. Transient overexpression of mdm-miR828 inhibited anthocyanin accumulation around the injection site, whereas no reduction in anthocyanin content was observed after injection of *Agrobacterium* containing the empty vector. Thus, mdm-miR828 was indicated to negatively regulate anthocyanin accumulation in apple. In addition, overexpression of *MdbHLH3* significantly promoted accumulation of anthocyanins in the peel (**Figures 2A,B**).

The expression of anthocyanin synthesis-related genes was further analyzed (**Figure 2C**). The expression levels of *MdbHLH3*, *MdbMYB1*, *MdDFR*, *MdANS*, and *MdUFGT* increased to varying degrees compared with those of the control (empty vector) in the peel of fruit overexpressing *MdbHLH3*. In addition, mdm-miR828, *MdTAS4*, and mdm-siR81(−) expression levels were higher than those of the control. In the peel of fruit overexpressing mdm-miR828, both mdm-miR828 and mdm-siR81(−) were up-regulated compared with those of the control (empty vector), whereas the expression levels of *MdTAS4*,

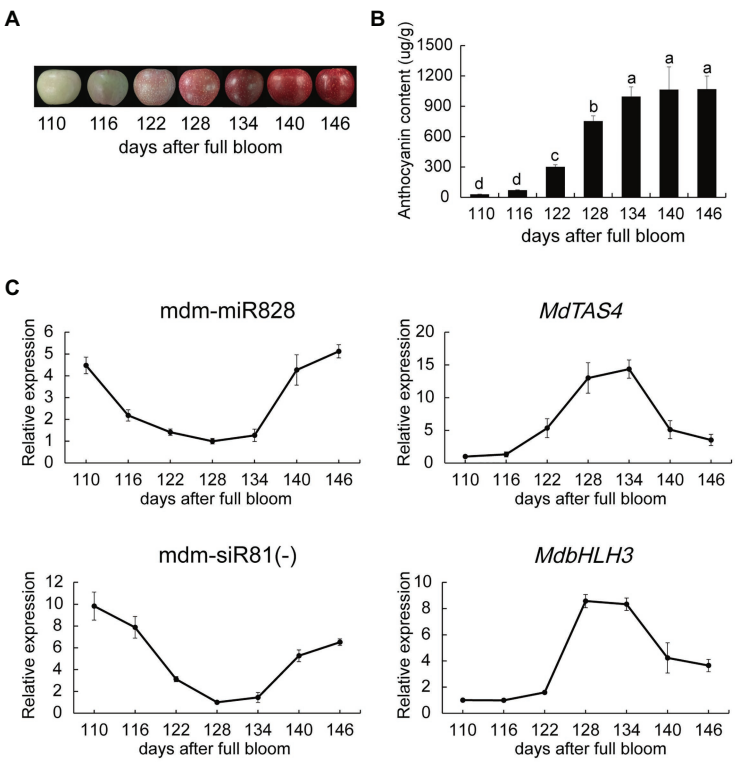

FIGURE 1 | Anthocyanin content and mdm-miR828 expression during coloration of apple "Starkrimson Delicious" fruit. **(A)** Change in peel color during coloration. **(B)** Peel anthocyanin content during coloration. **(C)** Relative expression levels of mdm-miR828, mdm-siR81(–), and anthocyanin biosynthesis-related genes. The expression levels of mdm-miR828 and mdm-siR81(–) were detected by the stem-loop method. *MdU6* was used as an internal reference. Error bars represent the standard deviation of three biological replicates. Different letters above the bars indicate a significant difference ($p < 0.05$; one-way ANOVA and LSD test).

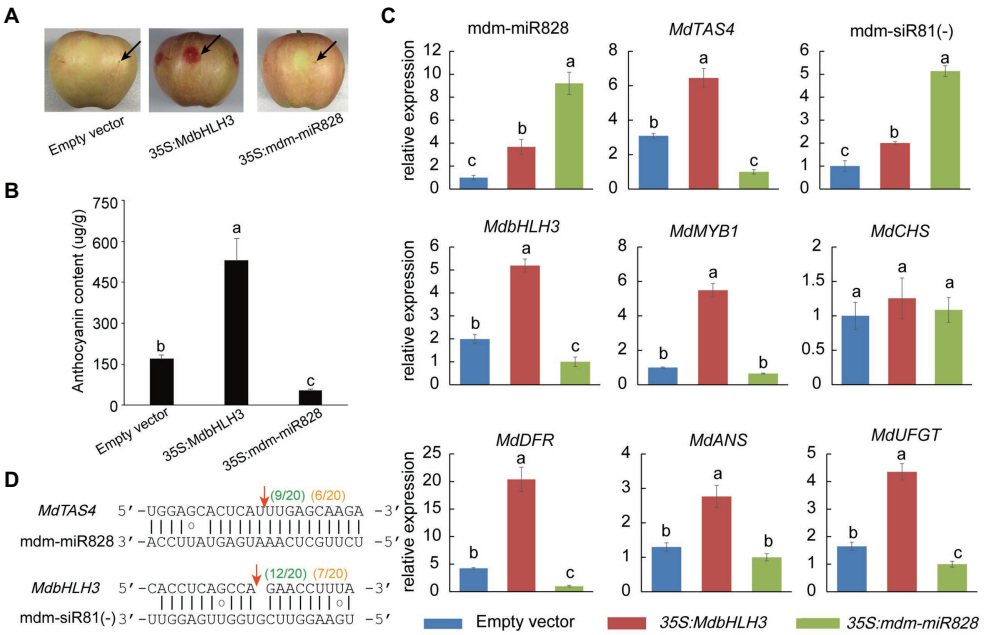

FIGURE 2 | Transient expression of mdm-miR828 in apple "Starkrimson Delicious" fruit. **(A)** Overexpression of mdm-miR828 inhibits anthocyanin accumulation in the peel. The black arrows represent the injection sites. **(B)** Anthocyanin content in the peel of fruit overexpressing mdm-miR828 or *MdbHLH3*. **(C)** Overexpression of mdm-miR828 and *MdbHLH3* affects the expression level of anthocyanin biosynthesis-related genes in the peel. **(D)** The binding positions of mdm-miR828 and mdm-siR81(−) in the target genes *MdTAS4* and *MdbHLH3*. Green numbers and orange numbers in parentheses indicate the proportion of 5'-RACE clones at the sites where the target gene is cleaved by mdm-miR828 and mdm-siR81(−) in mdm-miR828-overexpressing peel and mixed peels (harvested at 134, 140, and 146 days after blooming), respectively. Error bars represent the standard deviation of three biological replicates. Different letters above the bars indicate a significant difference ($p < 0.05$, one-way ANOVA and LSD test).

MdbHLH3, *MdDFR*, *MdANS*, and *MdUFGT* decreased compared with those of the control.

The RLM-RACE procedure was applied using the peel of fruit overexpressing mdm-miR828 to determine the cleavage site of mdm-miR828 on the *MdTAS4* transcript and that of mdm-siR81(−) on the *MdbHLH3* transcript. Some predicted 5' free ends of *MdTAS4* and *MdbHLH3* (between the 10th and 11th nucleotides of the miRNA binding site to the target gene) were detected (**Figure 2D**). In addition, we mixed samples of peels (collected at 134, 140, and 146 days) during the coloration period and performed the RACE experiment. The cleavage sites of mdm-miR828 and mdm-siR81(−) on the target genes *MdTAS4* and *MdbHLH3* were also detected.

Overexpression of mdm-miR828 in *Arabidopsis* Inhibits Anthocyanin Accumulation

To confirm the function of mdm-miR828 in anthocyanin biosynthesis, we generated mdm-miR828-overexpressing *Arabidopsis* transformants. Sucrose can induce the accumulation of anthocyanins in *Arabidopsis* seedlings (Teng et al., 2005). Therefore, seeds of WT and mdm-miR828-overexpression lines were cultured on half-strength Murashige and Skoog solid medium supplemented with 2% sucrose. The anthocyanin content was determined after culture for 4 days. The anthocyanin content in the mdm-miR828-overexpression lines was significantly lower than that of the WT. This result indicated that mdm-miR828 negatively regulated anthocyanin synthesis in *Arabidopsis* (**Figures 3A,B**).

Expression of siR81(−) in the mdm-miR828-overexpression *Arabidopsis* seedlings was increased compared with that of the WT. The expression levels of anthocyanin biosynthesis-related regulatory genes (*AtMYB75*, *AtMYB90*, and *AtMYB113*) and structural genes (*AtCHS*, *AtDFR*, *AtF3H*, and *AtLDOX*) were reduced compared with those of the WT. These results suggested that mdm-miR828 and ath-miR828 performed similar functions and negatively regulated anthocyanin synthesis in *Arabidopsis* (**Figure 3C**).

MdMYB1 Promotes Expression of mdm-miR828

An autoregulatory feedback loop operates in *Arabidopsis*, and up-regulated expression of PAP1/MYB75 can activate miR828 expression (Hsieh et al., 2009; Luo et al., 2012). As a consequence, miR828 generates a greater amount of siR81(−) by directing the cleavage of *TAS4* transcripts, thereby enhancing the cleavage of the target genes *PAP1/MYB75*, *PAP2/MYB90*, and *MYB113* (Hsieh et al., 2009; Luo et al., 2012). In the present experiment, the expression levels of mdm-miR828 and mdm-siR81(−) increased in the late stage of apple coloration (134–146 days; **Figure 1C**), whereas *MdMYB1* maintained a high level of expression in the late stage of coloration (**Figure 4A**). We speculated that MdMYB1 directly regulates expression of

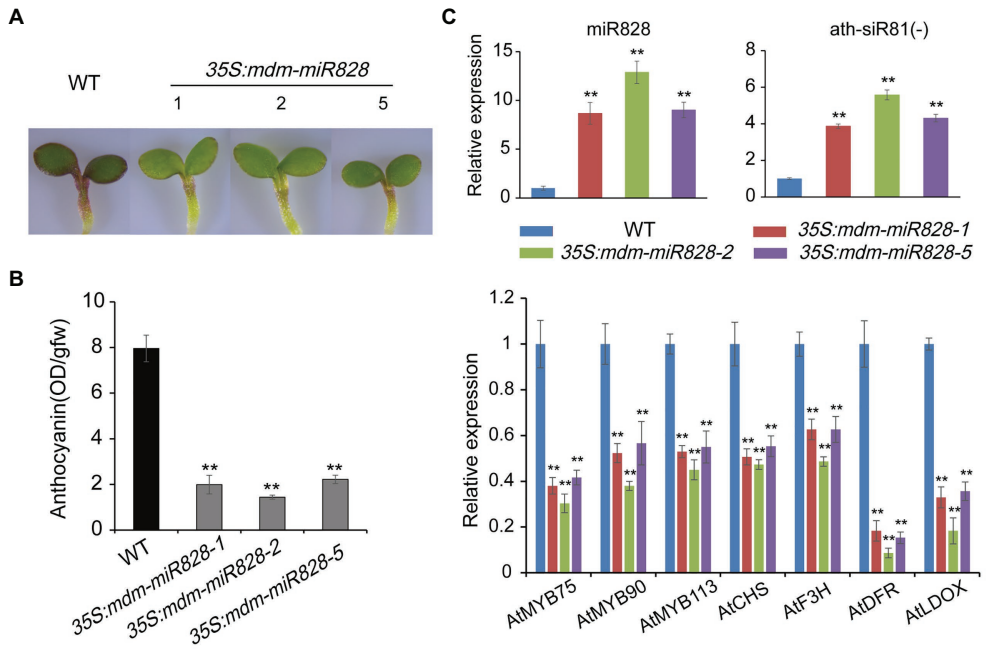

FIGURE 3 | Overexpression of mdm-miR828 inhibits anthocyanin accumulation in *Arabidopsis* seedlings. **(A)** Phenotypes of mdm-miR828-overexpressing and wild-type *Arabidopsis* grown on half-strength Murashige and Skoog medium supplemented with 2% sucrose for 4 days. 35S:mdm-miR828-1, 35S:mdm-miR828-2, and 35S:mdm-miR828-5 represent three transgenic lines. **(B)** Anthocyanin content in mdm-miR828-overexpressing and wild-type *Arabidopsis*. **(C)** Relative expression levels assessed by qRT-PCR analysis showed that mdm-miR828 overexpression affects expression of anthocyanin-related genes in *Arabidopsis*. Error bars represent the standard deviation of three biological replicates. $^{**}p < 0.01$ (Student *t*-test).

mdm-miR828. To test this hypothesis, we used a dual-luciferase assay to assess the correlation between MdMYB1 and the mdm-MIR828 promoter. First, we detected the expression levels of mdm-MIR828a (or pre–mdm-miR828a) and mdm-MIR828b. Expression of mdm-MIR828a was essentially undetectable, whereas the expression level of mdm-MIR828b was increased in the late fruit coloration stage (134–146 days; **Supplementary Figure S3**). Therefore, we cloned the 1,514-bp sequence of the mdm-MIR828b promoter and ligated it into the pGreenII 0800-LUC vector to form the fusion vector *pMIR828b:LUC* (**Figure 4B**). We then used *Agrobacterium* carrying *pMIR828b:LUC* and those harboring *35S:MdMYB1* to cotransform *N. benthamiana* leaves. Compared with the empty plasmid (pGreenII 62-SK) control, transient overexpression of *MdMYB1* significantly up-regulated LUC activity driven by the mdm-MIR828b promoter (**Figure 4C**). This result indicated that MdMYB1 promoted the expression of mdm-miR828. Using the PlantCARE database, we detected multiple potential MYB transcription factor recognition sites in the 645-bp segment upstream of the mdm-MIR828b promoter (**Figure 4D**). A yeast one-hybrid assay showed that MdMYB1 is able to bind to this segment in the mdm-MIR828b promoter (**Figure 4E**).

mdm-miR828 Participates in the Process of High-Temperature Inhibition of Anthocyanin Accumulation

High temperature can inhibit anthocyanin accumulation in apple (Fang et al., 2019). As a potential negative regulator of anthocyanin accumulation, it is not clear whether mdm-miR828 is involved in the inhibition of anthocyanin accumulation in response to high temperature. We placed previously bagged covered "Starkrimson Delicious" fruit in an incubator maintained at either 35 or 23°C (the control). Under continuous light for 24 h, apple anthocyanin accumulation was reduced at 35°C compared with that of the control group at 23°C, and the difference was more obvious at 48 h (**Figures 5A,B**). The expression level of mdm-miR828 was elevated under high temperature compared with that of the control. The expression level of mdm-siR81(−) also increased in response to high temperature, whereas the expression levels of the anthocyanin synthesis-related genes *MdTAS4*, *MdbHLH3*, *MdMYB1*, *MdCHS*, *MdDFR*, *MdANS*, and *MdUFGT* were significantly reduced under high temperature compared with those of the control (**Figure 5C**). These results suggested that mdm-miR828 may be involved in the process of high-temperature inhibition of anthocyanin accumulation.

DISCUSSION

Anthocyanins are an important subclass of flavonoids that not only provide rich pigmentation to flowers and fruits, but also have anticancer, antioxidation, antiaging, and other beneficial properties for human health (Venancio et al., 2017). In addition, anthocyanins are involved in resistance to diverse biotic and abiotic stresses (Guo et al., 2008; Kovinich et al., 2015).

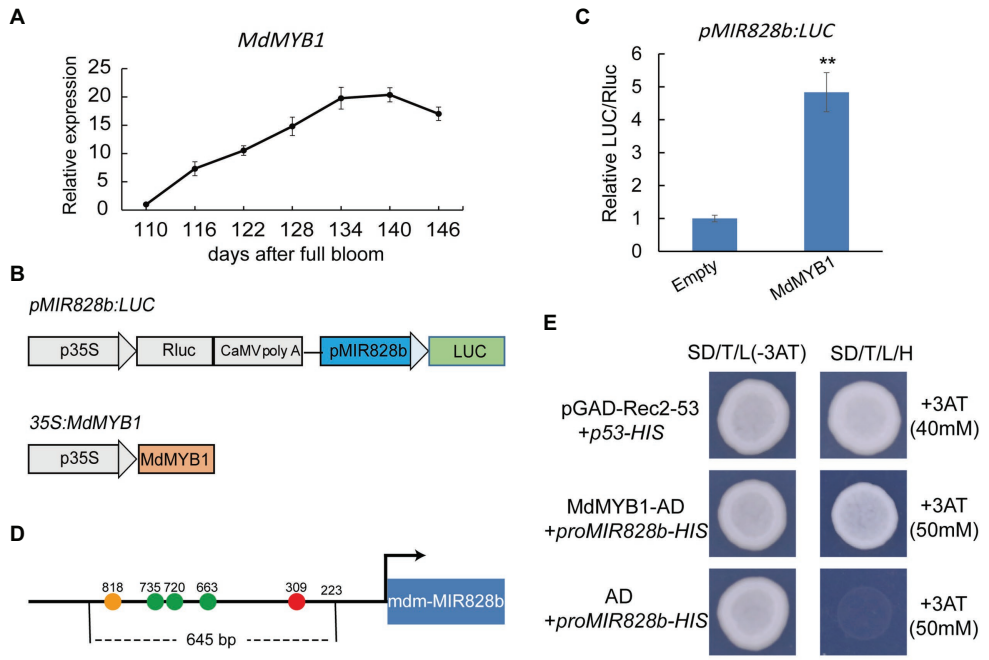

FIGURE 4 | MdMYB1 promotes mdm-MIR828b expression by directly binding to its promoter. **(A)** Expression pattern of *MdMYB1* during the coloration period of apple "Starkrimson Delicious" fruit. **(B)** Schematic representation of the construct used for the dual-luciferase assay. **(C)** Relative LUC/REN ratio. REN, *Renilla* luciferase activity; LUC, firefly luciferase activity. Error bars represent the standard deviation of three biological replicates. $^{**}p < 0.01$ (Student *t*-test). **(D)** Schematic diagram of the mdm-MIR828b promoter structure. Circles represent the potential MYB transcription factor recognition sites: orange (TAACCA), green (CAACCA), and red (TAACTG). The number represents the distance (bp) from the transcription start site. **(E)** Yeast one-hybrid assay for binding of the mdm-MIR828b promoter and MdMYB1. pGAD-Rec2-53 and p53-HIS were used as positive controls.

The synthesis of anthocyanins in plants is strictly regulated by sophisticated regulatory networks. Previous studies of anthocyanin synthesis have predominantly focused on regulation at the transcriptional level, whereas research at the posttranscriptional level is relatively rare. In plants, miRNAs play a vital regulatory role at the posttranscriptional level largely through cleavage of target genes. Therefore, identification of miRNAs associated with fruit coloration is important for improvement of fruit quality.

In plants, the MBW complex, composed of an MYB transcription factor, bHLH transcription factor, and WD40 protein, is highly conserved and coordinately regulates structural genes of the anthocyanin synthesis pathway (Gonzalez et al., 2008). A deletion mutation of *TT8* (a member of the bHLH transcription factor family) in *Arabidopsis* causes a severe decrease in the expression level of *DFR* and *BAN* genes, which causes a change in the seed coat color to yellow (Nesi et al., 2000). In apples, the transcription factor MdbHLH3 induced by low temperature can interact with MdMYB1 and promotes the expression of *MdDFR* and *MdUFGT* structural anthocyanin biosynthesis genes. In addition, MdbHLH3 may bind to the promoter of the regulatory gene *MdMYB1* to promote its expression (Xie et al., 2012). In the present study, during the period of rapid synthesis of anthocyanins, *MdbHLH3* was significantly increased, and overexpression of *MdbHLH3* in the fruit promoted anthocyanin accumulation. These findings are consistent with previous results (Xie et al., 2012) and reveal that *MdbHLH3* is essential for anthocyanin synthesis. The *TAS4* gene is conserved in many dicotyledonous plants. Luo et al. (2012) compared *TAS4* sequences from *A. thaliana*, *Theobroma cacao*, *Euphorbia esula*, *Actinidia chinensis*, *V. vinifera*, and *Mimulus guttatus* and observed that the binding site of miR828 is conserved. In the present study, RACE proved that mdm-miR828 directs cleavage of the *MdTAS4* transcript between the 10th and 11th nucleotides, which indicated that mdm-miR828-mediated cleavage of *MdTAS4* transcripts is conserved in apple. In *Arabidopsis*, ath-siR81(−) can target the anthocyanin-related transcription factors PAP1/MYB75, PAP2/MYB90, and MYB113. However, there is no sequence complement between mdm-siR81(−) and the anthocyanin-related transcription factor MdMYB1 in apples. Interestingly, mdm-siR81(−) is complementary to *MdbHLH3* sequence, whereas mdm-siR81(−) and *MdbHLH3* show an inverse expression profile to each other. In addition, the RACE experiment demonstrated that mdm-siR81(−)-directed cleavage of *MdbHLH3* transcripts occurs in the peel of apple fruit.

Regulation of anthocyanin synthesis by miRNAs has been observed in many plant species, but the effects of miRNAs on anthocyanins differ. Inhibition of miR156 expression in *Arabidopsis* results in increased expression of the target gene *SPL9* and thus inhibits accumulation of anthocyanins (Gou et al., 2011; Cui et al., 2014). In contrast, overexpression of mdm-miR156a in apple causes decreased expression of the *SPL2-like* and *SPL33* genes and inhibition of anthocyanin

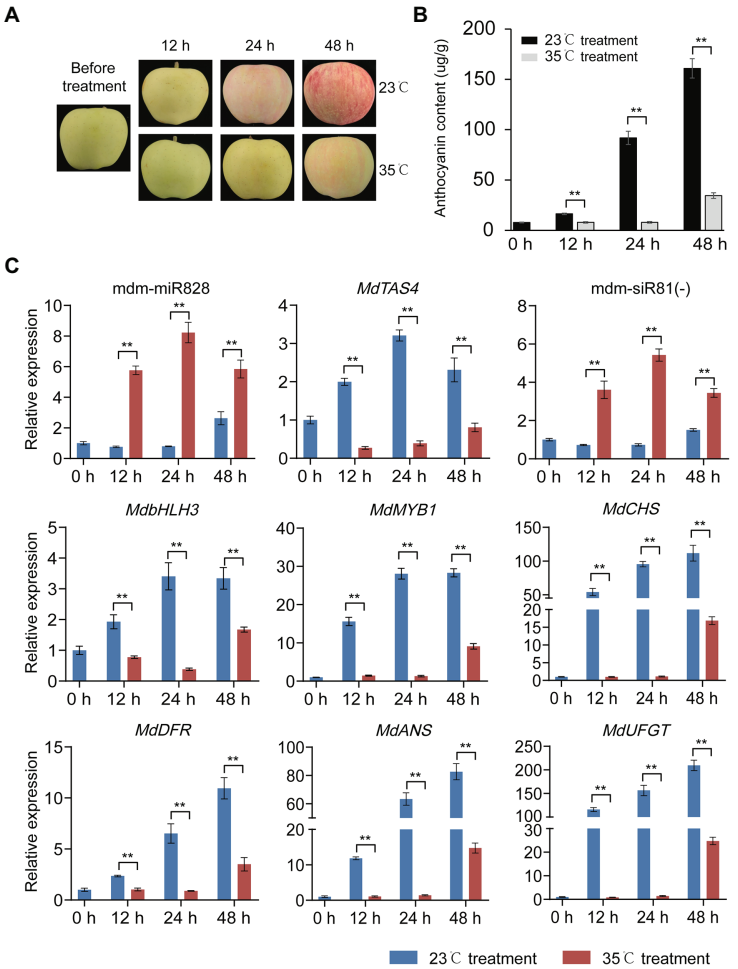

FIGURE 5 | High temperature inhibits anthocyanin accumulation in apple peel. **(A)** Change in color of "Starkrimson Delicious" peel under high temperature. **(B)** Anthocyanin content in the peel under high temperature. **(C)** Relative expression level of mdm-miR828, mdm-siR81(−), and anthocyanin-related genes (*MdTAS4*, *MdbHLH3*, *MdbMYB1*, *MdCHS*, *MdDFR*, *MdANS*, and *MdUFGT*) under high temperature. Error bars represent the standard deviation of three biological replicates. $^{**}p < 0.01$ (Student *t*-test).

accumulation (Yang et al., 2019). Overexpression of miR858 in *Arabidopsis* targets the anthocyanin inhibitor MYBL2 to promote anthocyanin accumulation (Wang et al., 2016b). However, miR858 in tomato plays a negative role in the biosynthesis of anthocyanins; blockage of miR858 binding increases *SlMYB7-like* transcript levels and ultimately promotes accumulation of anthocyanins (Jia et al., 2015). In *Arabidopsis*, siR81(−) derived from *TAS4* and miR828 together target the positive regulatory genes *PAP1*, *PAP2*, and *MYB113* for anthocyanin synthesis (Rajagopalan et al., 2006). Overexpression of miR828 reduces the accumulation of anthocyanins in *Arabidopsis* (Yang et al., 2013). However, miR828 targets the anthocyanin inhibitor MYB114 in grape to positively regulate anthocyanin accumulation (Tirumalai et al., 2019). In the present study, the mdm-miR828 expression level was relatively low during the rapid fruit coloration period (116 to 134 days), and mdm-siR81(−) expression was also relatively low. In addition, overexpression of mdm-miR828 in apple fruit inhibited anthocyanin synthesis, and overexpression of mdm-miR828 in *Arabidopsis* inhibited anthocyanin accumulation. These results indicated that mdm-miR828 negatively regulated anthocyanin synthesis in apple.

In the *Arabidopsis* mutant *pap1-D*, in which *PAP1* is strongly up-regulated, high quantities of anthocyanins accumulate, but the expression levels of ath-miR828 and ath-siR81(−) also increase (Hsieh et al., 2009). Phosphorus deficiency, nitrogen deficiency, and exogenous sugars treatment can increase the expression of the PAP1/MYB75 and PAP2/MYB90 MYB transcription factors, which in turn activates the expression of anthocyanin structural genes and increases anthocyanin synthesis in *Arabidopsis*. In these plants with high anthocyanin content, the expression levels of ath-miR828 and ath-siR81(−) increased, which induced negative feedback regulation of AtMYB75, AtMYB90, and AtMYB113 expression. These results indicate that a feedback regulatory mechanism in *Arabidopsis* may control anthocyanin accumulation (Hsieh et al., 2009; Luo et al., 2012). In the current study, the anthocyanin tended to remain stable

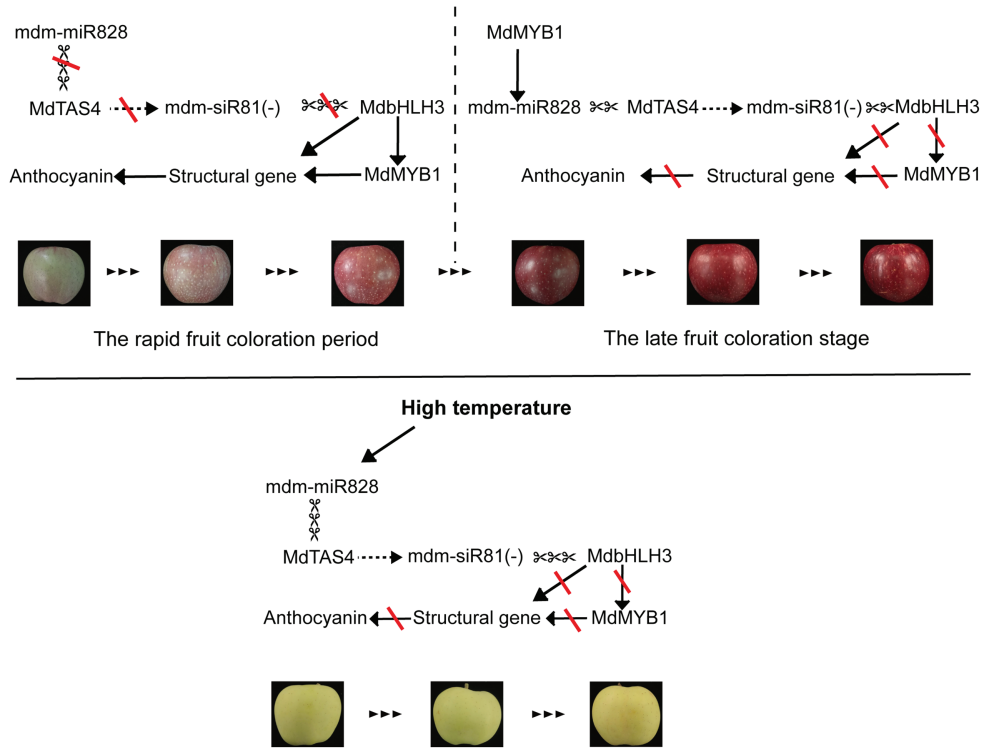

FIGURE 6 | Proposed model for mdm-miR828 participation in a feedback regulation pathway to balance anthocyanin accumulation in the peel of apple fruit. A solid arrow indicates activation, a dashed arrow indicates derivation, scissors indicate cleavage, and red lines indicate closure.

in the late stage of fruit coloration, whereas *MdMYB1* maintained a high level of expression from 134 to 140 days after blooming, and at this stage, the expression level of mdm-miR828 began to increase. The dual-luciferase and yeast one-hybrid assays showed that MdMYB1 is capable of binding to the promoter of mdm-MIR828b to induce its expression. These results indicated that a feedback regulatory mechanism operates in apple. In addition, *MdMYB1* and mdm-miR828 expression levels were increased in fruit overexpressing *MdbHLH3*, which suggested that the feedback pathway was activated.

In addition to sucrose, light (Takos et al., 2006) and low temperature (Xie et al., 2012; Gaiotti et al., 2018) promote anthocyanin synthesis. However, high temperature inhibits biosynthesis of anthocyanins in apple (Wang et al., 2016a). In the present experiment, anthocyanin accumulation under high temperature (35°C) was significantly inhibited compared with that of the control (23°C). The expression levels of mdm-miR828 and mdm-siR81(−) were higher under high temperature compared with those of the control, whereas the expression level of *MdbHLH3* was reduced under high temperature compared with that of the control. These findings indicated that mdm-miR828 may be involved in the process of inhibiting anthocyanin accumulation under high temperature.

Based on the present results, we propose a model for mdm-miR828 participation in a feedback regulatory mechanism to balance anthocyanin accumulation in the peel of apple fruit (**Figure 6**). In this model, mdm-miR828 regulates the expression of *MdbHLH3* by cleaving *MdTAS4* to derive mdm-siR81(−). During the period of rapid fruit coloration, the expression level of mdm-miR828 is low, and *MdbHLH3* and *MdMYB1* are expressed in large quantities, which promotes the rapid synthesis of anthocyanins. In the late fruit coloration stage, the high-level expression of MdMYB1, which binds to the promoter of mdm-MIR828b, promotes the expression of mdm-miR828, which in turn inhibits the expression of *MdbHLH3* so that anthocyanins do not accumulate indefinitely. In addition, high temperature can also induce the expression of mdm-miR828 and mdm-siR81(−), and then the highly expressed mdm-siR81(−) cleaves *MdbHLH3* to prevent the accumulation of anthocyanins.

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

BZ, H-JY, and Z-YZ conceived the original screening and research plans. BZ, Y-ZY, Z-ZZ, and Y-NL supervised the experiments. BZ, H-JY, and DQ analyzed the data. BZ wrote the article. H-JY revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the earmarked fund for the National Natural Science Foundation of China (32072555), and the Modern Agro-industry Technology Research System of China (CARS-27).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.608109/full#supplementary-material

REFERENCES

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

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

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

Allen, E., Xie, Z., Gustafson, A. M., and Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121, 207–221. doi: 10.1016/j.cell.2005.04.004

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

Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* 39, 366–380. doi: 10.1111/j.1365-313X.2004.02138.x

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12, 2383–2394. doi: 10.1105/tpc.12.12.2383

Chaves-Silva, S., Santos, A. L. D., Chalfun-Júnior, A., Zhao, J., Peres, L. E. P., and Benedito, V. A. (2018). Understanding the genetic regulation of anthocyanin biosynthesis in plants - tools for breeding purple varieties of fruits and vegetables. *Phytochemistry* 153, 11–27. doi: 10.1016/j.phytochem.2018.05.013

Chen, C., Ridzon, D. A., Broomer, A. J., Zhou, Z., Lee, D. H., Nguyen, J. T., et al. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* 33:e179. doi: 10.1093/nar/gni178

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Cui, L. G., Shan, J. X., Shi, M., Gao, J. P., and Lin, H. X. (2014). The miR156-SPL9-DFR pathway coordinates the relationship between development and abiotic stress tolerance in plants. *Plant J.* 80, 1108–1117. doi: 10.1111/tpj.12712

Cuperus, J. T., Fahlgren, N., and Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *Plant Cell* 23, 431–442. doi: 10.1105/tpc.110.082784

de Vetten, N., Quattrocchio, F., Mol, J., and Koes, R. (1997). The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Genes Dev.* 11, 1422–1434. doi: 10.1101/gad.11.11.1422

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

Fang, H., Dong, Y., Yue, X., Chen, X., He, N., Hu, J., et al. (2019). MdCOL4 interaction mediates crosstalk between UV-B and high temperature to control fruit coloration in apple. *Plant Cell Physiol.* 60, 1055–1066. doi: 10.1093/pcp/pcz023

Gaiotti, F., Pastore, C., Filippetti, I., Lovat, L., Belfiore, N., and Tomasi, D. (2018). Low night temperature at veraison enhances the accumulation of anthocyanins in Corvina grapes (*Vitis vinifera* L.). *Sci. Rep.* 8:8719. doi: 10.1038/s41598-018-26921-4

Goff, S. A., Cone, K. C., and Chandler, V. L. (1992). Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. *Genes Dev.* 6, 864–875. doi: 10.1101/gad.6.5.864

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

Gou, J. Y., Felippes, F. F., Liu, C. J., Weigel, D., and Wang, J. W. (2011). Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell* 23, 1512–1522. doi: 10.1105/tpc.111.084525

Guo, J., Han, W., and Wang, M. -H. (2008). Ultraviolet and environmental stresses involved in the induction and regulation of anthocyanin biosynthesis: a review. *Afr. J. Biotechnol.* 7, 4966–4972.

Hichri, I., Heppel, S. C., Pillet, J., Léon, C., Czemmel, S., Delrot, S., et al. (2010). The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. *Mol. Plant* 3, 509–523. doi: 10.1093/mp/ssp118

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7, 1071–1083. doi: 10.2307/3870058

Hsieh, L. C., Lin, S. I., Shih, A. C., Chen, J. W., Lin, W. Y., Tseng, C. Y., et al. (2009). Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing. *Plant Physiol.* 151, 2120–2132. doi: 10.1104/pp.109.147280

Jia, X., Shen, J., Liu, H., Li, F., Ding, N., Gao, C., et al. (2015). Small tandem target mimic-mediated blockage of microRNA858 induces anthocyanin accumulation in tomato. *Planta* 242, 283–293. doi: 10.1007/s00425-015-2305-5

Kim, S. H., Lee, J. R., Hong, S. T., Yoo, Y. K., An, G., and Kim, S. R. (2003). Molecular cloning and analysis of anthocyanin biosynthesis genes preferentially expressed in apple skin. *Plant Sci.* 165, 403–413. doi: 10.1016/S0168-9452(03)00201-2

Kobayashi, S., Goto-Yamamoto, N., and Hirochika, H. (2004). Retrotransposon-induced mutations in grape skin color. *Science* 304:982. doi: 10.1126/science.1095011

Kovinich, N., Kayanja, G., Chanoca, A., Otegui, M. S., and Grotewold, E. (2015). Abiotic stresses induce different localizations of anthocyanins in *Arabidopsis*. *Plant Signal. Behav.* 10:e1027850. doi: 10.1080/15592324.2015.1027850

Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T. K., et al. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in rosaceae. *BMC Plant Biol.* 10:50. doi: 10.1186/1471-2229-10-50

Liu, Y., Che, F., Wang, L., Meng, R., Zhang, X., and Zhao, Z. (2013). Fruit coloration and anthocyanin biosynthesis after bag removal in non-red and red apples (*Malus* × *domestica* Borkh.). *Molecules* 18, 1549–1563. doi: 10.3390/molecules18021549

Liu, R., Lai, B., Hu, B., Qin, Y., Hu, G., and Zhao, J. (2016). Identification of MicroRNAs and their target genes related to the accumulation of anthocyanins in *Litchi chinensis* by high-throughput sequencing and degradome analysis. *Front. Plant Sci.* 7:2059. doi: 10.3389/fpls.2016.02059

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., et al. (2017). Advances in the MYB-bHLH-WD repeat (MBW) pigment regulatory model: addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. *Plant Cell Physiol.* 58, 1431–1441. doi: 10.1093/pcp/pcx075

Luo, Q. J., Mittal, A., Jia, F., and Rock, C. D. (2012). An autoregulatory feedback loop involving *PAP1* and *TAS4* in response to sugars in *Arabidopsis*. *Plant Mol. Biol.* 80, 117–129. doi: 10.1007/s11103-011-9778-9

Matus, J. T., Poupin, M. J., Cañón, P., Bordeu, E., Alcalde, J. A., and Arce-Johnson, P. (2010). Isolation of WDR and bHLH genes related to flavonoid synthesis in grapevine (*Vitis vinifera* L.). *Plant Mol. Biol.* 72, 607–620. doi: 10.1007/s11103-010-9597-4

Medina-Puche, L., Cumplido-Laso, G., Amil-Ruiz, F., Hoffmann, T., Ring, L., Rodríguez-Franco, A., et al. (2014). *MYB10* plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria* x *ananassa* fruits. *J. Exp. Bot.* 65, 401–417. doi: 10.1093/jxb/ert377

Nesi, N., Debeaujon, I., Jond, C., Pelletier, G., Caboche, M., and Lepiniec, L. (2000). The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in *Arabidopsis* siliques. *Plant Cell* 12, 1863–1878. doi: 10.1105/tpc.12.10.1863

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

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

Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* 20, 3407–3425. doi: 10.1101/gad.1476406

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

Shirley, B. W., Kubasek, W. L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F. M., et al. (1995). Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *Plant J.* 8, 659–671. doi: 10.1046/j.1365-313X.1995.08050659.x

Song, L., Axtell, M. J., and Fedoroff, N. V. (2010). RNA secondary structural determinants of miRNA precursor processing in *Arabidopsis*. *Curr. Biol.* 20, 37–41. doi: 10.1016/j.cub.2009.10.076

Spelt, C., Quattrocchio, F., Mol, J. N., and Koes, R. (2000). anthocyanin1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* 12, 1619–1632. doi: 10.1105/tpc.12.9.1619

Sun, Y., Qiu, Y., Duan, M., Wang, J., Zhang, X., Wang, H., et al. (2017). Identification of anthocyanin biosynthesis related microRNAs in a distinctive Chinese radish (*Raphanus sativus* L.) by high-throughput sequencing. *Mol. Gen. Genomics.* 292, 215–229. doi: 10.1007/s00438-016-1268-y

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

Teng, S., Keurentjes, J., Bentsink, L., Koornneef, M., and Smeekens, S. (2005). Sucrose-specific induction of anthocyanin biosynthesis in *Arabidopsis* requires the MYB75/PAP1 gene. *Plant Physiol.* 139, 1840–1852. doi: 10.1104/pp.105.066688

Tian, J., Peng, Z., Zhang, J., Song, T., Wan, H., Zhang, M., et al. (2015). McMYB10 regulates coloration via activating *McF3'H* and later structural genes in ever-red leaf crabapple. *Plant Biotechnol. J.* 13, 948–961. doi: 10.1111/pbi.12331

Tirumalai, V., Swetha, C., Nair, A., Pandit, A., and Shivaprasad, P. V. (2019). miR828 and miR858 regulate VvMYB114 to promote anthocyanin and flavonol accumulation in grapes. *J. Exp. Bot.* 70, 4775–4792. doi: 10.1093/jxb/erz264

Venancio, V. P., Cipriano, P. A., Kim, H., Antunes, L. M., Talcott, S. T., and Mertens-Talcott, S. U. (2017). Cocoplum (*Chrysobalanus icaco* L.) anthocyanins exert anti-inflammatory activity in human colon cancer and non-malignant colon cells. *Food Funct.* 8, 307–314. doi: 10.1039/C6FO01498D

Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* 136, 669–687. doi: 10.1016/j.cell.2009.01.046

Wang, Y., Wang, Y., Song, Z., and Zhang, H. (2016b). Repression of *MYBL2* by both microRNA858a and HY5 leads to the activation of anthocyanin biosynthetic pathway in *Arabidopsis*. *Mol. Plant* 9, 1395–1405. doi: 10.1016/j.molp.2016.07.003

Wang, L., Zeng, J. H., Song, J., Feng, S. J., and Yang, Z. M. (2015). miRNA778 and *SUVH6* are involved in phosphate homeostasis in *Arabidopsis*. *Plant Sci.* 238, 273–285. doi: 10.1016/j.plantsci.2015.06.020

Wang, N., Zhang, Z., Jiang, S., Xu, H., Wang, Y., Feng, S., et al. (2016a). Synergistic effects of light and temperature on anthocyanin biosynthesis in callus cultures of red-fleshed apple (*Malus sieversii* f. *niedzwetzkyana*). *Plant Cell Tissue Org. Cult.* 127, 217–227. doi: 10.1007/s11240-016-1044-z

Xia, R., Zhu, H., An, Y. Q., Beers, E. P., and Liu, Z. (2012). Apple miRNAs and tasiRNAs with novel regulatory networks. *Genome Biol.* 13:R47. doi: 10.1186/gb-2012-13-6-r47

Xie, X. B., Li, S., Zhang, R. F., Zhao, J., Chen, Y. C., Zhao, Q., et al. (2012). The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. *Plant Cell Environ.* 35, 1884–1897. doi: 10.1111/j.1365-3040.2012.02523.x

Yang, F., Cai, J., Yang, Y., and Liu, Z. (2013). Overexpression of microRNA828 reduces anthocyanin accumulation in *Arabidopsis*. *Plant Cell Tissue Org. Cult.* 115, 159–167. doi: 10.1007/s11240-013-0349-4

Yang, T., Ma, H., Zhang, J., Wu, T., Song, T., Tian, J., et al. (2019). Systematic identification of long noncoding RNAs expressed during light-induced anthocyanin accumulation in apple fruit. *Plant J.* 100, 572–590. doi: 10.1111/tpj.14470

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.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 13 January 2021
doi: 10.3389/fphar.2020.616378

Formononetin Activates the Nrf2/ARE Signaling Pathway Via Sirt1 to Improve Diabetic Renal Fibrosis

Kai Zhuang[1], Xiyu Jiang[1], Renbin Liu[2], Cunsi Ye[1], Yumei Wang[1], Yunhan Wang[3], Shijian Quan[1] and Heqing Huang[4]**

[1]School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China, [2]Department of Traditional Chinese Medicine, Renmin Hospital, Hubei University of Medicine, Shiyan, China, [3]Science and Technology Innovation Center, Guangzhou University of Chinese Medicine, Guangzhou, China, [4]Laboratory of Pharmacology and Toxicology, School of Pharmaceutical Science, Sun Yat-Sen University, Guangzhou, China

OPEN ACCESS

Edited by:
M. Carmen González-Mas,
University of Valencia, Spain

Reviewed by:
Yonghong Shi,
Hebei Medical University, China
Vittorio Calabrese,
University of Catania, Italy

***Correspondence:**
Shijian Quan
quansj@gzucm.edu.cn
Heqing Huang
huangheq@mail.sysu.edu.cn

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: 12 October 2020
Accepted: 30 November 2020
Published: 13 January 2021

Citation:
Zhuang K, Jiang X, Liu R, Ye C, Wang Y, Wang Y, Quan S and Huang H (2021) Formononetin Activates the Nrf2/ARE Signaling Pathway Via Sirt1 to Improve Diabetic Renal Fibrosis. Front. Pharmacol. 11:616378. doi: 10.3389/fphar.2020.616378

Oxidative stress is the main factor responsible for the induction of diabetic renal fibrosis. Thus, improving the state of oxidative stress can effectively prevent the further deterioration of diabetic nephropathy (DN). Previous research has shown that formononetin (FMN), a flavonoid with significant antioxidant activity and Sirt1 activation effect, can improve diabetic renal fibrosis. However, the exact mechanisms underlying the effect of FMN on diabetic renal fibrosis have yet to be elucidated. In this study, we carried out *in vivo* experiments in a db/db (diabetic) mouse model and demonstrated that FMN activated the nuclear factor E2-related factor 2 (Nrf2)/antioxidant response element (ARE) signaling pathway and improved oxidative stress by increasing levels of sirtuin-1 (Sirt1) protein level in renal tissue. We also found that this process reversed the up-regulation of fibronectin (FN) and intercellular adhesion molecule 1 (ICAM-1) and led to an improvement in renal insufficiency. *In vitro* results further showed that FMN significantly reversed the upregulation of FN and ICAM-1 in glomerular mesangial cells (GMCs) exposed to high glucose. FMN also promoted the expression of Nrf2 and widened its nuclear distribution. Thus, our data indicated that FMN inhibited hyperglycemia-induced superoxide overproduction by activating the Nrf2/ARE signaling pathway. We also found that FMN up-regulated the expression of Sirt1 and that Sirt1 deficiency could block the activation of the Nrf2/ARE signaling pathway in GMCs induced by high glucose. Finally, we found that Sirt1 deficiency could reverse the down-regulation of FN and ICAM-1 induced by FMN. Collectively, our data demonstrated that FMN up-regulated the expression of Sirt1 to activate the Nrf2/ARE signaling pathway, improved oxidative stress in DN to prevent the progression of renal fibrosis. Therefore, FMN probably represents an efficient therapeutic option of patients with DN.

Keywords: formononetin, diabetic nephropathy, renal fibrosis, oxidative stress, Sirt1/Nrf2/ARE signaling pathways

INTRODUCTION

Diabetic nephropathy (DN) is one of the most serious chronic microvascular complications and the key factor that can lead to end-stage nephropathy (Perkovic et al., 2019; Ruiz-Ortega et al., 2020). Although the occurrence and progression of DN can be delayed by the strict control of blood glucose levels and blood pressure levels, these measures cannot completely prevent patients with diabetes from developing into renal failure. Consequently, it is vital to develop new drugs to prevent and treat DN. Renal fibrosis is the main pathological feature of DN (Wei-Jian et al., 2015). In the diabetic state, glomerular mesangial cell proliferation, along with glomerular hypertrophy, gradually lead to the deposition of extracellular matrix (ECM) in the glomerular mesangial region. This leads to the progressive accumulation of inflammatory fibrosis components such as fibronectin (FN) and intercellular cell adhesion molecule-1 (ICAM-1), and can develop into renal fibrosis with glomerulosclerosis as the predominant feature eventually. Oxidative stress is considered to be the main pathogenic factor underlying renal fibrosis in diabetes mellitus (Kultigin, 2017; Sagoo and Gnudi, 2018). In the pathological state of diabetes, high levels of glucose stimulates cells to produce a large number of reactive oxygen species (ROS), which activates a variety of downstream inflammatory signaling pathways, thus inducing and accelerating the occurrence of renal inflammatory fibrosis. Therefore, regulating or inhibiting oxidative stress is an effective method to prevent the progression of diabetic renal fibrosis and thus prevent DN.

The nuclear factor E2-related factor 2 (Nrf2)/antioxidant response element (ARE) signaling pathway is the most important intracellular antioxidant pathway (Hu et al., 2020; Nwakiban et al., 2020). And the activation of Nrf2/ARE pathway can promote the expression of a range of downstream antioxidant enzymes, including heme oxygenase 1 (HO-1) and superoxide dismutase 1 (SOD-1), thereby reducing ROS level and resisting the renal fibrosis injury induced by oxidative stress (Ahmed et al., 2017). However, during end-stage diabetes, the damage caused by the endogenous antioxidant system inevitable leads to the attenuation of Nrf2. Therefore, the activation of Nrf2/ARE signaling pathway exerts a vital protective effect on DN. The family of mammal silent information regulator two proteins (Sirtuins) are deacetylase enzymes that are dependent on nicotinamide adenine dinucleotide (NAD+). Researches have shown that Sirt1 plays a key role in the activation of Nrf2/ARE signaling and the protection of DN (Kim et al., 2018; Arioz et al., 2019; Ma et al., 2019). In addition, studies have shown that Sirt1 could up-regulate the levels of Nrf2 protein and improve oxidative stress (Ma et al., 2018). As such, Sirt1 is considered to be a promising target for the prevention and treatment of DN.

Formononetin (FMN) is the main active ingredient of *Astragalus membranaceus*, Niudali, and other forms of Traditional Chinese medicines with known antioxidant activities (Jin et al., 2017). Formononetin is also known to improve hepatic cholestasis by up-regulating the expression of Sirt1 (Yang et al., 2019). Recent studies have also shown that FMN could improve renal function in diabetic animal models, although the specific mechanisms involved in this effect remain unclear (Qiu et al., 2017). Thus, current evidences indicate that the Sirt1/Nrf2/ARE signaling pathway is important in the prevention and treatment of DN, and that the antioxidant activity of FMN and the agonistic effect of FMN on Sirt1 probably plays a significant role in this process. We hypothesized that the activation of Sirt1/Nrf2/ARE signaling pathway induced by FMN could improve oxidative stress and that FMN probably prevents the progression of diabetic renal fibrosis and thereby represents a useful therapeutic option for patients with DN. In the present study, we used the db/db (diabetic) mouse model and demonstrated that FMN effectively improved the indices of inflammatory fibrosis and oxidative stress. We also found that FMN up-regulated the expression of Sirt1 protein in the renal tissue of db/db diabetic mice and in glomerular mesangial cells (GMCs) under the conditions of high glucose, thus activating the Nrf2/ARE signaling pathway. Our results therefore confirmed that FMN improved diabetic renal fibrosis by activating the of Sirt1/Nrf2/ARE signaling pathway and by relieving oxidative stress.

MATERIALS AND METHODS

Reagents and Antibodies

FMN (purity > 98.0%, HPLC), metformin hydrochloride (Met) (purity > 98.0%, HPLC), and sodium carboxymethyl cellulose (CMC) (purity > 98.0%, HPLC) were provided by Vicci Biotechnology Co., Ltd. (Sichuan, China). Resveratrol (Res) (purity > 99.7%, HPLC) was purchased from MedChemExpress (Monmouth Junction, United States). FN (catalog: 66042-1-1 g, diluted 1:1,000), ICAM-1 (catalog: 16174-1-AP, diluted 1:1,000), Nrf2(catalog: 16,396-1-AP, diluted 1:1,000), Keleh-like ECH-associated protein (Keap1) (catalog: 10503-2-AP, diluted 1:2,500), HO-1 (catalog: 10701-1-AP, diluted 1:2,000) and SOD-1 (catalog: 10269-1-AP, diluted 1:2,000) were purchased from Proteintech Group (Wuhan, China). Sirt1 (catalog: ab110304, diluted 1:1,000) and goat anti-rabbit IgG H&L (Alexa Fluor® 594) (catalog: ab150080) were purchased from Abcam (Cambridge, MA, United States). A superoxide anion fluorescent probe (catalog: S0063), Methylthiazolyldiphenyl-tetrazolium bromide (MTT, catalogue: ST316) and primary anti-diluent, were all provided by Shanghai Biyuntian Biotechnology Co., Ltd. (Shanghai, China). Paraformaldehyde (4%) (catalog: AR1,068) was purchased from Wuhan Baodu Biological Engineering Co., Ltd. (Wuhan, China). RNAiMAX (catalogue: 13778-150) was provided by Seymour Fisher Technologies (Rockford, IL, United States). PAS dye, Masson's dye, and hematoxylin and eosin (H&E) dye, were provided by Servicebio (Wuhan, China). For the animal work, FMN was dissolved in 0.5% sodium carboxymethyl cellulose. For the *in vitro* work, FMN and Res were both dissolved in Dimethyl sulfoxide (DMSO); Met was dissolved in distilled water.

Animal Experiments

The mice used in this research were provided by Changzhou Cavins Experimental Animals Co., Ltd. (License number: SCXK 2016–0010). Male db/db mice were divided into five groups: wild db/m (normal control group), db/db (diabetes model group), Met (positive control group; 100 mg/kg), and FMN (FMN treatment group; 25 and 50 mg/kg). Mice in the Met and FMN groups received drugs by oral gavage for eight weeks. Blood and kidney samples were collected from all mice at the end of the eighth week for biochemical and histopathological analyses. All mice were bred at the Laboratory Animal Center of Guangzhou University of Chinese Medicine (License Number: SCXK 2018-0085). The temperature of the animal room was controlled at 23 ± 1°C, relative humidity was controlled at 55 ± 5%, and the photoperiod was controlled at 12 h light/12 h dark. Food and water were available *ad libitum*. All animal experiments are conducted in accordance with relevant regulations relating to experimental animal ethics and were approved by the Animal Experiment Management and Use Committee of Guangzhou University of Traditional Chinese Medicine (Experimental approval reference number: 00222199).

Biochemical Analyses

Serum levels of blood urea nitrogen (BUN), creatinine (Cr), triglyceride (TG) and total cholesterol (TC) were detected using appropriate detection kits in accordance with the manufacturer's instructions.

Cell Culture

The glomeruli from primary GMCs and Sprague-Dawley (SD) rats were isolated and identified by a specificity assay, as described earlier (Xiao et al., 2020). Prior to experiments, GMCs were placed in an incubator set to 37°C and 5% CO_2. We also supplemented the GMCs with 10% fetal bovine serum (FBS) dulbecco's modified eagle medium (DMEM) to provide appropriate conditions for growth. The GMCs were then cultured for 2–3 days until they achieved 90% aggregation and were then subcultured at a ratio of 1:6. Cells from the 5th to 14th passages were used for experiments. FMN was dissolved in 100 mM DMSO to provide the stock liquor; this was stored at −20 C. At around 80% confluency, the cells were incubated in sera for 12–16 h and then exposed to hyper-glycemic conditions or FMN for a specified period of time. In this study, the control groups experienced 5.6 mM glucose while the model group experienced 30 mM glucose.

MTT Assay

In this study, 3-(4, 5-dimethylthiazole-2-yl)-2, 5-diphenyl tetrazolium bromide (MTT, Sigma, United States) was used to evaluate the effect of FMN on the activity of GMCs cells and to determine the appropriate concentration of FMN. Cells were first inoculated in a 96-well plate. Once adherent to the walls of the plate, the cells were cultured in serum-free medium for 12 h and then treated with or without different concentrations of FMN (0, 5, 10, 20, 40, 60, 80, 100 and 150 μM) for 24 h. Next, 20 μM MTT (5 mg/ml) was then added to each well and incubated at 37°C for 4 h before discarding the medium carefully. Finally, DMSO (200 μm) was added to each well to dissolve formazan crystals and a microplate reader (Bio-Tek, United States) was used to measure the absorbance of each well at an optimized wavelength of 570 nm.

Detection of Intracellular Superoxide

We used a fluorescence probe, dihydroethidium (DHE, Beyotime, Shanghai, China), to detect intracellular levels of ROS. After different treatments (according to experimental requirements), the GMCs cells were washed twice with phosphate buffer solution (PBS); the cells were then incubated with DHE (10 μM) in serum-free DMEM at room temperature for 30 min. Finally, we determined fluorescence intensity by high-content screening (ArrayScan VTI 600 Plus, Thermo Fisher, Rockford, United States).

Small Interfering RNA and Transient Transfection

Next, we designed a specific small interfering RNA (siRNA) to target Nrf2, along with a negative control; these siRNAs were synthesized by Genepharma (Shanghai, China). The sequences for Sirt1-siRNA were as follows: sense: 5′-CCAGUAGCACUAAUUCCAATT-3′, antisense: 5′-UUGGAAUUAGUGCUACUGGTT-3′. GMCs were inoculated in a 60 mm plate 24 h before the experiment began. Next, we used the RNA iMAX transfection reagent to transfect the cells (at 60% confluency) with appropriate siRNA constructs; the transfection reagent was used in accordance with the manufacturer's instructions. After 36 h, the medium was removed and 2 ml serum-free DMEM was added to restore cell growth for a further 12 h. Following stimulation, the cells were harvested, and the next experiment was carried out.

Western Blotting

For western blotting, cultured GMCs or animal tissues were first washed twice with pre-cooled PBS, followed by the addition of RIPA lysis buffer combined with a mixture of proteases or phosphatase inhibitors to help lyse the total protein content of each tissue sample. An equivalent amount of total protein from cells or tissues were then separated by 8–12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and then transferred to PVDF membranes (Bio Rad, Hercules, CA, United States). Membranes were first blocked with 5% skimmed milk and then incubated with appropriate primary antibodies at 4 C overnight. The following morning, membranes were washed and then incubated for 1 hat room temperature with the corresponding secondary antibodies. Finally, we used a GE ImageQuant LAS4000mini (GE Healthcare; Waukesha, United States) to observe target bands and then analyzed the relative gray values of different bands with Quantity One protein analysis software (Bio Rad, Hercules, CA, United States).

Immunofluorescence

GMCs were inoculated on glass cover slips and placed in a 24-well plate in which the cells were exposed to different stimuli. Following treatment, the cells were washed three times in PBS (5 min per wash). The GMCs were then fixed with 4% paraformaldehyde at room temperature for 15 min, dried for 10 min and permeabilized with 1% Triton X-100 dissolved in PBS for 10 min. After washing with PBS, the cells were blocked with goat serum at room temperature for 1 h and then incubated overnight with a rabbit polyclonal antibody raised against Nrf2 (1:200) in a refrigerator at 4 C. The following day, the GMCs were

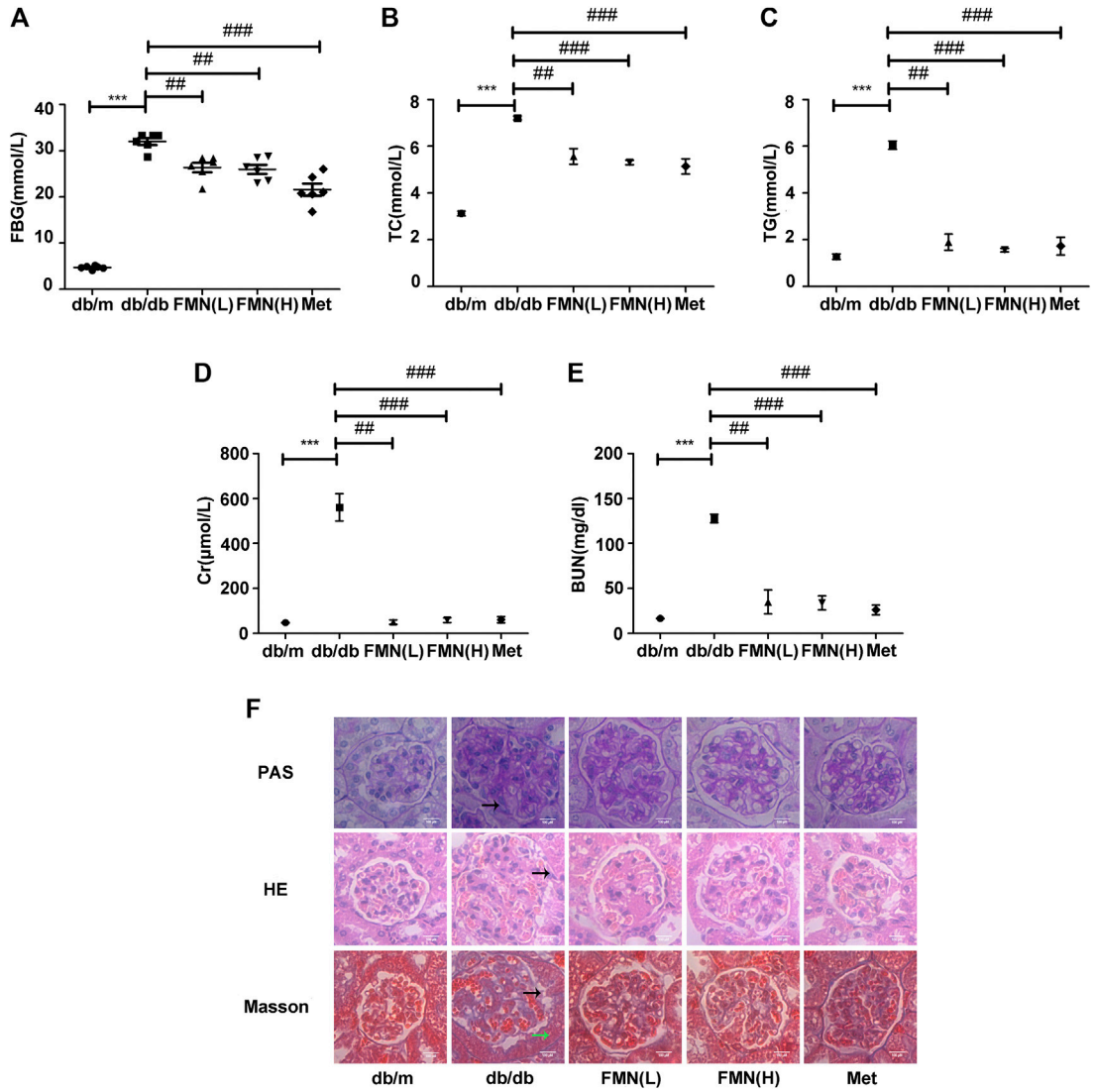

FIGURE 1 | FMN significantly improved indicators of renal function in db/db mice and reduced renal fibrosis. **(A–E)** Levels of FBG, TC, TG, BUN, and Cr, in experimental mice. Data are expressed as ± SD. n = 6, Diabetes: diabetes group; FMN (L): FMN treatment group (low dose: 25 mg/kg); FMN **(H)**: FMN treatment group (medium dose: 50 mg/kg); Met: metformin treatment group (100 mg/kg). ***$p < 0.001$ vs. db/m, ##$p < 0.01$, ###$p < 0.001$ vs. db/db. **(F)** Histopathological analysis of glomeruli using H&E, PAS, and Masson's staining (400×).

incubated with a secondary antibody for 1 h [Goat anti-rabbit IgG H&L (Alexa Fluor® 594)]. Cell nuclei were labeled with DAPI (5 mg/ml in PBS, Sigma, St. Louis, Mo) for 10 min at room temperature in the dark. Finally, the slides were fixed, and images were acquired using a Zeiss LSM 510 laser confocal fluorescence microscope (Carl Zeiss, Oberkochen, Germany).

Data Analysis

All data were evaluated by Graphpad Prism 5.0 software (GraphPad Software, Inc., San Diego, CA, United States). Unpaired Student's t-tests were used to compare parameters between two groups. Multiple comparisons were analyzed by one-way analysis of variance (ANOVA) with Bonferroni post hoc tests. $p < 0.05$ was considered to be statistically significant.

RESULTS

Formononetin Significantly Improved Indicators of Renal Function in db/db Mice and Reduced Renal Fibrosis

Compared with the control group, the levels of fasting blood glucose (FBG) were significantly higher in the model group (**Figure 1A**) ($p < 0.001$). However, FMN treatment caused a significant reduction in FBG levels ($p < 0.01$); there was no statistically significant difference between the groups treated with FMN and Met (the positive drug control group). Compared with the model group, FMN significantly reduced the levels of TG, TC, Cr, and BUN (**Figures**

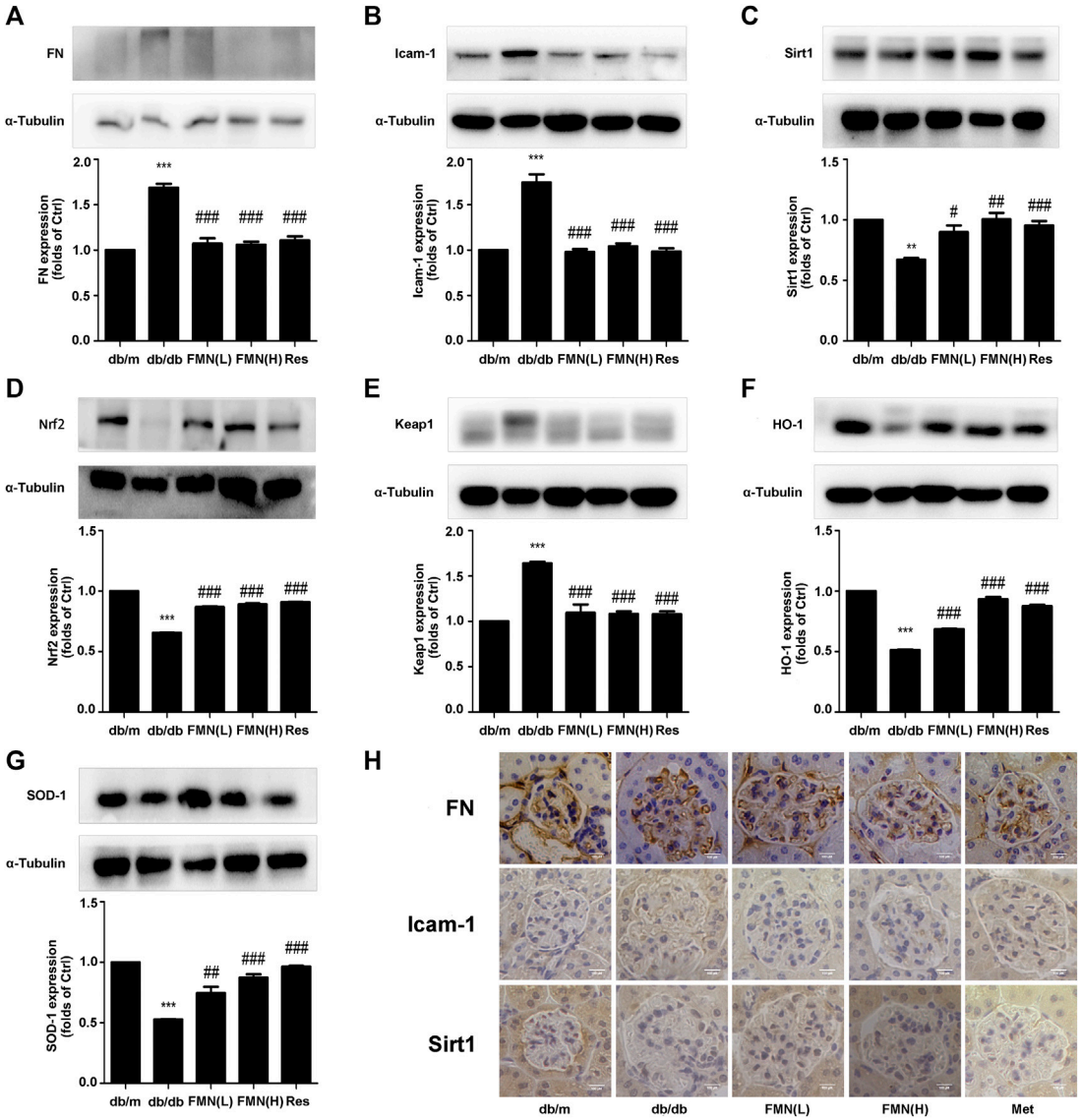

FIGURE 2 | FMN increased the expression of Sirt1 and Nrf2, and reduced the expression of FN and ICAM-1 in the kidneys of db/db mice. **(A–G)** Protein levels of FN, ICAM-1, Sirt1, Nrf2, Keap1, HO-1, and SOD-1, in renal tissues from diabetic mice, as detected by western blotting. $^{**}p < 0.01$, $^{***}p < 0.001$ vs. db/m, $\#p < 0.05$, $\#\#p < 0.01$, $\#\#\#p < 0.001$ vs. db/db. **(H)** Immunohistochemical staining of FN, Icam-1, and Sirt1 in glomeruli (400×). Independent experiments were performed at least three times (with similar results).

1B–E) ($p < 0.01$). As shown in **Figure 1F**, H&E, PAS, and Masson's staining showed that the glomerular structure of mice in the control group remained intact, and no proliferation or glycogen accumulation was observed in the glomerular mesangium. However, in the model group, we observed abnormal hypertrophy of the glomeruli, proliferation of the mesangial cells, an increase in the mesangial matrix, and interstitial fibrosis. Following treatment with FMN, the morphology of the glomeruli gradually returned to normal; the degree of inflammation and fibrosis within the glomeruli also improved (by variable extents). Therefore, these results demonstrated that FMN treatment can significantly improve the pathology of the kidneys in db/db mice.

Formononetin Increased the Expression of Sirt1 and Nrf2, and Reduced the Expression of Fibronectin and Intercellular Adhesion Molecule 1-1 in the Kidneys of db/db Mice

In order to further explore the potential mechanisms that might underlie the ability of FMN to resist oxidative stress and alleviate diabetic kidney fibrosis, we tested a range of associated proteins. As shown in **Figures 2A–G**, the protein levels of FN, ICAM-1, and Keap1, decreased significantly in mice treated with FMN ($p < 0.001$), while the protein levels of Nrf2, HO-1, SOD-1, and Sirt1 increased ($p < 0.05$). Immunohistochemical results (**Figure 2H**) also confirmed that FMN treatment improved Sirt1 expression and inhibited the high expression levels of FN and ICAM-1 in db/db

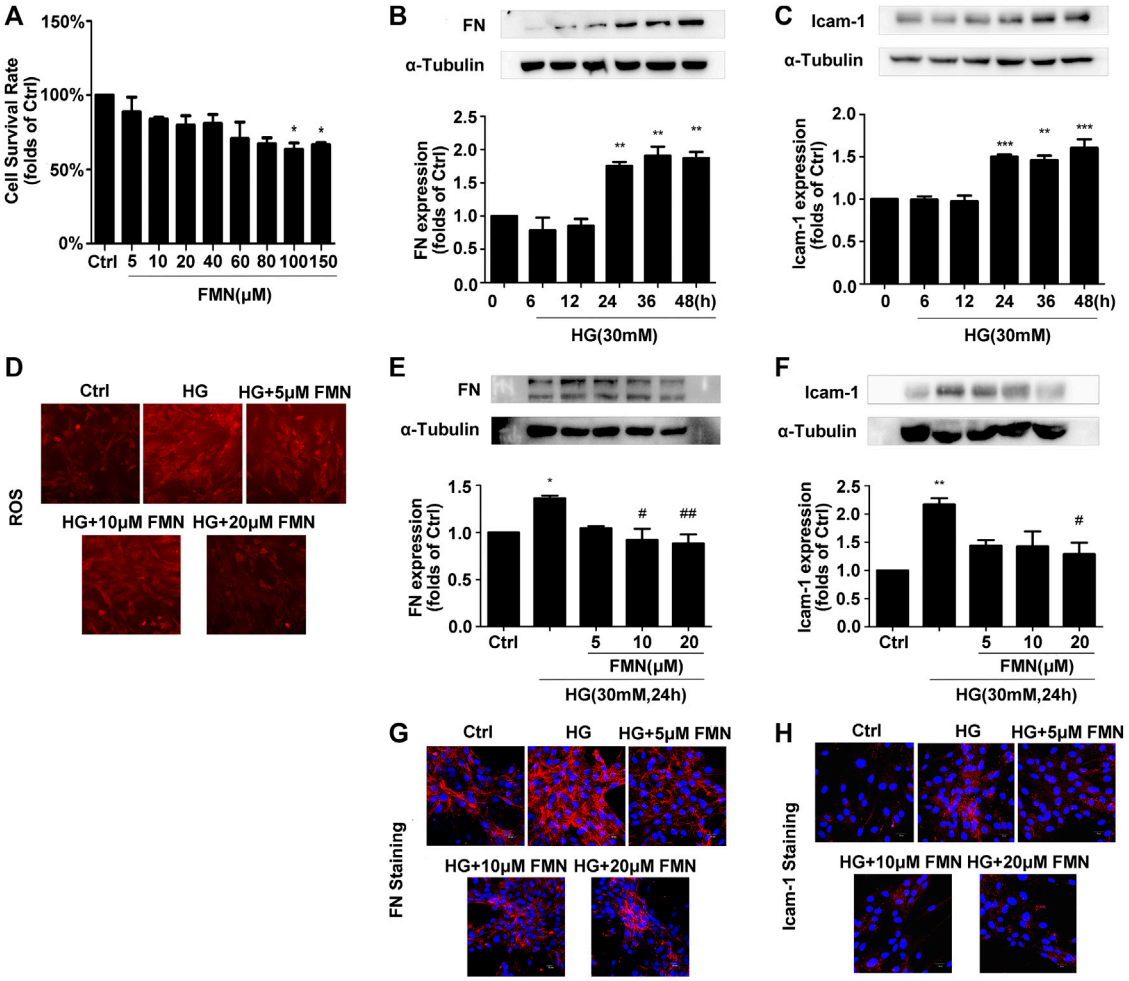

FIGURE 3 | FMN inhibited FN and ICAM-1 protein levels and reduced levels of ROS in GMCs induced by high glucose. **(A)** The viability of GMCs, as detected by MTT assay, $^{*}p < 0.05$ vs. control. **(B,C)** The expression levels of FN and ICAM-1 in GMCs, as detected by western blotting, $^{**}p < 0.01$, $^{***}p < 0.001$ vs. 0 h. **(D)** FMN inhibited the high ROS levels induced by high glucose conditions. **(E,F)** The protein levels of FN and ICAM-1 in GMCs, as determined by western blotting, $^{*}p < 0.05$, $^{**}p < 0.01$ vs. Ctrl, $\#p < 0.05$, $\#\#p < 0.01$ vs. HG. **(G,H)** The distribution of FN and ICAM-1, as determined by immunofluorescence (magnification 400×). Red fluorescence indicates the localization of FN and ICAM-1. Independent experiments are performed at least three times with similar results.

mice. These results suggest that the mechanism by which FMN improves renal fibrosis in diabetic mice may be related to the up-regulation of Sirt1 and the activation of the Nrf2/ARE signaling pathway. To further confirm this hypothesis, we next conducted a series of *in vitro* experiments.

Formononetin Inhibited Fibronectin and Intercellular Adhesion Molecule-1 Protein Levels and Reduced Levels of Reactive Oxygen Species in Glomerular Mesangial Cells Induced by High Glucose

MTT data showed that when applied at a concentration of 100 μM, FMN had no cytotoxic effects on GMCs (**Figure 3A**) ($p < 0.05$). The results of a preliminary dose test showed that FMN could inhibit the expression of FN and ICAM-1 in GMCs stimulated by high glucose at a concentration of 5 μM (**Figures 3E,F**) ($p < 0.05$). Therefore, 5, 10 and 20 μM of FMN were used in subsequent studies. The expression levels of FN and ICAM-1 in GMCs increased in a time-dependent manner following stimulation with high glucose at specified times (0, 6, 12, 24, 36, and 48 h) ($p < 0.01$). Compared with the control group, the upregulation of these indicators was triggered by 30 mM hyper-glycemia for 24 h (**Figures 3B,C**) ($p < 0.05$), as shown by western blotting. Therefore, stimulation with high glucose levels (30 mM) for 24 h was used as a modelling condition for all subsequent experiments. In addition, western blotting and immunofluorescence experiments further showed that FMN treatment reduced the up-regulated expression of FN and ICAM-1 induced by high glucose stimulation for 24 h in a dose-dependent manner (**Figures 3E–H**) ($p < 0.05$). Furthermore, the levels of ROS in GMCs stimulated by high glucose conditions were significantly higher than those in the control group, while

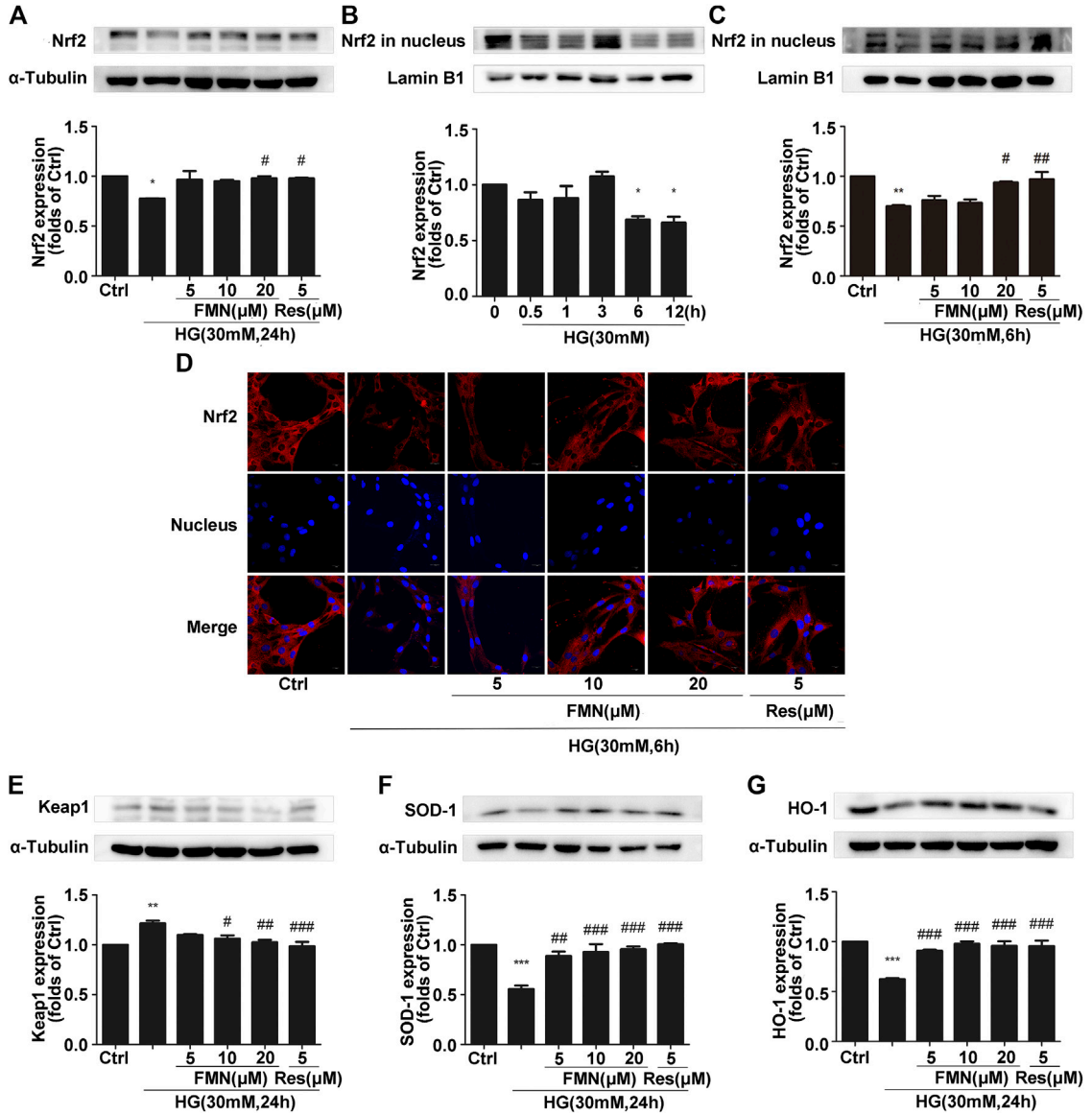

FIGURE 4 | FMN increased the nuclear translocation of Nrf2 to activate the Nrf2/ARE signaling pathway. **(A)** The protein levels of Nrf2 in GMCs, as determined by western blotting, *$p < 0.05$ vs. control, #$p < 0.05$ vs. HG. **(B)** The expression of Nrf2 in GMCs, as determined by western blotting. *$p < 0.05$ vs. 0 h. **(C)** The protein levels of Nrf2 in GMCs, as determined by western blotting, **$p < 0.01$ vs. control (Ctrl), #$p < 0.05$, ##$p < 0.01$ vs. HG. **(D)** The distribution of Nrf2 in GMCs, as determined by immunofluorescence analysis. Scale bar: 20 μm. Red fluorescence indicates the localization of Nrf2. **(E–G)** The protein levels of Keap1, SOD-1, and HO-1, in GMCs, as determined by western blotting, **$p < 0.01$, ***$p < 0.001$ vs. control, ##$p < 0.01$, ###$p < 0.01$ vs. HG. Independent experiments were performed at least three times (with similar results).

levels of ROS in the FMN-treated group were significantly reduced (**Figure 3D**).

Formononetin Increased the Nuclear Translocation of Nrf2 to Activate the Nrf2/ARE Signaling Pathway

In order to verify the correlation between the activation of the Nrf2/ARE signaling pathway by FMN and the enhancement of Sirt1 expression by FMN, we used resveratrol, a classic agonist of Sirt1, as a positive control. As shown in **Figure 4A**, the total Nrf2 protein content in GMCs from the FMN treatment group was significantly increased compared with that of the control group ($p < 0.05$). In order to explore the activation effect of FMN on the nuclear expression of Nrf2, we conducted aging experiments on Nrf2 nuclear protein; data showed that 30 mM high glucose stimulation for 6 h represented the optimal conditions for subsequent experiments (**Figure 4B**) ($p < 0.05$). Western blotting showed that FMN treatment increased the nuclear expression of Nrf2 ($p < 0.05$). Immunofluorescence experiments also showed that Nrf2 was distributed inside and outside of the signaling nucleus in GMCs under physiological conditions, but

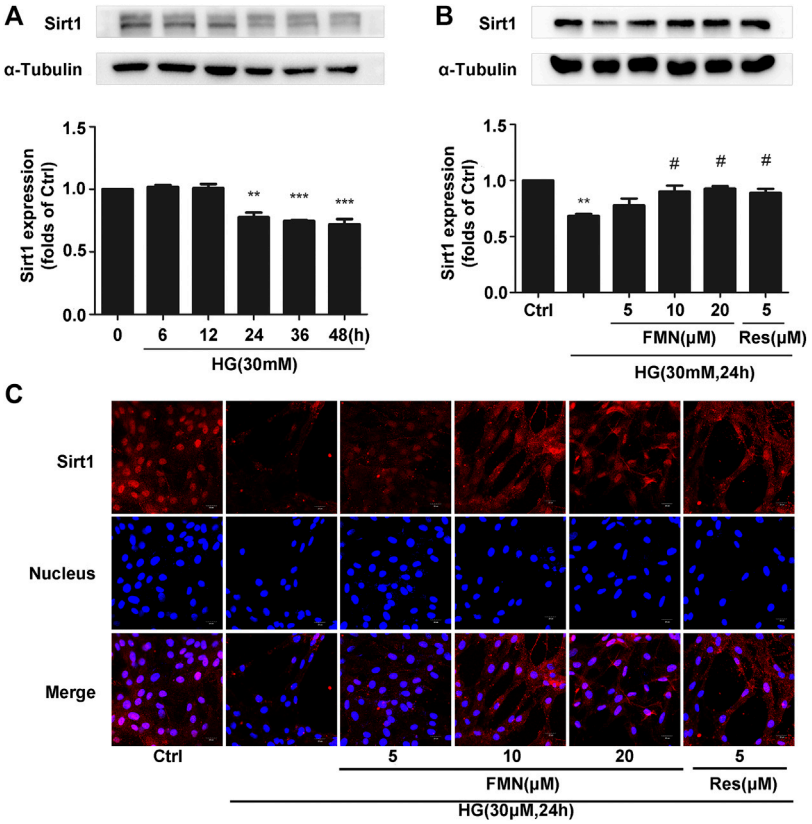

FIGURE 5 | FMN increased the expression of Sirt1 in GMCs induced by high glucose. **(A)** The protein levels of Sirt1 in GMCs, as determined by western blotting, $^{**}p < 0.01$, $^{***}p < 0.001$ vs. 0 h. **(B)** The protein levels of Sirt1 in GMCs, as determined by western blotting, $^{**}p < 0.01$ vs. control, $\#p < 0.05$ vs. HG. **(C)** Distribution of Sirt1 in GMCs, as determined by immunofluorescence. Scale bar: 20 μm. Red fluorescence indicates the localization of Sirt1. Independent experiments were performed at least three times (with similar results).

levels of Nrf2 were reduced in the nucleus under high glucose stimulation. Following FMN treatment, the expression of Nrf2 protein in the nucleus increased (**Figures 4C,D**) ($p < 0.05$). The expression of the Nrf2 target genes, HO-1 and SOD-1, were also increased following the administration of FMN (**Figures 4F,G**) ($p < 0.001$). Keap1 has the ability to regulate the stability of Nrf2 protein and inhibit its expression, so the inhibitory effect of formononetin on Keap1 is also worthy of attention (Vittorio et al., 2010). Meanwhile, in this experiment, the increased expression of Keap1 induced by high glucose stimulation was further reduced in the FMN-treated group (**Figure 4E**) ($p < 0.05$).

Formononetin Increased the Expression of Sirt1 in Glomerular Mesangial Cells Induced by High Glucose

Collectively, the previous experiments demonstrated that FMN was clearly able to activate the Nrf2/ARE signaling pathway. Furthermore, our animal experiments showed that FMN was also activated the expression of Sirt1. In order to verify this more accurately, we conducted aging experiments and finally determined the appropriate conditions to use in subsequent experiments; we ascertained that we should apply 30 mM high-glucose stimulation for a period of 24 h (**Figure 5A**) ($p < 0.01$). As shown in **Figure 5B**, FMN alleviated the inhibitory effect of Sirt1 caused by high-glucose stimulation ($p < 0.05$); this was consistent with data derived from animal experiments. This was further supported by our immunofluorescence data (**Figure 5C**).

The Depletion of Sirt1 Inhibited the Activation of the Nrf2/ARE Signaling Pathway Induced by Formononetin

According to the previous experiments, we ascertained that FMN plays an important role in activating the Nrf2/ARE signaling pathway and improving the process of diabetic renal fibrosis. In order to further clarify the significance of Sirt1 in this process, we used siRNA constructs to interfere with the expression of Sirt1 (**Figure 6A**) ($p < 0.001$). Data showed that the total Nrf2 protein content was significantly reduced when FMN and Sirt1-siRNA were administered at the same time (**Figure 6B**) ($p < 0.001$). Immunofluorescence and western blotting data also showed that the distribution of Nrf2 protein in the nucleus also decreased following the administration of FMN and Sirt1-siRNA (**Figures 6C,D**) ($p < 0.001$); the expression of Keap1 increased, and the levels of the Nrf2 target proteins, HO-1 and SOD-1, were also reduced to different degrees (**Figures 6E–G**) ($p < 0.05$).

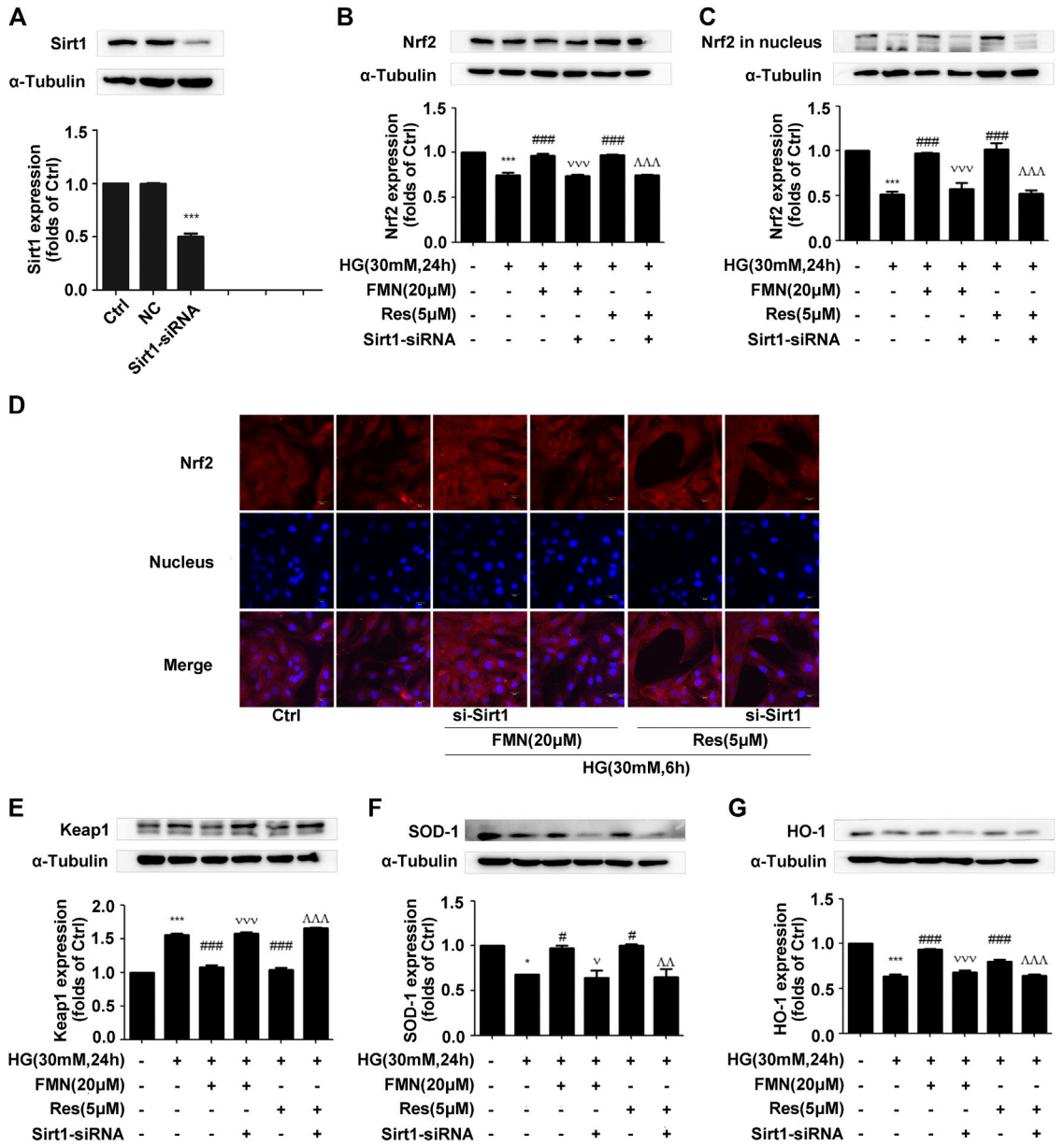

FIGURE 6 | The depletion of Sirt1 inhibited the activation of the Nrf2/ARE signaling pathway induced by FMN. **(A)** The protein levels of Sirt1 in GMCs, as determined by western blotting, ***$p < 0.001$ vs. control (Ctrl). **(B,C)** The protein levels of Nrf2 in GMCs, as determined by western blotting, ***$p < 0.001$ vs. control, ###$p < 0.001$ vs. HG, vvv $p < 0.001$ vs. HG with FMN, ^^^$p < 0.001$ vs. HG with Res. **(D)** The distribution of Nrf2 in GMCs, as determined by immunofluorescence. Scale bar: 20 μm. Red fluorescence indicates the localization of Nrf2. **(E–G)** The protein levels of Nrf2 in GMCs, as determined by western blotting, *$p < 0.05$, ***$p < 0.001$ vs. control (Ctrl), #$p < 0.05$, ###$p < 0.001$ vs. HG, v $p < 0.05$, vvv $p < 0.001$ vs. HG with FMN, ^$p < 0.01$, ^^$p < 0.001$ vs. HG with Res. Independent experiments were performed at least three times (with similar results).

The Depletion of Sirt1 Reversed the Effects of Formononetin on the Inhibition of Oxidative Stress

Activation of the Nrf2/ARE pathway by FMN was inhibited following interference of the Sirt1 protein. In order to verify whether activation of the Nrf2/ARE pathway was also inhibited, we investigated the protein levels of FN and ICAM-1. Western blotting and immunofluorescence experiments showed that Sirt1 interference alleviated the reduced expression levels of FN and ICAM-1 induced by FMN (**Figures 7A–D**) ($p < 0.001$). Sirt1 interference also prevented FMN from lowering the intracellular levels of ROS (**Figure 7E**).

DISCUSSION

Studies have shown that isoflavones could improve the oxidative stress effect by inhibiting the heat shock signaling pathway (Vittorio et al., 2010). FMN as an isoflavone compound also

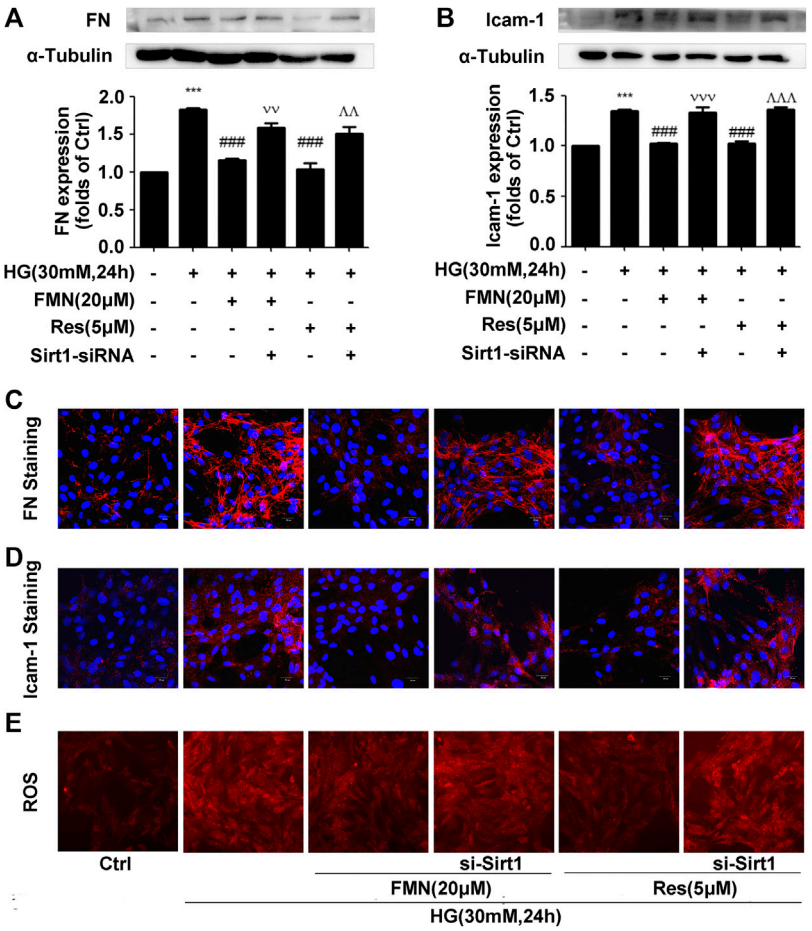

FIGURE 7 | The depletion of Sirt1 reversed the effects of FMN on the inhibition of oxidative stress. **(A,B)** The protein levels of FN and ICAM-1 in GMCs, as determined by western blotting, ***$p < 0.001$ vs. control (Ctrl), ###$p < 0.001$ vs. HG, $p < 0.001$ vs. HG with FMN, ^^$p < 0.001$ vs. HG with Res. **(C,D)** The distribution of FN and ICAM-1 in GMCs, as determined by immunofluorescence. Scale bar: 20 μm. Red fluorescence indicates the localization of FN and ICAM-1. **(E)** The depletion of Sirt1 further enhanced ROS levels induced by HG. Independent experiments were performed at least three times (with similar results).

has been widely reported to improve the renal function of patients suffering from diabetes, and enhance resistance to oxidative stress (Kim et al., 2018; Zhen et al., 2019). In addition, several previous studies have reported that FMN could exert a range of pharmacological activities, including anti-atherosclerosis, anti-tumor, and selective neuroprotective effects (El-Bakoush and Olajide, 2018; Ong et al., 2019; Ma et al., 2020). Oxidative stress is a key link in the intensification of renal fibrosis and can also accelerate the progression of diabetic nephropathy. Given that FMN has been reported to exert significant antioxidant capacity, we hypothesized that FMN could also improve diabetic nephropathy via its antioxidant capabilities probably.

In this study, we used db/db mice as a model for type 2 diabetic nephropathy as experimental animals. We found that FMN significantly reduced the levels of BUN, Cr, TG, and TC in the model group, thus indicating that FMN could clearly improve the renal function of diabetic mice. Furthermore, HE, PAS, and Masson's staining showed that after FMN treatment, there was clear improvements in renal tubular interstitial fibrous tissue hyperplasia and glycogen accumulation compared with the model group. Endogenous cellular defense mechanism plays an important role in resisting oxidative stress and maintaining cellular state. The Nrf2/ARE signaling pathway is one of the most well-studied oxidative stress pathways. Studies have shown that Nrf2 could maintain protein homeostasis under various pathological conditions and improve oxidative stress by encoding the antioxidant pathway of Vitagenes (Calabrese et al., 1996; Calabrese et al., 2010; Calabrese and Calabrese, 2013). It is generally believed that when Nrf2 is stimulated by oxidation, it is separated from endogenous inhibitor (Keap1) and then Nrf2 translocates into the nucleus to mediate the transcription of its target genes, including HO-1 and SOD-1 (Gong et al., 2018). In order to study the antioxidant effects of FMN, we carried out western blotting and immunofluorescence experiments using both mouse kidney tissue and GMCs that had been induced by high glucose. Data showed that FMN treatment could activate the Nrf2/ARE signaling pathway and significantly enhance the nuclear translocation of Nrf2. FMN also alleviated renal fibrosis

by reducing the expression levels of FN and ICAM1both *in vitro* and *in vivo*. Some studies have shown that Sirt1 is closely associated with the pathogenesis of diabetic nephropathy and represents a new potential therapeutic target for diabetic nephropathy (Munehiro et al., 2013; Sun et al., 2018). Sirt1 agonists are known to have protective effects on the metabolism of some diabetic animals, including glucose tolerance, fasting glucose level, and insulin resistance (Marina et al., 2018). Sirt1 knockout mice showed increased levels of proteinuria and renal impairment (Hasegawa et al., 2013). Furthermore, *in vitro* and *in vivo* studies have shown that Res, an agonist of Sirt1, could reduce ROS production and up-regulate the expression of endogenous antioxidant genes and vascular protective molecules (Vella et al., 2015; Yang et al., 2018). In this study, we found that FMN promotes the expression of Sirt1 both *in vitro* and *in vivo*. In addition, we knocked out Sirt1 in GMCs induced by high glucose, and found that the antioxidant and anti-fibrosis effects induced by FMN were negated. Although we have demonstrated that FMN can prevent the deposition of ECM and improve oxidative stress by regulating the Sirt1/Nrf2/ARE pathway, how does FMN directly activate Sirt1 and its association with Vitagenes network have not been fully clarified. This issue needs further investigation. Collectively, these results suggest that Sirt1 is necessary for FMN to activate the Nrf2/ARE signaling pathway.

CONCLUSION

In vitro and *in vivo* experiments confirmed that the inhibitory mechanism underlying the effects of FMN on diabetic renal fibrosis was closely related to the up-regulation of Sirt1 expression and activation of the Nrf2/ARE signaling pathway. Our results provide an experimental basis for the application of FMN in the prevention and treatment of DN and the development of new drugs.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by Animal Ethics Committee of Guangzhou University of Chinese Medicine. Written informed consent was obtained from the owners for the participation of their animals in this study.

AUTHOR CONTRIBUTIONS

KZ and HH designed and performed the experiments, acquired and analyzed the data, and drafted the manuscript. XJ and CY helped to perform the animal experiments. YW, RL, and YMW helped to draft the manuscript. SQ contributed reagents, materials, and analysis tools. All authors read and approved the final manuscript.

FUNDING

This research was supported by the 2019 Guangdong University Characteristic Innovation Project (Reference number: 2019KTSCX023).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2020.616378/full#supplementary-material.

REFERENCES

Ahmed, S. M., Luo, L., Namani, A., Wang, X. J., and Tang, X. (2017). Nrf2 signaling pathway: pivotal roles in inflammation. *Biochim. Biophys. Acta* 1863, 585. doi:10.1016/j.bbadis.2016.11.005

Arioz, B. I., Tastan, B., Tarakcioglu, E., Tufekci, K. U., Olcum, M., Ersoy, N., et al. (2019). Melatonin attenuates LPS-induced acute depressive-like behaviors and microglial NLRP3 inflammasome activation through the SIRT1/Nrf2 pathway. *Front. Immunol.* 10, 1511. doi:10.3389/fimmu.2019.01511

Calabrese, E. J., and Calabrese, I. I. (2013). Hormesis: its impact on medicine and health. *Hum. Exp. Toxicol.* 32. doi:10.1177/0960327112455069

Calabrese, V., Renis, M., Calderone, A., Russo, A., Barcellona, M. L., and Rizza, V. (1996). Stress proteins and SH-groups in oxidant-induced cell damage after acute ethanol administration in rat. *Free Radic. Biol. Med.* 20, 391. doi:10.1016/0891-5849(95)02095-0

Calabrese, V., Cornelius, C., Maiolino, L., Luca, M., Chiaramonte, R., Toscano, M. A., et al. (2010). Oxidative stress, redox homeostasis and cellular stress response in Ménière's disease: role of vitagenes. *Neurochem. Res.* 35, 2208. doi:10.1007/s11064-010-0304-2

El-Bakoush, A., and Olajide, O. A. (2018). Formononetin inhibits neuroinflammation and increases estrogen receptor beta (ERβ) protein expression in BV2 microglia. *Int. Immunopharm.* 61, 325. doi:10.1016/j.intimp.2018.06.016

Gong, W., Chen, Z., Zou, Y., Zhang, L., Huang, J., Liu, P., et al. (2018). CKIP-1 affects the polyubiquitination of Nrf2 and Keap1 via mediating Smurf1 to resist HG-induced renal fibrosis in GMCs and diabetic mice kidneys. *Free Radic. Biol. Med.* 115, 338. doi:10.1016/j.freeradbiomed.2017.12.013

Hasegawa, K., Wakino, S., Simic, P., Sakamaki, Y., Minakuchi, H., Fujimura, K., et al. (2013). Renal tubular Sirt1 attenuates diabetic albuminuria by epigenetically suppressing Claudin-1 overexpression in podocytes. *Nat. Med.* 19, 1496. doi:10.1038/nm.3363

Hu, W., Wang, J., Guo, W., Liu, Y., Guo, Z., Miao, Y., et al. (2020). Studies on characteristics and anti-diabetic and -nephritic effects of polysaccharides isolated from Paecilomyces hepiali fermentation mycelium in db/db mice. *Carbohydr. Polym.* 232, 115766. doi:10.1016/j.carbpol.2019.115766

Jin, F., Wan, C., Li, W., Yao, L., Zhao, H., Zou, Y., et al. (2017). Formononetin protects against acetaminophen-induced hepatotoxicity through enhanced NRF2 activity. *PLoS One* 12. doi:10.1371/journal.pone.0170900

Kim, E. N., Lim, J. H., Kim, M. Y., Ban, T. H., Jang, I. A., Yoon, H. E., et al. (2018). Resveratrol, an Nrf2 activator, ameliorates aging-related progressive renal injury. *Aging* 10. doi:10.18632/aging.101361

Kim, C., Lee, S. G., Yang, W. M., Arfuso, F., Um, J. Y., and Kumar, A. P. (2018). Formononetin-induced oxidative stress abrogates the activation of STAT3/5 signaling axis and suppresses the tumor growth in multiple myeloma preclinical model. *Canc. Lett.* 431, 123. doi:10.1016/j.canlet.2018.05.038

Kultigin, T. (2017). Inflammation, oxidative stress, apoptosis, and autophagy in diabetes mellitus and diabetic kidney disease: the four horsemen of the apocalypse. *Int. Urol. Nephrol.* 49, 837. doi:10.1007/s11255-016-1488-4

Ma, R., Liang, W., Sun, Q., Qiu, X., Ying, L., Ge, X., et al. (2018). Sirt1/Nrf2 pathway is involved in oocyte aging by regulating Cyclin B1. *Aging* 10. doi:10.18632/aging.101609

Ma, F., Wu, J., Jiang, Z., Huang, W., Jia, Y., Sun, W., et al. (2019). P53/NRF2 mediates SIRT1's protective effect on diabetic Nephropathy. *Biochim. Biophys. Acta Mol. Cell Res.* 1866, 1272. doi:10.1016/j.bbamcr.2019.04.006

Ma, C., Xia, R., Yang, S., Liu, L., Zhang, J., Feng, K., et al. (2020). Formononetin attenuates atherosclerosis via regulating interaction between KLF4 and SRA in apoE-/- mice. *Theranostics*. 10. doi:10.7150/thno.38115

Munehiro, K., Shinji, K., Ai, T.-W., Keizo, K., and Daisuke, K. (2013). Sirtuins and renal diseases: relationship with aging and diabetic Nephropathy. *Clin. sci.* 124. doi:10.1042/CS20120190

Nwakiban, A. P. A., Cicolari, S., Piazza, S., Gelmini, F., Sangiovanni, E., Martinelli, G., et al. (2020). Oxidative stress modulation by Cameroonian spice extracts in HepG2 cells: involvement of Nrf2 and improvement of glucose uptake. *Metabolites* 10. doi:10.3390/metabo10050182

Ong, S. K. L., Shanmugam, M. K., Fan, L., Fraser, S. E., Arfuso, F., Ahn, K. S., et al. (2019). Focus on formononetin: anticancer potential and molecular targets. *Cancers* 11. doi:10.3390/cancers11050611

Perkovic, V., Jardine, M. J., Neal, B., Bompoint, S., Heerspink, H. J. L., Charytan, D. M., et al. (2019). Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. *N. Engl. J. Med.* 380. doi:10.1056/NEJMoa1811744

Qiu, G., Tian, W., Huan, M., Chen, J., and Fu, H. (2017). Formononetin exhibits anti-hyperglycemic activity in alloxan-induced type 1 diabetic mice. *Exp. Biol. Med.* 242, 223. doi:10.1177/1535370216657445

Ruiz-Ortega, M., Rayego-Mateos, S., Lamas, S., Ortiz, A., and Rodrigues-Diez, R. R. (2020). Targeting the progression of chronic kidney disease. *Nat. Rev. Nephrol.* 16, 269. doi:10.1038/s41581-019-0248-y

Sagoo, M. K., and Gnudi, L. (2018). Diabetic nephropathy: is there a role for oxidative stress? *Free Radic. Biol. Med.* 116, 50. doi:10.1016/j.freeradbiomed.2017.12.040

Sun, Z., Ma, Y., Chen, F., Wang, S., Chen, B., Shi, J., et al. (2018). miR-133b and miR-199b knockdown attenuate TGF-β1-induced epithelial to mesenchymal transition and renal fibrosis by targeting SIRT1 in diabetic nephropathy. *Eur. J. Pharmacol.* doi:10.1016/j.ejphar.2018.08.022

Vella, K. R., Pullen, C., Coulson, F. R., and Fenning, A. S. (2015). Resveratrol prevents cardiovascular complications in the SHR/STZ rat by reductions in oxidative stress and inflammation. *BioMed. Res. Int.* doi:10.1155/2015/918123

Vittorio, C., Cornelius, C., Kostova, A. T. D., Calabrese, E. J., and Mattson, M. P. (2010). Cellular stress responses, the hormesis paradigm, and vitagenes: novel targets for therapeutic intervention in neurodegenerative disorders. *Antioxid. Redox Signal.* 13. doi:10.1089/ars.2009.3074

Wei-Jian, N., Li-Qin, T., and Wei, W. (2015). Research progress in signalling pathway in diabetic nephropathy. *Diabetes/metabolism research and reviews* 31. doi:10.1002/dmrr.2568

Wu, M., Luca, P., and Ariela, B. (2018). Sirtuins in renal health and disease. *J. Am. Soc. Nephrol.* 29. doi:10.1681/ASN.2017111218

Xiao, H., Sun, X., Liu, R., Chen, Z., Lin, Z., Yang, Y., et al. (2020). Gentiopicroside activates the bile acid receptor Gpbar1 (TGR5) to repress NF-kappaB pathway and ameliorate diabetic nephropathy. *Pharmacol. Res.* 151, 104559. doi:10.1016/j.phrs.2019.104559

Yang, S., Wei, L., Xia, R., Liu, L., Chen, Y., Zhang, W., et al. (2019). Formononetin ameliorates cholestasis by regulating hepatic SIRT1 and PPARα. *Biochem. Biophys. Res. Commun.* 512. doi:10.1016/j.bbrc.2019.03.131

Yang, G., Chang, C. C., Yang, Y., Yuan, L., Xu, L., Ho, C. H., et al. (2018). Resveratrol alleviates rheumatoid arthritis via reducing ROS and inflammation, inhibiting MAPK signaling pathways, and suppressing angiogenesis. *J. Agric. Food Chem.* doi:10.1021/acs.jafc.8b05047

Zhen, Z., Xinjian, Z., Youhong, D., Mingyi, L., and Yancheng, X. (2019). Formononetin ameliorates high glucose-induced endothelial dysfunction by inhibiting the JAK/STAT signaling pathway. *Mol. Med. Rep.* 20. doi:10.3892/mmr.2019.10512

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 February 2021
doi: 10.3389/fpls.2021.632557

Iris domestica (iso)flavone 7- and 3′-O-Glycosyltransferases Can Be Induced by $CuCl_2$

Xiang Zhang[1,2†], *Yan Zhu*[1,2†], *Jun Ye*[1,2], *Ziyu Ye*[1,2], *Ruirui Zhu*[1,2], *Guoyong Xie*[1,2], *Yucheng Zhao*[1,2*] *and Minjian Qin*[1,2*]

[1] *Department of Resources Science of Traditional Chinese Medicines, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing, China,* [2] *Key Laboratory of Modern Traditional Chinese Medicines (Ministry of Education), China Pharmaceutical University, Nanjing, China*

In many plants, isoflavones are the main secondary metabolites that have various pharmacological activities, but the low water solubility of aglycones limits their usage. The O-glycosylation of (iso)flavones is a promising way to overcome this barrier. O-glycosyltransferases (UGTs) are key enzymes in the biosynthesis of (iso)flavonoid O-glycosides in plants. However, limited investigations on isoflavonoid O-UGTs have been reported, and they mainly focused on legumes. *Iris domestica (L.)* Goldblatt et Mabberley is a non-legume plant rich in various isoflavonoid glycosides. However, there are no reports regarding its glycosylation mechanism, despite the *I. domestica* transcriptome previously being annotated as having non-active isoflavone 7-O-UGTs. Our previous experiments indicated that isoflavonoid glycosides were induced by $CuCl_2$ in *I. domestica* calli; therefore, we hypothesized that isoflavone O-UGTs may be induced by Cu^{2+}. Thus, a comparative transcriptome analysis was performed using *I. domestica* seedlings treated with $CuCl_2$, and eight new active BcUGTs were obtained. Biochemical analyses showed that most of the active BcUGTs had broad substrate spectra; however, substrates lacking 5-OH were rarely catalyzed. Real-time quantitative PCR results further indicated that the transcriptional levels of BcUGTs were remarkably induced by Cu^{2+}. Our study increases the understanding of UGTs and isoflavone biosynthesis in non-legume plants.

Keywords: ***Iris domestica (Belamcanda chinensis)*** **copper chloride, induce, (iso)flavones, transcriptome, UGT**

OPEN ACCESS

Edited by:
María Pilar López-Gresa, Universitat Politècnica de València, Spain

Reviewed by:
Alexander Wilson, Northern Michigan University, United States
Bing-kai Hou, Shandong University, China

***Correspondence:**
Minjian Qin
minjianqin@163.com
Yucheng Zhao
zhaoyucheng1986@126.com

†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: *23 November 2020*
Accepted: *18 January 2021*
Published: *09 February 2021*

Citation:
Zhang X, Zhu Y, Ye J, Ye Z, Zhu R, Xie G, Zhao Y and Qin M (2021) Iris domestica (iso)flavone 7- and 3′-O-Glycosyltransferases Can Be Induced by $CuCl_2$. Front. Plant Sci. 12:632557. doi: 10.3389/fpls.2021.632557

INTRODUCTION

Secondary metabolites derived from plants play very important roles in organisms' physiological activities, as well as human health. Isoflavonoid glycosides are a class of significant secondary metabolites mainly found in legume plants, which are associated with the interactions between legumes and both symbiotic and pathogenic microorganisms (Krämer et al., 1984; Graham et al., 1990; Udomsuk et al., 2010; Li et al., 2014; Clúa et al., 2018; Karre et al., 2019). However, the low water solubility of isoflavonoid aglycones is a barrier to the clinical applications. Glycosylation is an effective way to increase water solubility and bioavailability. For example, the apparent solubility of puerarin 7-O-glycoside is 18-fold that of puerarin (Jiang et al., 2008), and genistein 7-O-glycoside showed greater oral bioavailability than genistein (Kwon et al., 2007). O-glycosylation is the main glycosylation form, and O- UDP- sugar glycosyltransferases (O-UGTs) are responsible for the

biosynthesis of various O-glycosides in plants. However, only several genes with isoflavone O-glycosylation ability have been reported, and moreover, they are mainly investigated in leguminous plants. For instance, eight *Medicago* UGTs have been identified as being active against various isoflavones and flavones using a functional genomics approach (Modolo et al., 2007). In soybean (*Glycine max*), GmUGT4 is highly specific for isoflavones, while GmUGT1 and GmUGT7 have broad substrate spectra (Funaki et al., 2015). *Pueraria lobate* UGT1 plays a role in isoflavone 7-O-glycosylation (Li et al., 2014), while PlUGT2 is a bifunctional isoflavone O-UGT responsible for both the 4′-O- and 4′,7-O-glycosylation of isoflavonoids (Wang et al., 2016). In non-legumes, isoflavone O-UGTs have only been reported in pomegranate and rice. UGT84A23, UGT84A24, and PgUGT95B2 cloned from pomegranate exhibit 7-O-glycosylation activity of genistein *in vitro* (Ono et al., 2016; Wilson et al., 2019). In the monocotyledonous plant rice, UGT709A4 and UGT706D1 could use isoflavones as substrates *in vitro*, and produce isoflavonoid 7-O-glycosides (Ko et al., 2008).

Iris domestica (L.) Goldblatt et Mabberley [syn. *Belamcanda chinensis* (L.) DC] forms a traditional Chinese medicine and isoflavones are the main active ingredients. Previous *in vitro* investigations showed that the isoflavones in rhizomes have a number of biological activities, including anti-mutagenic, anti-inflammatory, anti-angiogenic and anti-tumor activities, as well as inhibiting lipid peroxidation and scavenging free radicals (Seidlová-Wuttke et al., 2004; Zhang et al., 2005; Kang et al., 2009; Wozniak et al., 2010; Wang et al., 2013; Xie et al., 2014). Among the active isoflavones, O-glycosylated compounds, such as tectoridin and iridin, are the main ingredients (Lee et al., 2011; Chen et al., 2014). This implies that some isoflavone O-UGT genes may be abundant in *I. domestica*. In addition, it is unusual that *I. domestica* accumulates isoflavonoid glycosides as a non-legume. Thus, it is important for us to explore the mechanisms of glycosylation in *I. domestica*. However, *I. domestica* is a monocotyledonous plant; therefore, it is very difficult to obtain the active BcUGTs using homology cloning strategies owing to their low homology levels. Transcriptomes have aided in the discovery of isoflavone-related-UGTs in *I. domestica*. However, after analyzing the transcriptomes of six *I. domestica* organs, we did not identify any active (iso)flavone genes (Tian et al., 2018).

Previously, we found that the isoflavonoid glycoside content in *I. domestica* calli was significantly induced by Cu^{2+} stress. Specifically, when the isoflavones contents of calli were monitored for 49 days, tectoridin and iridin increased up to 1.14 mg/g and 0.95 mg/g, respectively, at 42 days after a 1.2-mM $CuCl_2$ treatment. However, no tectoridin or iridin was detected in the control group (Zhu et al., 2020). The same dynamic content variation was observed in the adventitious roots of *Iris germanica* after a $CuCl_2$ treatment (Tomoyoshi et al., 2005). Thus, $CuCl_2$ appears to promote the biosynthesis of isoflavone glycosides by inducing related gene expression. Consequently, in this study, we chose $CuCl_2$ as the elicitor to induce the expression of (iso)flavone 7-O-UGT genes in *I. domestica*. Using RNA sequencing (RNA-seq) and differentially expressed gene analyses, several BcUGT complementary DNAs (cDNAs), which belonged to the UGT706L family, were obtained. Biochemical analyses showed that eight active BcUGTs shared similar coding DNA sequences, but they exhibited different catalytic activities against different substrates. For most substrates, 7-O-glycoside compounds were the main products. Interestingly, a new compound was catalyzed by BcUGT4 and BcUGT5, and catalysis by the BcUGTs was 5-OH dependent. In addition, real-time quantitative PCR was performed to predict the functions of BcUGTs in *I. domestica* growth, especially its mechanism in $CuCl_2$-induced gene expression. Our study elucidated the mechanisms of isoflavonoid glycoside biosynthesis in a non-legume plant through Cu^{2+} induction. Furthermore, it offers the potential to engineer new compounds using synthetic biological strategies.

Abbreviations: cDNA, complementary DNA; $CuCl_2$, copper chloride; UDPG, uridine diphosphate glucose; UGT, UDP-sugar glycosyltransferase.

MATERIALS AND METHODS

Plant Materials and Stress Treatment

Newly harvested seeds of *I. domestica* were collected from the Botanic Garden of China Pharmaceutical University, Nanjing, China (31°54′36.36 N, 118°54′38.42 E). The seeds were identified by Prof. Minjian Qin (China Pharmaceutical University, Nanjing, China). After being sown in nutrient soil mixed with perlite, the seeds were placed in an artificial climate box at an ambient temperature of 25°C under a 14-h/10-h photoperiod. After 45 days of culturing, all the seedlings were transferred into liquid 1/2 Murashige and Skoog medium (MS, Murashige and Skoog, 1962) for 48 h. Then, 1.5 mM $CuCl_2$ was added to the medium, while nothing was added to the control group medium. The concentration of $CuCl_2$ was established in accordance with to the results of our previous copper stress-induced experiment (Zhu et al., 2020). Roots were collected after a 24-h $CuCl_2$ treatment for RNA-seq and gene expression analyses. Another induction experiment was carried out to rule out the effect of the chlorine ion in the induction, 48-day-old seedlings were transferred into liquid 1/2 MS for 48 h. Then 1.5 mM $MgCl_2$, 1.5 mM $CaCl_2$, 1.5 mM $ZnCl_2$, and 1.5 mM $CuCl_2$ was added to the medium respectively, while nothing was added to the control group medium. Roots were collected after a 24-h treatment for gene expression analyses. All the samples were collected in three independent biological replicates. Tissues from different organs (flower, fruit, leaf, rhizome, and root) were collected separately in three biological replicates from three independent plants, these samples were frozen immediately in liquid nitrogen after washing and maintained at −80°C until use.

RNA Extraction and cDNA Library Preparation

Total RNAs were extracted using an RNA simple total RNA kit (TIANGEN, Beijing, China), and the roots were collected from the seedlings treated with $CuCl_2$ and the control group. RNA degradation and contamination were examined using 1% agarose gels. RNA purity was checked using a NanoPhotometer® spectrophotometer (Implen, Carlsbad, CA, USA). The RNA

concentration was measured using a Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Sequencing libraries were generated using the NEB Next Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Seattle, WA, USA).

Illumina Sequencing, Data Filtering, Transcriptome Assembly, and Gene Functional Annotation

Six sequencing libraries of *I. domestica* were subjected to the Illumina Novaseq 6000 platform (Illumina, San Diego, CA, USA) to generate paired-end reads. After removing low-quality reads and reads containing adapters or poly-N, we obtained clean data that were then assembled into contigs using the Trinity assembler (Grabherr et al., 2011). A gene functional annotation was performed in seven public databases: NCBI non-redundant protein sequences, NCBI non-redundant nucleotide sequences, Pfam, euKaryotic Ortholog Groups/Clusters of Orthologous Groups of proteins, Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes, and Gene Ontology.

Screening of Candidate BcUGTs

Clean data were mapped back onto the assembled transcriptome. The read count for each gene was obtained from the mapping results using the RNA-seq by Expectation–Maximization (RSEM) software package (Li and Dewey, 2011). The expression level of each transcript in each sample was described in terms of fragments per kilobase of exon model per million mapped reads, which is a calculation of normalized read counts. A differential expression analysis of two samples was performed using the DESeq R package (Anders and Huber, 2010). Genes having an adjusted p-value < 0.005 and $|\log2 \text{ Foldchange}| > 1$ were considered as differentially expressed. To search for genes involved in the biosynthesis of isoflavone 7-O-glycosides in *I. domestica*, we established a local database using the transcriptome data, and then performed a local BLAST algorithm-based search to identify genes that shared similar conserved sequences with isoflavone 7-O-UGT genes of legume plants. To further determine the evolutionary relationships among these genes, the BcUGTs proteins were aligned with sequences of other plant (iso)flavone O-UGT genes having known functions using MEGA7 (https://www.megasoftware.net/). A phylogenetic tree was constructed using the neighbor-joining method, with a bootstrap of 1,000. The evolutionary distances were computed using the Poisson correction (Kumar et al., 2016).

Functional Analysis of BcUGTs Candidates *in vitro*

The open reading frames of BcUGTs were cloned individually into the prokaryotic expression vector pCold-TF, which helps the proteins fold correctly owing to the expression of the coupled trigger factor (chaperone). Consequently, the efficient expression of soluble proteins that are difficult to deal with in other systems may be achieved. The primers used for the cloning are listed in **Supplementary Table 1**. All the BcUGT-pCold-TF constructs were transferred into *Escherichia coli* after sequencing. The recombinant strains were then inoculated into 100 mL Luria Broth liquid culture medium at 37°C and 220 rpm in an orbital shaker incubator. When the OD_{600} of the strains reached 0.6, 0.5 mM isopropyl 1-β-D thiogalactoside (IPTG) was added into the culture to induce the expression of recombinant proteins. This was followed by further incubation at 15°C for 24 h. The recombinant *E. coli* cells were harvested by centrifugation (10,000 g for 10 min at 4°C), resuspended with 50 mM phosphate-buffered saline (PBS; pH 8.0) and disrupted using a sonicator. The recombinant proteins were purified using Ni-NTA columns (Sangon, Shanghai, China). The purified proteins were subjected to SDS-PAGE in accordance with the method of Laemmli (1970), and the protein concentrations were determined using the Bradford method (Bradford, 1976).

Enzyme Assays

The catalytic UGT reactions were performed in 200 μL. The reaction components included 2 mM uridine diphosphate glucose (UDPG; Sigma-Aldrich, Shanghai, China), 1 mM dithiothreitol, 50 mM PBS, 10 μg of the purified recombinant protein, and 100 μM (iso)flavonoid acceptors. The reactions were incubated for 1 h at 30°C and then 400 μL ethyl acetate was added to stop the reactions. Ethyl acetate was vaporized to dryness in the vacuum drying oven, and finally, the samples were dissolved in 100 μL methanol for HPLC analysis. Firstly, we calculated the amount of residual substrate, then the ratio of residual substrate in total initial substrate was obtained, and the ratio is the substrate conversion rate. For kinetic studies of BcUGTs, typical assays containing 2 mM UDPG, 1 mM dithiothreitol, 50 mM PBS, 2 μg purified recombinant protein, and varying concentrations of substrates were performed. The reactions were incubated for 10 min at 30°C, and the solutions subjected to HPLC were prepared as above. Kinetic parameters were determined by hyperbolic regression analyses using the Prism program (GraphPad, https://www.graphpad.com/). The reaction products were analyzed using an Agilent 1100 HPLC system on SB C18 columns (5 μm; 250 × 4.6 mm; Agilent Technologies) in which solvent A was 0.1% formic acid in water and solvent B was 100% acetonitrile. The substrates and products were eluted using a linear gradient of 30% B to 60% B in 15 min at a flow rate of 1 mL/min. The chromatograms were obtained with detections at 270 nm (for isoflavones), 280 nm (for flavanones), 340 nm (for flavones), and 360 nm (for flavonols).

Homology Modeling and Auto-Docking Analysis

We chose BcUGT4 to generate a protein model using SWISS-MODEL (https://swissmodel.expasy.org/, Biasini et al., 2014), the templates were searched using BLAST (Camacho et al., 2009) and HHblits (Steinegger et al., 2019). We used GMQE (Global Model Quality Estimation) to estimate the template quality, GMQE score is expressed as a number between 0 and 1, higher numbers indicate higher reliability. The quality of protein model was evaluated *via* GMQE and QMEAN (Benkert et al., 2008).

Docking of naringenin and liquiritigenin to BcUGT4 model was performed using AutoDock Tools (Trott and Olson, 2010, version 1.5.6). We used the default parameters in AutoDock Tools to prepare the ligands and macromolecule, then ran AutoDock to acquire the docking structures. Finally, we used PyMOL to present the docking results (https://pymol.org/2/).

Transcriptional Aanalysis

The RNAs of seedling roots treated with $MgCl_2$, $CaCl_2$, $ZnCl_2$, and $CuCl_2$ were isolated as above. Then, different organs of *I. domestica* and different parts of rhizomes were extracted to detect the expression levels of BcUGTs, and samples were all collected in three independent biological replicates. The quality and the quantity of total RNAs were determined using a NanoDrop Spectrophotometer (Eppendorf, Germany) and 1% agarose gels. The first-strand cDNA for real-time PCR was synthesized using HiScript II Q RT SuperMix for qPCR (Vazyme, Nanjing, China). The PCR reaction was carried out in a total volume of 20 μL containing ChamQ SYBR qPCR Master Mix (Vazyme) on an ABI StepOne plus Real-time PCR system (Applied Biosystem, USA), with cycling parameters 95°C for 30 s and 40 cycles of 95°C for 5 s and 60°C for 30 s. The relative quantification method ($\Delta\Delta^{-CT}$) was used to calculate the expression level of each unigene (Livak and Schmittgen, 2001). The specificity of primers was tested using a melting curve program. The slope of the relative standard curve was calculated to ensure the amplification efficiency. The primers used for qPCR are listed in **Supplementary Table 2**, and each gene has three biological replications and three technical replications.

RESULTS

De novo Sequencing and Functional Annotation of Unigenes

Our previous study showed that calli of *I. domestica* treated with $CuCl_2$ accumulated 2.09 mg/g of tectoridin and iridin, while no isoflavonoid glycosides were detected in the control group. Thus, 45-day-old *I. domestica* seedling roots treated with $CuCl_2$ were used for Illumina sequencing after the RNA quality met specific criteria (**Supplementary Figure 1**). The adapter, unknown and low-quality reads were removed to obtain clean reads. In total, 25.03 and 25.57 Gb clean bases from the control and treated groups, respectively, were obtained and used for transcript assembly (**Supplementary Table 3**). A total of 73,132 unigenes were annotated in the seven referenced public databases, with 24.6% annotated unigenes being homologous to *Asparagus officinalis*, followed by *Elaeis guineensi* and *Phoenix dactylifera*. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes annotations assigned the unigenes to different metabolic pathways. This was the basis for the preliminarily functional predictions for these unigenes (**Supplementary Figures 2, 3**). These results provided a good foundation for the screening of candidate BcUGTs.

Selection of Candidate BcUGTs Induced by $CuCl_2$

Genes involved in the same biosynthetic route are temporally and spatially induced by effective treatments, which indicates that the gene functions may be characterized by a differential

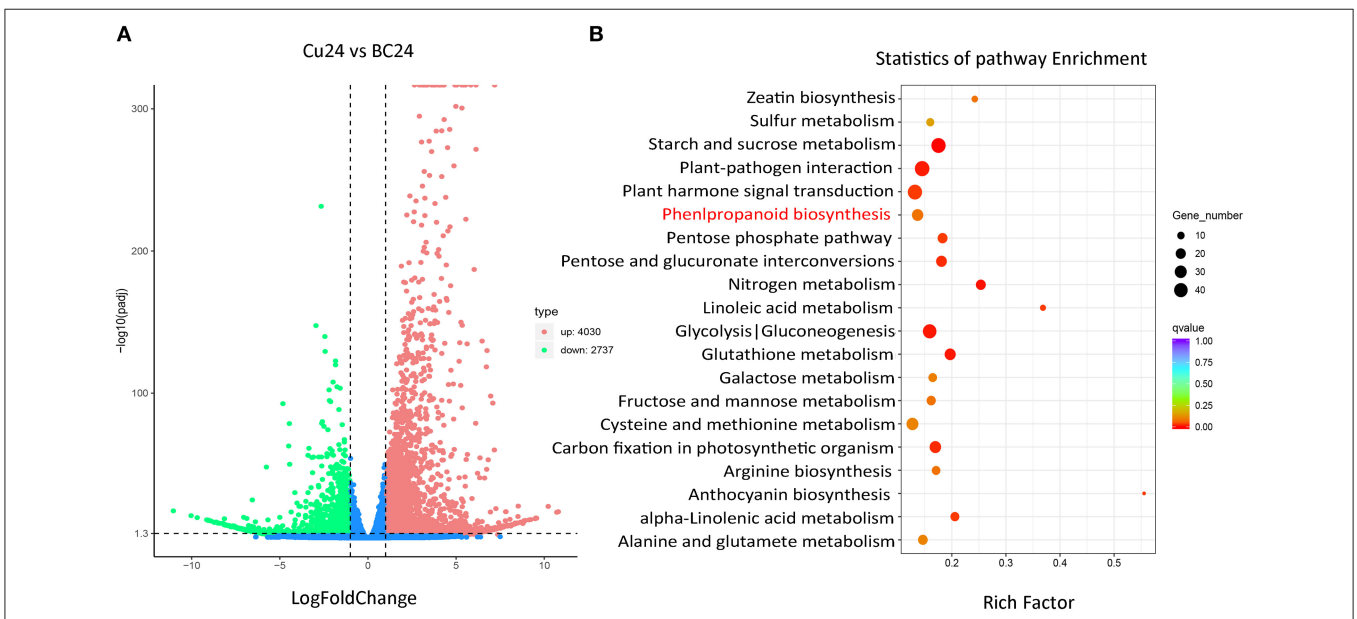

FIGURE 1 | Differential expression analysis of the *I. domestica* transcriptome induced by $CuCl_2$. **(A)** Volcano plots of differentially expressed genes; **(B)** Statistics of enrichment by KEGG. Cu24, the root samples treated by $CuCl_2$ for 24 h; BC24, the control groups. The control groups and the experimental groups are both in three independent biological replicates.

expression analysis. Using the standard criteria of an adjusted *p*-value < 0.005 and |log2 Foldchange| > 1, we obtained 6,767 differentially expressed genes. Among them, 4,030 were up-regulated in *I. domestica* roots treated with $CuCl_2$, and many unigenes were enriched in phenylpropanoid biosynthesis (**Figure 1**). Then, 19 candidate unigenes were selected from the up-regulated unigenes annotated as O-UGTs. To fully screen for possible unigenes involved in isoflavonoid glycoside biosynthesis, two putative UGTs (Cluster-1732.40383 and Cluster-1732.18310) were identified using a local BLAST algorithm-based search, and they shared similar conserved sequences with GmUGTs. The expression levels of these 21 candidate genes were up-regulated after $CuCl_2$ treatments, except Cluster-1732.18310, and the expression levels of most of the selected genes increased remarkably, having log2 Foldchanges > 2 (**Figure 2**). The sequences of the 21 genes were used as the query in a BLAST algorithm-based search of National Center for Biotechnology Information (NCBI) databases to predict their gene families and possible functions. We finally selected 11 BcUGTs putatively identified as (iso)flavone-O-UGTs to determine whether they were capable of transferring UDPG to (iso)flavone aglycones (**Figure 2**).

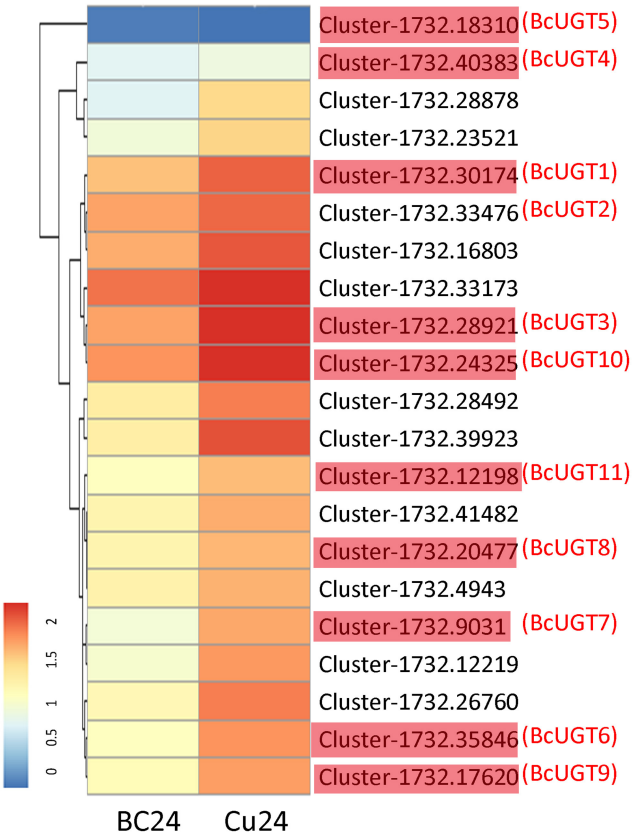

FIGURE 2 | Analysis of differentially expressed unigenes from *I. domestica*. Hierarchical clustering and corresponding heatmaps of the differentially expressed (iso)flavone glycosyltransferases across all pairwise library comparisons. The clusters in red (BcUGT1–11) were selected to verify the activity against flavones *in vitro*. Cu24, the root samples treated by $CuCl_2$ for 24 h; BC24, the control groups. The control groups and the experimental groups are both in three independent biological replicates.

Functional Characterization of Active BcUGTs *in vitro*

For *in vitro* biochemical assays, the open reading frame sequences of the selected 11 BcUGTs were all cloned independently into the prokaryotic expression vector pCold-TF and expressed in *E. coli* cells. The recombinant proteins purified with Ni-NTA columns showed strong and single protein bands at 105 kDa by SDS-PAGE (**Supplementary Figure 4**). The deduced molecular weights of the BcUGTs were ∼53 kDa, with the fusion protein having an additional soluble label and His-tag of ∼52 kDa. The protein bands identified by SDS-PAGE indicate that these soluble proteins may be used to determine the enzymes' activity levels and kinetic properties. The functional BcUGT assays were performed against the different (iso)flavone substrates: naringenin, apigenin, kaempferol, irigenin and tectorigenin. BcUGT1–8 showed activities toward (iso)flavones, while Cluster-1732.24325, Cluster-1732.17620 and Cluster-1732.12198 did not transfer a glycosyl group onto (iso)flavones, although they were annotated as (iso)flavone UGTs. BcUGT1–8, except BcUGT8, shared very high deduced amino acid sequence similarities, ranging from 81.7 to 97.3% (**Supplementary Table 4**), all of them were submitted to the UGT Nomenclature Committee and designated as UGT706L family genes.

The substrate conversion rates were determined using the flavones: naringenin, liquiritigenin, apigenin, baicalein, luteolin, quercetin, kaempferol, irigenin, tectorigenin, genistein, daidzein, and 6,7,4′-trihydroxyisoflavone (**Supplementary Figure 5**). BcUGT1, −2 and −4–7 displayed broad activities toward various (iso)flavone substrates, with flavonols being the most preferred (**Supplementary Figure 6**). However, for different types of substrates, BcUGTs exhibited different catalytic efficiencies. As shown in **Figure 3**, BcUGT1 exhibited similar catalytic activities toward different sugar acceptors, as did BcUGT2, while BcUGT4 was analogous to BcUGT5. Among these substrates, BcUGT3 only accepted luteolin and kaempferol as sugar acceptors, while BcUGT8 only exhibited relatively low activities toward luteolin and irigenin. The other flavonoids were all inert substrates for these enzymes (**Table 1**). The conversion rates of BcUGT4 and BcUGT5 for irigenin and tectorigenin were greater than other BcUGTs, by 2–5 times (**Table 2**).

None of the BcUGTs displayed any activity when daidzein and 6,7,4′-trihydroxyisoflavone were used as substrates. These isoflavones are rich in soybean, and their chemical structures lack 5-OH (**Table 1**). BcUGT1 and BcUGT5 converted 2.5-fold and 5.6-fold more naringenin, respectively, to their 7-O-glycosides compared with liquiritigenin. In addition, for daidzein and 6,7,4′-trihydroxy-isoflavone, none of the BcUGTs showed any activity. Thus, the existence of 5-hydroxy of (iso)flavone aglycones was very important for catalysis by the BcUGTs. In this study, BcUGT4 was chosen to investigate

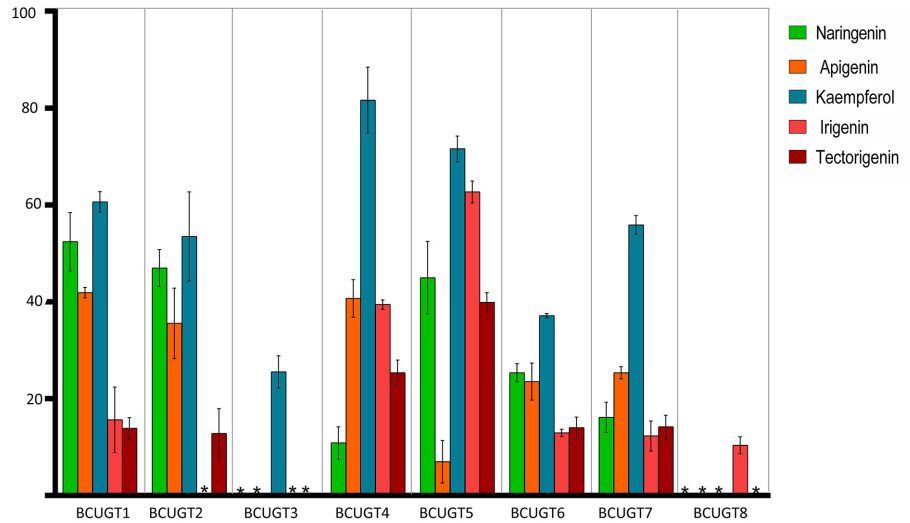

FIGURE 3 | Substrate conversion rates of BcUGTs with flavonoids. The reactions were performed at pH 8.0 in 30°C for 1 h with 100 μM each substrate, naringenin, apigenin, kaempferol, irigenin, and tectorigenin. The values were calculated as the averages of three replicates of each reaction. Asterisks, activity not detected.

TABLE 1 | Substrate conversion rates of BcUGTs with various flavonoids.

Substrate	Conversion rate %							
	BcUGT1	**BcUGT2**	**BcUGT3**	**BcUGT4**	**BcUGT5**	**BcUGT6**	**BcUGT7**	**BcUGT8**
Flavanone								
Naringenin	52.4 ± 6.0	47.0 ± 3.8	0.0	47.0 ± 3.8	45.0 ± 7.5	25.3 ± 1.9	16.1 ± 3.1	0.0
Liquiritigenin	20.6 ± 5.0	11.2 ± 3.4	0.0	16.1 ± 3.1	7.8 ± 2.4	16.2 ± 1.1	11.3 ± 1.5	0.0
Flavone								
Apigenin	41.9 ± 1.1	35.6 ± 7.3	0.0	40.7 ± 3.9	7.0 ± 4.4	25.5 ± 3.8	25.4 ± 1.2	0.0
Baicalein	42.5 ± 3.5	41.7 ± 1.5	0.0	9.1 ± 3.9	7.9 ± 6.4	15.9 ± 3.7	25.9 ± 2.1	0.0
Luteolin	29.9 ± 3.5	28.0 ± 4.0	13.7 ± 3.9	44.8 ± 9.2	21.3 ± 8.2	18.6 ± 3.7	35.2 ± 1.7	9.7 ± 2.4
Flavonol								
Quercetin	53.6 ± 7.5	62.5 ± 4.3	0.0	39.2 ± 7.1	26.3 ± 4.2	12.2 ± 0.8	12.4 ± 1.8	0.0
Kaempferol	60.6 ± 2.1	53.5 ± 9.2	25.5 ± 3.3	81.6 ± 6.9	71.6 ± 2.6	37.1 ± 0.5	55.8 ± 1.9	0.0
Isoflavone								
Tectorigenin	13.9 ± 2.2	12.8 ± 5.2	0.0	25.3 ± 2.7	39.9 ± 2.0	14.0 ± 2.2	14.2 ± 2.4	0.0
Irigenin	15.6 ± 6.8	0.0	0.0	39.4 ± 1.0	62.7 ± 2.3	12.9 ± 0.8	12.3 ± 3.1	10.4 ± 1.8
Genistein	11.1 ± 2.8	16.2 ± 7.6	0.0	39.9 ± 1.0	13.5 ± 1.9	24.8 ± 0.2	17.5 ± 2.3	0.0
Daidzein	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6,7,4'-Trihydroxy-isoflavone	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Reactions of BcUGTs with 100 μM substrate at pH 8.0 and 30°C for 1 h. Data are the means ± SDs of one experiment with three technical replicates.

the critical role of the 5-hydroxy in protein binding with substrates, we identified the top 50 templates according to the coverage, identity, and GMQE. A newly reported TcCGT1 (PDB) crystal structure was used as the template for homologous molecular modeling in SWISS-MODEL, the GMQE score of TcCGT was 0.75. A C-glycosyltransferase found in *Trollius chinensis*, TcCGT1, regio-specifically produces 8-C-glycosylation flavones with broad substrate spectra, and site-directed mutagenesis at just two residues switched C- to O-glycosylation (He et al., 2019). The GMQE of the model was 0.72, and the GMEAN was −2.59, GMEAN scores > −4 is an indication of models with high quality. Afterward, we modeled the structure of BcUGT4 docking with liquiritigenin and naringenin (**Figure 4**). The spacious binding pocket explained the broad substrate spectra of BcUGT4, with seven residues interacting with naringenin through hydrogen bonds having 1.8–2.8 Å distances. The residues GLN-394 and TYR-391 interacted with naringenin C-5-OH (2.8 Å, 2.1 Å), but for liquiritigenin, five residues, HIS-369, GLU-393, TRP-372, ASN-373, and GLU-377, were discovered around the ligand and they were also identified in the naringenin docking.

TABLE 2 | Kinetic parameters of BcUGTs with tectorigenin and irigenin.

Substrate	kat(s^{-1})	km(mM)	kcat/km($M^{-1}s^{-1}$)
BcUGT1			
Tectorigenin	57.29 ± 4.94	0.592 ± 0.049	9.68×10^4
Irigenin	58.88 ± 3.24	0.210 ± 0.013	2.80×10^5
BcUGT2			
Tectorigenin	70.10 ± 9.86	0.392 ± 0.007	1.79×10^5
Irigenin	–	–	–
BcUGT4			
Tectorigenin	155.18 ± 12.26	0.115 ± 0.023	1.35×10^6
Irigenin	57.23 ± 18.21	0.190 ± 0.007	3.01×10^5
BcUGT5			
Tectorigenin	116.03 ± 6.98	0.201 ± 0.015	5.77×10^5
Irigenin	296.52 ± 11.8	0.121 ± 0.009	2.45×10^6
BcUGT7			
Tectorigenin	15.86 ± 7.53	1.269 ± 0.009	1.98×10^4
Irigenin	17.73 ± 9.89	0.894 ± 0.011	1.98×10^4
BcUGT8			
Tectorigenin	–	–	—
Irigenin	11.97 ± 1.31	1.423 ± 0.021	8.41×10^3

Kinetic parameters were determined at pH 8.0 in 30°C with UDP-glucose, as described in the Materials and methods section. Data are the means ± SDs of one experiment with three technical replicates.

Structural Analysis of the New Compound

Reactions of BcUGTs with most of the sugar acceptors yielded single products, each of which was identified as its 7-O-glycoside on the basis of their co-chromatography with authentic samples using HPLC. However, as shown in **Figure 5**, when irigenin was used as acceptor substrate, BcUGT4 and BcUGT5 tended to yield double products, one of which (1a) was identified as iridin according to the mass spectrum and co-chromatography with authentic iridin, the molecular formula $C_{24}H_{26}O_{13}$ was deduced from the HR-ESI-MS ion peak at *m/z* 521.1832 $[M-H]^-$. The other main product (1b) was confirmed to be irigenin 3′-O glycoside using HR-ESI-MS, $_1$H-NMR, $_{13}$C-NMR, HMBC, and NEOSY. This structure has not been found previously in *I. domestica* or any other plant. Its UV (MeOH) λmax values were 268 nm, which is the typical absorption peak of isoflavones, and its IR absorptions at 3,424, 1,627, and 1,510 revealed the existence of hydroxyl and carbonyl groups (**Supplementary Figure 7**). The molecular formula $C_{24}H_{26}O_{13}$ was deduced from the HR-ESI-MS ion peak at *m/z* 545.1275 $[M+Na]^+$ (calcd for $C_{24}H_{26}O_{13}$ as 522.4544) (**Supplementary Figure 8**). In the $_1$H NMR spectrum, the most representative signal was the characteristic signal for isoflavone at δ8.43. There were three methoxyl group signals at δ3.81, δ3.76, and δ3.75. The signals at δ13.01 and δ10.93 were assigned to the hydroxy groups at the five and seven positions of isoflavone, and the signal of the end group hydrion of glucose occurred at δ4.91. The signals at δ6.96 (1H, d, 2.8) and δ6.96 (1H, d, 2.8) represented two AX coupling systems, which indicated the existence of incomplete symmetry tri-substituted groups in the B ring of the isoflavone skeleton (**Supplementary Figure 9**). The $_{13}$C NMR spectrum indicated 24 carbon resonances, which contained 19 carbon resonances of the isoflavone skeleton, 3 methoxyl group signals, and 6 carbon signals of glucose (**Supplementary Figure 10**). In the HMBC experiment, the resonating signal at δH 4.91 correlated with δC 151.2, which indicated that the glucose was transferred to isoflavonoid 3′-hydroxyl position (**Supplementary Figure 11**). $_1$H-NMR and $_{13}$C-NMR data for compound 1b and key HMBC correlations were shown in **Supplementary Table 5**. The NEOSY spectrum showed a resonating signal correlation between two hydrogens at a distance of two chemical bonds, and the resonating signal of δH 4.91 displayed a strong correlation with δH 6.96, which implied that 1′ H of glucose was close to 2′ H in irigenin. The NEOSY result further validated the structure of 3′-O glycoside (**Supplementary Figure 12**).

Kinetic Analysis of BcUGTs for Isoflavones

Although most of the BcUGTs exhibited broad substrate ranges and preference for the kaempferol aglycone, no data showed the existence of flavone O-glycosides or flavonol O-glycosides in *I. domestica,* only small amounts of flavones and flavonols have been reported, such as hispidulin, apigenin, rhamnazin, kaempferol, and quercetin (Jin et al., 2008). Consequently, we selected irigenin and tectorigenin as the substrates for kinetic parameter determinations (**Table 2**). The UDPG concentration was set at 2 mM, and the substrate concentration ranged from 10 to 400 μM for the assays. The reaction time was strictly controlled at 10 min. The km values of BcUGT4 and BcUGT5 for irigenin and tectorigenin were the lowest among all the active BcUGTs, and BcUGT5 exhibited the highest kcat/km ratio for irigenin, at 2.45×10^6, which was approximately four times the kcat/km ratio for tectorigenin. Using BcUGT4, the kcat/km ratio for tectorigenin was approximately 4.5 times that for irigenin. Thus, although BcUGT4 shared a 94.2% sequence identity with BcUGT5, their substrate specificities were quite different. In addition, the sequence identity between BcUGT1 and BcUGT3 was 97.3%, but their catalytic specificities were totally different. BcUGT3 did not show activity against isoflavones, but BcUGT1 exhibited broad substrate ranges, although its kcat/km ratio for irigenin was just 2.8×10^5. BcUGT8 shared only about 70% sequence identity with other BcUGTs, so irigenin and tectorigenin may not be it's natural substrates, and BcUGT8 could almost not glycosylate tectorigenin (**Supplementary Table 4**).

Phylogenetic Analysis of BcUGTs

According to the phylogenetic tree, BcUGTs showed a closer relationship with the UDP-UGT706 family (**Figure 6**). UGT706D1 and UGT709A4 were reported to be capable of forming isoflavone 7-O glucosides *in vitro*, while UGT706C2 could use flavones to produce flavone 3-O glucosides (Ko et al., 2008), all BcUGTs showed a closer relationship with UGT706D1 and UGT706C2, and sequence identities were 40–50%. Meanwhile, BCUGTs had 40–50% amino acid sequence similarities with the UGTs of *G. max* and *P. lobata*, which have been confirmed to have activities toward (iso)flavone aglycones. Although all the BcUGTs selected for phylogenetic tree construction had similar sequence similarities with UGT88 family group and UGT706 family group, they clustered into rice

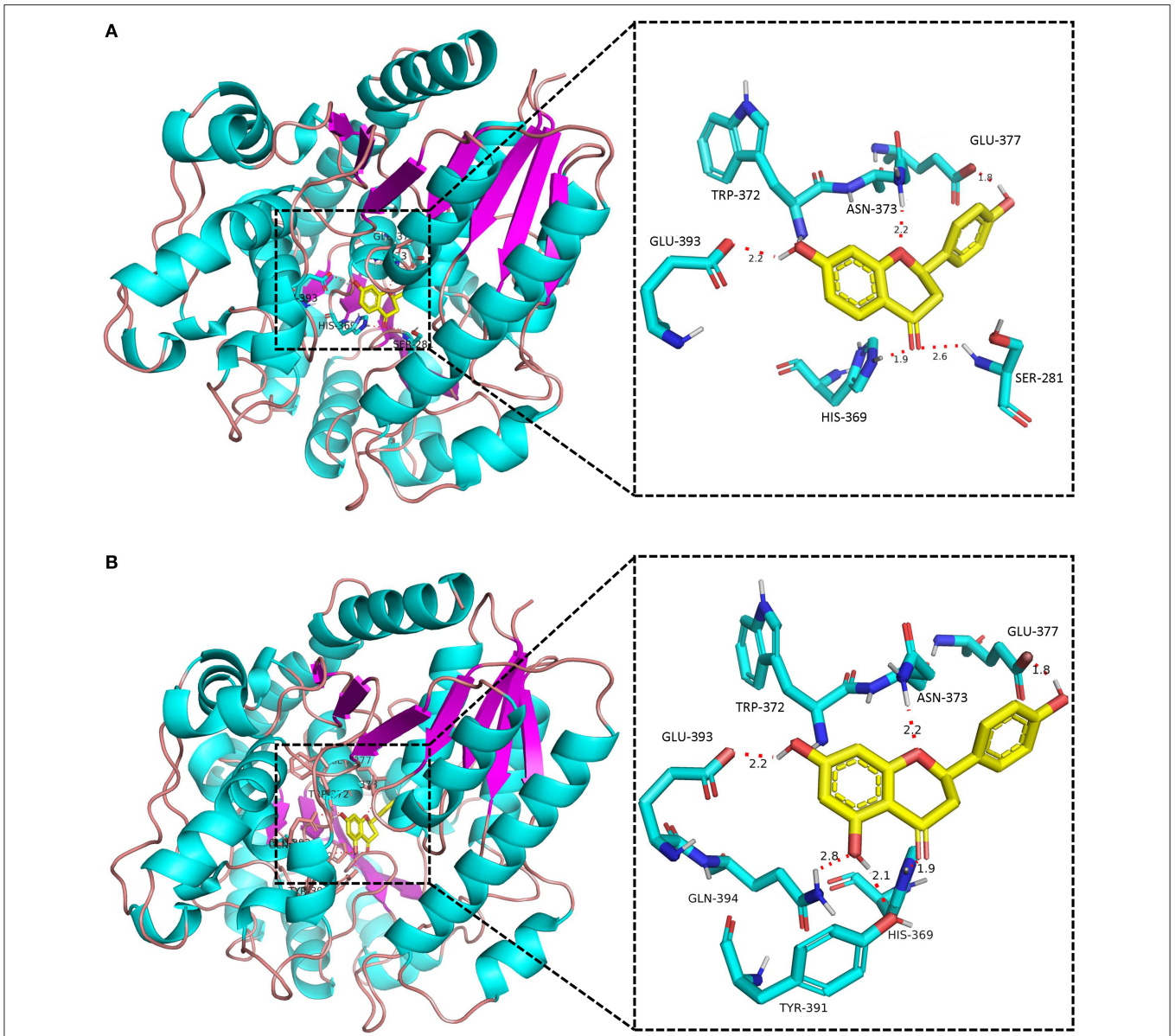

FIGURE 4 | Docking model of BcUGT4 with liquiritigenin and naringenin. The protein model of BcUGT4 was established using SWISS-MODEL. The enlarged region shows polar interactions of substrates binding with BcUGT4 residues. **(A)** Autodocking result of BcUGT4 with liquiritigenin; **(B)** Autodocking result of BcUGT4 with naringenin.

UGT706 clade, perhaps owing to the genetic distance between monocots and eudicots. Genes chosen in the phylogenetic tree were mainly those of *G. max* and *P. lobata*, but they exhibited different substrate specificities and regio-selectivities. PlUGT1 and PlUGT13 showed enzyme activities for isoflavone substrates at the 7-hydroxy group, while other flavonoids were hardly accepted. PlUGT2 was later shown to be the isoflavone 4′,7-O-diglucosides' UGT. GmUGT4 was specific for isoflavones at the 7-hydroxy group, but other GmUGTs showed broad glycosyl-acceptor specificities (Li et al., 2014; Funaki et al., 2015; Wang et al., 2016). The PgUGT84A23 and PgUGT84A24 played important roles in formation of hydrolyzable tannins (HTs) in pomegranate, but they also exhibited genistein 7-O-glycosylation activity *in vitro* (Ono et al., 2016), and another PgUGT95B2 showed genistein 7-O-glycosylation activity *in vitro* as well (Wilson et al., 2019). These reports suggested BcUGTs' ability to glycosylate different flavones or isoflavones at their 7-hydroxy positions, but *in vitro* experiments revealed that their specificities and regio-selectivities for these sugar acceptors differed. A further division of the clusters among the 10 unigenes from *I. domestica* revealed that Cluster-1732.24325, Cluster-1732.17620, and BcUGT8 did not fall into the same

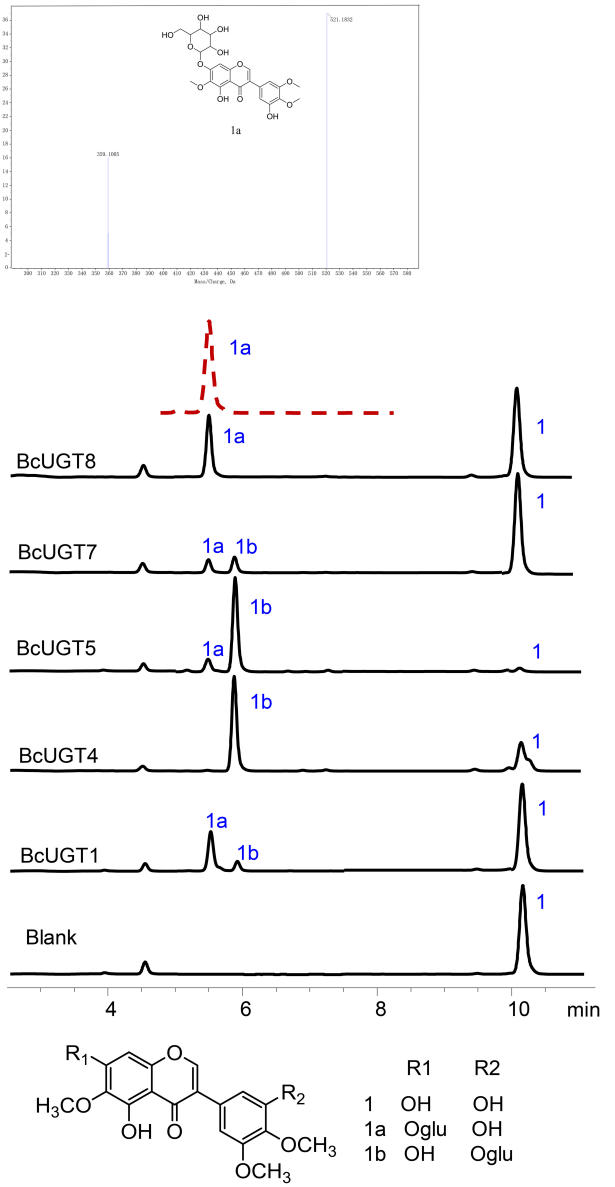

FIGURE 5 | HPLC analyses of BcUGT reactions with irigenin and mass spectrum of 1a. A reaction mixture of irigenin and UDP-glucose with BcUGTs. 1, Irigenin; 1a, Irigenin-7-O-glycoside; 1b, Irigenin-3′-O-glycoside; red dashed line, Authentic irigenin-7-O-glycoside; Chromatograms were obtained with absorption at 268 nm.

subcluster with the other BcUGTs, and Cluster-1732.24325 and Cluster-1732.17620 had no activities toward flavones. Additionally, BcUGT8 only displayed a weak catalytic efficiency toward irigenin. BcUGT1 and BcUGT2 clustered into the same branch, and BcUGT6 and BcUGT7 displayed a close genetic relationship. Cluster-1732.12198 had a relatively higher homology to PlUGT43, which was reported to be a daidzein C-glycosyltransferase cloned from *P. lobata* (Wang et al., 2017). However, to date, there have been no isoflavone C-UGTs detected in *I. domestica*, and Cluster-1732.12198 did not transfer glucose to any flavones.

Expression Level Analyses

The transcription levels of the active BcUGTs (BcUGT1–8) in *I. domestica* seedling roots were detected using quantitative real-time RT-PCR with primers having good specificity and amplification efficiency levels (**Supplementary Figure 13**). After treating with $CuCl_2$, the transcript abundances in seedling roots increased to different degrees, especially for BcUGT1 and BcUGT3, in addition, the expression levels of BcUGT1 and BcUGT3 were significantly higher than those of other BcUGTs (**Figure 7A**). The results showed good correlations with the FPKM (fragments per kilobase of exon model per million mapped reads) of each BcUGT calculated from the RNA-seq data. We found that the expression levels of all the active BcUGTs in seedlings not receiving $CuCl_2$ treatments were very low, and some of them, such as BcUGT5 and BcUGT8, were almost undetectable. Thus, $CuCl_2$ plays a crucial role in inducing the expression of the eight active BcUGTs. The expression levels of BcUGTs in seedlings receiving $MgCl_2$, $ZnCl_2$ and $CaCl_2$ treatments were very low as well (**Supplementary Figure 14**). Moreover, $ZnCl_2$ and $MgCl_2$ suppressed the expression of all BcUGTs compared with the control group, $CaCl_2$ showed almost no effect in inducing the expression of the eight BcUGTs. These data indicated that Cu^{2+} plays a crucial role in inducing the expression of the BcUGTs, and we could rule out the effect of the chlorine ion in the induction.

The transcript abundances of all the BcUGTs in the underground parts were significantly greater than in the aerial parts, which correlated well with the isoflavone glycosides' distribution pattern (**Figure 7B**). The rhizome accumulated the most isoflavones, but the BcUGTs were more highly expressed in roots than rhizomes (**Figure 8**). To further explore the relationship of BcUGT expression levels between roots and rhizomes, we divided the triennial rhizomes into five parts, lateral, bulb, vertical, sprout primordium, and cork layer (Chen et al., 2014). The vertical part is the first and sprout primordium is the last during rhizome growth, and the lateral and bulb are biennial parts in rhizomes. Among the five rhizome parts, the sprout primordium and vertical part displayed higher BcUGT expression levels than other parts. However, it was obvious that the lateral and bulb parts had the most roots, while the other parts, especially the sprout primordium, had almost no roots. In the cork layer, previous experiments revealed that isoflavone aglycones were the most abundant and almost no isoflavone glycosides were detected (Chen et al., 2014). This may be explained by the low BcUGT expression levels.

DISCUSSION

To date, the biosynthesis of isoflavonoids, including isoflavone 7-O-, 4′-O-, and C-UGTs, has been mainly studied in legumes, with the only exceptions being studies from pomegranate and rice. Thus, little is known regarding the genes involved in isoflavonoid biosynthesis in other plants containing abundant isoflavones.

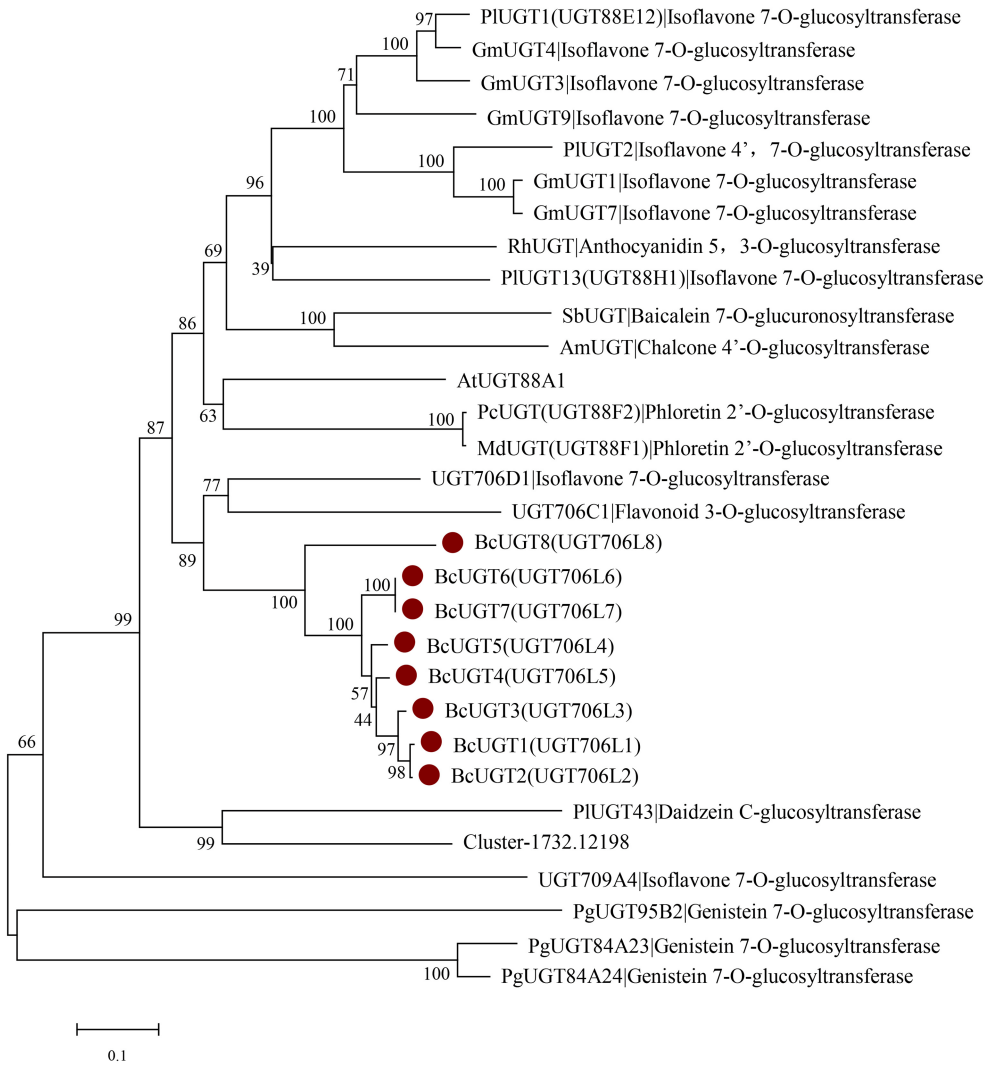

FIGURE 6 | Rooted phylogenetic tree of BcUGTs, UGT88-, UGT706-, and UGT84-related (iso)flavone O-glycosyltransferase genes. *Glycine max*: GmUGT1 (A6BM07.1), GmUGT3 (NP_001353995.1), GmUGT4 (NP_001304440.2), GmUGT7 (NP_001304487.1), GmUGT9 (accession: NP_001280039.1), *Pueraria montana* var. *lobate*: PlUGT1 (A0A067YB04.1), PlUGT2 (A0A172J2D0.1), PlUGT13 (A0A067YBQ3.1), *Arabidopsis thaliana*: AtUGT88A1 (Q9LK73.1), *Scutellaria baicalensis*: SbUGT (Q76MR7.1), *Antirrhinum majus*: AmUGT (Q33DV3.1), *Rosa hybrid* cultivar: RhUGT (Q4R1I9.1), *Malus domestica*: MdUGT (B3TKC8.1), *Pyrus communis*: PcUGT (D3UAG3.1), *Punica granatum*: PgUGT95B2 (MH507175), PgUGT84A23 (KT159805.1), PgUGT84A24 (KT159807.1), *Oryza sativa* Japonica Group: UGT709A4 (BAC80066.1), UGT706D1 (BAB68093.1), UGT706C1 (BAB68090.1).

Iris domestica accumulates many isoflavonoid glycosides, such as tectoridin, iridin, iristectorin A, and iristectorin B, in its rhizomes. Using copper stress, we obtained eight unigenes that were active toward (iso)flavones. A sequence alignment showed that the active BcUGTs shared a very high sequence homology (**Supplementary Figure 15**).

Cu^{2+} Mechanism for Up-Regulating BcUGT Expression

Plants are continuously exposed to abiotic stresses during their growth in the wild, including various heavy metal stresses (Piscopo et al., 2016). Consequently, they have evolved a series of skills to deal with the rapid changes in the environment. Elements such as Fe, Cu, Zn, Mo, and Ni are considered to be essential heavy metals (HMs) (Hall, 2002; Arif et al., 2016), but excessive HMs may be toxic for plants, resulting in large-scale molecular, biochemical and physiological responses (Hayat et al., 2012; Hussain et al., 2013; Clemens and Ma, 2016). The perception of stress stimuli is the initial plant response when exposed to toxic HMs. Then, the amplified signal is transduced and transmitted into cells to bring about cascade reactions of molecular, biochemical and physiological activities (Wani et al., 2018). Through transcription analyses of *I. domestica* roots, many P450 genes related to the biosynthesis and catabolism of methyl jasmonate and allene oxide (Maksymiec et al., 2005), including CYP94 (cluster1732.31596 and cluster1732.31816) and CYP74

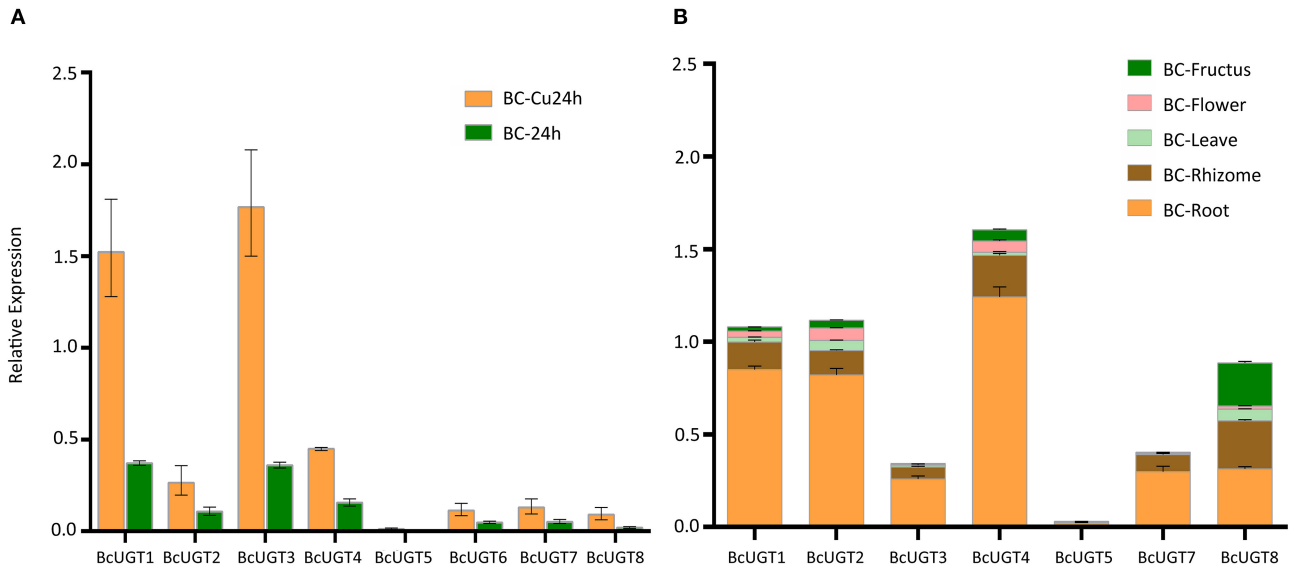

FIGURE 7 | Transcriptional levels of genes coding for flavonoid-active BcUGTs. **(A)** Transcription levels of BcUGTs in the seedlings of *Iris domestica* treated with $CuCl_2$ for 24 h (BC-Cu24h) and the control group (BC-24h). **(B)** Transcription levels in fructus, flower, leaf, rhizome, and root of *I. domestica* collected from the wild. Data are presented as the means ± SDs from three independent biological replicates and three technical replicates.

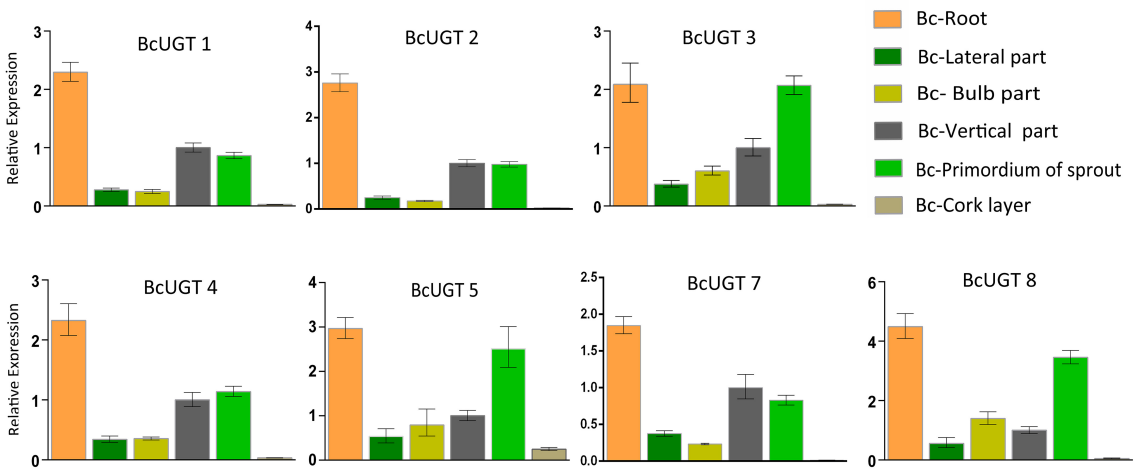

FIGURE 8 | Relative transcription levels of active BcUGTs in different rhizome parts. The transcription levels of flavonoid-active BcUGTs in the lateral part, blub part, vertical part, sprout primordium, and cork layer of the rhizome. Data are presented as the means ± SDs from three independent biological replicates and three technical replicates.

(cluster 1732.20141), were found to be up-regulated after $CuCl_2$ treatments (Utsunomiya et al., 2000; Koo et al., 2011).

At the biochemical level, plants need more secondary metabolites to cope with excess HMs, resulting in the high expression levels of many unigenes involved in the biosynthesis of secondary metabolites. $CuCl_2$ was first applied to induce isoflavonoids in *P. lobata* stems, and the daidzein and genistein contents in the $CuCl_2$-treated stem increased 5–10-fold compared with the control group (Hakamatsuka et al., 1991). The cotyledons of *Lupinus albus* have been treated with a fungal elictor and $CuCl_2$, which induced the isoflavone 2′-hydroxygenistein. When challenged with $CuCl_2$, isoflavone glycosidic conjugates are largely degraded to isoflavone aglycones, which indicates the important roles of isoflavonoid glycosides in plants coping with Cu^{2+} toxicity (Wojtaszek and Stobiecki, 1997). Additionally, the dynamic change in isoflavonoid glycosides contents in *I. domestica* calli suggested that $CuCl_2$ promotes the biosynthesis of isoflavone glycosides during the first stage of $CuCl_2$ stress. Afterward, when the Cu^{2+} concentration in cells reaches a high level, the

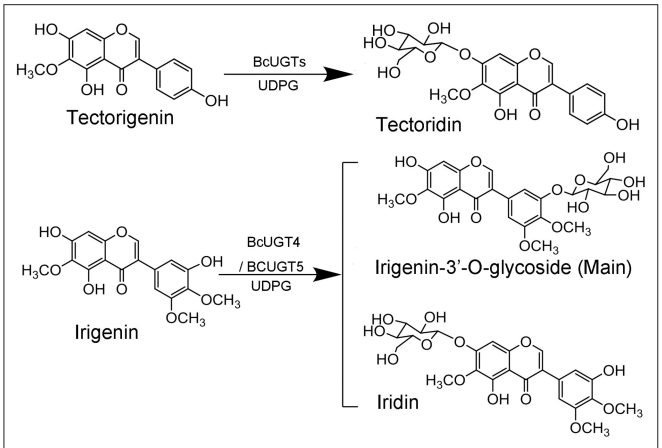

FIGURE 9 | Isoflavone O-glucosyltransferase-catalyzed reaction.

isoflavone glycoside is hydrolyzed to its aglycone. Thus, we speculated that isoflavone glycosides play very important roles in plant defenses against Cu^{2+} stress. Through the differential expression analysis, several up-regulated unigenes annotated as (iso)flavone 7-O-UGTs, which belonged to the UGT706 family, were cloned. Then, these unigenes were expressed in a recombinant *E. coli* expression system, and their functions were verified using *in vitro* catalytic experiments. These results implied that the high $CuCl_2$ concentration enhanced the expression of (iso)flavone O-UGTs in *I. domestica*.

Substrate Preference and Regio-Specificity of BcUGTs

Many flavone UGTs show more regio-specificity than substrate specificity (Kim et al., 2006; Li et al., 2014; Funaki et al., 2015; Wang et al., 2016). In soybean, GmUGT1 and GmUGT7 efficiently catalyze regio-specific glucosyl transfers to isoflavones at their 7-hydroxy positions, and they accept naringenin, apigenin, quercetin, and kaempferol as substrates (Funaki et al., 2015). In an important crop, *Camellia sinensis*, CsUGT73A20 was identified as a flavonoid UGT. CsUGT73A20 transfers UDP-glucose to produce 3-O- and 7- O-glycosides, depending mainly on the reaction's *in vitro* pH value. Kaempferol is the most preferred substrate, but naringenin and apigenin are also accepted (Zhao et al., 2017). Recently, three new (iso)flavonoid glycosyltransferases (PlUGT4, −15 and −57) were identified in *P. lobate*. PlUGT15 and PlUGT57 exhibit much higher isoflavone specificity than other 7-O-UGTs found in *P. lobate* (Wang X. et al., 2019).

Kaempferol is the most preferred acceptor among the 12 substrates tested for all the BcUGTs. The major product yielded by BcUGTs was (iso)flavone 7-O-glycoside, which indicates that the BcUGTs preferred to transfer the sugar moiety to the hydroxy of ring A, but no (iso)flavone 4′-O-glycosides or 4′,7-O-di-glycosides were obtained. Thus, most of the BcUGTs had no activity toward the hydroxy of ring B. Tectorigenin is enriched in the rhizomes and roots of *I. domestica*, and all the BcUGTs, except BcUGT3 and BcUGT8, converted tectorigenin to tectoridin, which exists in large amounts in *I. domestica*. Irigenin is another main isoflavone aglycone in the rhizomes and roots of *I. domestica*, but we unexpectedly found that the main product of BcUGT4 and BcUGT5 using irigenin as the substrate was not iridin (**Figure 9**). We identified its structure as irigenin 3′-O-glycoside. This is a new compound that has never been identified in *I. domestica* or other plants. However, other BcUGTs hardly glycosylated irigenin at the 3′-OH position. BcUGT4 yielded two different products using luteolin as the substrate, but the main product was luteolin-7-O-glycoside. Thus, we speculated that the binding conformation of the substrate determines the glycosylation position for BcUGT4 and BcUGT5.

Interestingly, substrates lacking a 5-hydroxy were hardly glycosylated by BcUGTs. Recently, a promiscuous flavonoid-3-O-UGT (sb3GT1) was identified in *Scutellaria baicalensis*, and it accepts many sugar donors to glycosylate 17 flavonols. Through homologous molecular modeling and docking, 18 residues within a 5-Å distance of the binding pocket were selected as candidates, and residue G15 was identified as the critical residue in the sugar donor promiscuity of Sb3GT1 by comparisons with the docking results of *Vitis vinifera* GT1 (Wang Z. L. et al., 2019). In this study, the docking results of BcUGT4 with liquiritigenin and naringenin indicated that the residues GLN-394 and TYR-391 could form hydrogen bond with 5-hydroxy, which may stabilize the substrate in the binding pocket, further study could confirm the role of the 5-hydroxy and the coordinating amino acids in substrate binding.

Predicting the Function of BcUGTs

The different expression levels of BcUGTs in the transcriptome of *I. domestica* seedlings indicate that BcUGTs may play important roles in responses to abiotic stresses. In plants, glycosylation is a key modification that allows the formation of a myriad of secondary metabolites, which play many important roles in plant responses to challenges faced during growth and development (Bowles et al., 2005; Gachon et al., 2005; Campos et al., 2019). PlUGT4, −15 and −57 are up-regulated in *P. lobata* roots after methyl jasmonate treatments (Wang X. et al., 2019), and PlUGT1 may be related to the defense response against the MeJA treatment. In this paper, all the BcUGTs were up-regulated after $CuCl_2$ stress, especially BcUGT1 and BcUGT2, which have highly homologous nucleotide sequences and broad substrate ranges. The unigenes encoding the core metabolic pathway usually form small families consisting of unique genes. For example, isoflavone synthases belong to the CYP93C family, which only consists of a few members in leguminous plants (Jung et al., 2000). This is in sharp contrast to the multiple gene families of UGTs, which tend to contain hundreds of members. In addition, the broad substrate ranges may ensure that plants respond rapidly to both abiotic and biotic stresses in the environment.

For wild *I. domestica*, we found that the transcription levels of most BcUGTs in the vertical parts (triennial rhizome) and sprout primordium were greater than in the lateral and bulb parts (biennial rhizome). Additionally, we found that lateral and bulb parts had more roots than the other parts; consequently, we speculated that this is the reason for the differences in BcUGT expression among these rhizome parts. Roots form the main

organ that delivers water and nutrients from the soil to the aerial plant parts; therefore, we hypothesized that many isoflavone glycosides synthetized in roots by BcUGTs are delivered to the lateral and bulb parts, while the sprout primordium and vertical part have no isoflavone glycoside source. Consequently, they have to increase the BcUGT expression levels to defend against the stress. Recently, isoflavone membrane transformation in *Trifolium pratense* was shown to be related to the presence of ABC proteins and vesicular transport (Kubes et al., 2019). By analyzing global transcriptomes of six organs from *I. domestica*, many ABC transporter family unigenes were obtained specifically from the roots.

DATA AVAILABILITY STATEMENT

The sequencing project has been deposited at the SRA database under the accession number PRJNA596865. The SRA accession numbers are SRR10743030 and SRR107430. The accession numbers for the eight BcUGTs are as follows: BcUGT1 (MN894539), BcUGT2 (MN894540), BcUGT3 (MN894541), BcUGT4 (MN894542), BcUGT5 (MN894543), BcUGT6 (MN894544), BcUGT7 (MN894545), and BcUGT8 (MN894546).

AUTHOR CONTRIBUTIONS

XZ, YaZ, GX, and MQ designed research. XZ, JY, and ZY performed research and all authors analyzed the data. XZ, YaZ, YuZ, and MQ wrote the paper. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 81872958, 81373918, and 81703637).

ACKNOWLEDGMENTS

We thank International Science Editing (https://www.internationalscienceediting.com/) for editing this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.632557/full#supplementary-material

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi: 10.1186/gb-2010-11-10-r106

Arif, N., Yadav, V., Singh, S., Singh, S., Ahmad, P., Mishra, R., et al. (2016). Influence of high and low levels of plant-beneficial heavy metal ions on plant growth and development. *Front. Env. Sci. Switz.* 4:69. doi: 10.3389/fenvs.2016.00069

Benkert, P., Tosatto, S. C. E., and Schomburg, D. (2008). QMEAN: a comprehensive scoring function for model quality assessment. *Proteins* 71, 261–277. doi: 10.1002/prot.21715

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 42, W252–W258. doi: 10.1093/nar/gku340

Bowles, D., Isayenkova, J., Lim, E. K., and Poppenberger, B. (2005). Glycosyltransferases: managers of small molecules. *Curr. Opin. Plant Biol.* 8, 254–263. doi: 10.1016/j.pbi.2005.03.007

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

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:1. doi: 10.1186/1471-2105-10-421

Campos, L., López-Gresa, M. P., Fuertes, D., Bellés, J. M., Rodrigo, I., and Lisón, P. (2019). Tomato glycosyltransferase Twi1 plays a role in flavonoid glycosylation and defence against virus. *BMC Plant Biol.* 19:450. doi: 10.1186/s12870-019-2063-9

Chen, Y. J., Liang, Z. T., Zhu, Y., Xie, G. Y., Tian, M., Zhao, Z. Z., et al. (2014). Tissue-specific metabolites profiling and quantitative analyses of flavonoids in the rhizome of *Belamcanda chinensis* by combining laser-microdissection with UHPLC-Q/TOF-MS and UHPLC–QqQ-MS. *Talanta* 130, 585–597. doi: 10.1016/j.talanta.2014.07.004

Clemens, S., and Ma, J. F. (2016). Toxic heavy metal and metalloid accumulation in crop plants and foods. *Annu. Rev. of Plant Biol.* 67, 489–512. doi: 10.1146/annurev-arplant-043015-112301

Clúa, J., Roda, C., Zanetti, M. E., and Blanco, F. A. (2018). Compatibility between legumes and rhizobia for the establishment of a successful nitrogen-fixing symbiosis. *Genes* 9:125. doi: 10.3390/genes9030125

Funaki, A., Waki, T., Noguchi, A., Kawai, Y., Yamashita, S., Takahashi, S., et al. (2015). Identification of a highly specific isoflavone 7-O-glucosyltransferase in the soybean (*Glycine max (L.)* Merr.). *Plant Cell Physiol.* 56, 1512–1520. doi: 10.1093/pcp/pcv072

Gachon, C. M., Langlois-Meurinne, M., and Saindrenan, P. (2005). Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends Plant Sci.* 10, 542–549. doi: 10.1016/j.tplants.2005.09.007

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

Graham, T. L., Kim, J. E., and Graham, M. Y. (1990). Role of constitutive isoflavone conjugates in the accumulation of glyceollin in soybean infected with Phytophthora-Megasperma. *Mol. Plant Microbe Interact.* 3, 157–166. doi: 10.1094/MPMI-3-157

Hakamatsuka, T., Ebizuka, Y., and Sankawa, U. (1991). Induced isoflavonoids from copper chloride-treated stems of *Pueraria lobata*. *Phytochemistry* 30, 1481–1482. doi: 10.1016/0031-9422(91)84191-T

Hall, J. L. (2002). Cellular mechanisms for heavy metal detoxification and tolerance. *J. Exp. Bot.* 53, 1–11. doi: 10.1093/jexbot/53.366.1

Hayat, S., Khalique, G., Irfan, M., Wani, A. S., and Tripathi, B. N. (2012). Physiological changes induced by chromium stress in plants: an overview. *Protoplasma* 249, 599–611. doi: 10.1007/s00709-011-0331-0

He, J. B., Zhao, P., Hu, Z. M., Liu, S., Kuang, Y., Zhang, M., et al. (2019). Molecular and structural characterization of a promiscuous C-glycosyltransferase from *Trollius chinensis*. *Angew Chem. Int. Edit.* 58, 11513–11520. doi: 10.1002/anie.201905505

Hussain, M. B., Ali, S., Azam, A., Hina, S., Farooq, M. A., Ali, B., et al. (2013). Morphological, physiological and biochemical responses of plants to nickel stress: a review. *Afr. J. Agric. Res.* 8, 1596–1602. doi: 10.5897/AJAR12.407

Jiang, J. R., Yuan, S., Ding, J. F., Zhu, S. C., Xu, H. D., Chen, T., et al. (2008). Conversion of puerarin into its 7-O-glycoside derivatives by Microbacterium oxydans (CGMCC 1788) to improve its water solubility and pharmacokinetic properties. *Appl. Microbiol. Biot.* 81, 647–657. doi: 10.1007/s00253-008-1683-z

Jin, L., Chen, H. S., Jin, Y. S., Liang, S., Xiang, Z. B., and Lu, J. (2008). Chemical constituents from *Belamcanda chinensis*. *J. Asian Nat. Prod. Res.* 10, 89–94. doi: 10.1080/10286020701273619

Jung, W., Yu, O., Lau, S. M. C., O'Keefe, D. P., Odell, J., Fader, G., et al. (2000). Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nat. Biotechnol.* 18, 208–212. doi: 10.1038/72671

Kang, K., Lee, S. B., Jung, S. H., Cha, K. H., Park, W. D., Sohn, Y. C., et al. (2009). Tectoridin, a poor ligand of estrogen receptor α, exerts its estrogenic effects via an ERK-dependent pathway. *Mol. Cells* 27, 351–357. doi: 10.1007/s10059-009-0045-8

Karre, S., Kumar, A., Yogendra, K., Kage, U., Kushalappa, A., and Charron, J. B. (2019). HvWRKY23 regulates flavonoid glycoside and hydroxycinnamic acid amide biosynthetic genes in barley to combat Fusarium head blight. *Plant Mol. Biol.* 100, 591–605. doi: 10.1007/s11103-019-00882-2

Kim, J. H., Kim, B. G., Park, Y., Ko, J. H., Lim, C. E., Lim, J., et al. (2006). Characterization of flavonoid 7-O-glucosyltransferase from *Arabidopsis thaliana*. *Biosci. Biotech. Bioch.* 70, 1471–1477. doi: 10.1271/bbb.60006

Ko, J. H., Kim, B. G., Kim, J. H., Kim, H., Lim, C. E., Lim, J., et al. (2008). Four glucosyltransferases from rice: cDNA cloning, expression, and characterization. *J. Plant Physiol.* 165, 435–444. doi: 10.1016/j.jplph.2007.01.006

Koo, A. J. K., Cooke, T. F., and Howe, G. A. (2011). Cytochrome P450 CYP94B3 mediates catabolism and inactivation of the plant hormone jasmonoyl-L-isoleucine. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9298–9303. doi: 10.1073/pnas.1103542108

Krämer, R. P., Hindorf, H., Jha, H. C., Kallage, J., and Zilliken, F. (1984). Antifungal activity of soybean and chickpea isoflavones and their reduced derivatives. *Phytochemistry* 23, 2203–2205. doi: 10.1016/S0031-9422(00)80520-8

Kubes, J., Skalicky, M., Tumova, L., Martin, J., Hejnak, V., and Martinkova, J. (2019). Vanadium elicitation of *Trifolium pratense L.* cell culture and possible pathways of produced isoflavones transport across the plasma membrane. *Plant Cell Rep.* 38, 657–671. doi: 10.1007/s00299-019-02397-y

Kumar, S., Stecher, G., and Tamura, K. (2016). Mega7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Kwon, S. H., Kang, M. J., Huh, J. S., Ha, K. W., Lee, J. R., Lee, S. K., et al. (2007). Comparison of oral bioavailability of genistein and genistin in rats. *Int. J. Pharmaceut.* 337, 148–154. doi: 10.1016/j.ijpharm.2006.12.046

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0

Lee, Y. S., Kim, S. H., Kim, J. K., Lee, S., Jung, S. H., and Lim, S. S. (2011). Preparative isolation and purification of seven isoflavones from *Belamcanda chinensis*. *Phytochem. Anal.* 22, 468–473. doi: 10.1002/pca.1306

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323

Li, J., Li, Z. B., Li, C. F., Gou, J. B., and Zhang, Y. S. (2014). Molecular cloning and characterization of an isoflavone 7-O-glucosyltransferase from *Pueraria lobate*. *Plant Cell Rep.* 33, 1173–1185. doi: 10.1007/s00299-014-1606-7

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

Maksymiec, W., Wianowska, D., Dawidowicz, A. L., Radkiewicz, S., Mardarowicz, M., and Krupa, Z. (2005). The level of jasmonic acid in *Arabidopsis thaliana* and *Phaseolus coccineus* plants under heavy metal stress. *J. Plant Physiol.* 162, 1338–1346. doi: 10.1016/j.jplph.2005.01.013

Modolo, L. V., Blount, J. W., Achnine, L., Naoumkina, M. A., Wang, X., and Dixon, R. (2007). A functional genomics approach to (iso)flavonoid glycosylation in the model legume *Medicago truncatula*. *Plant Mol. Biol.* 64, 499–518. doi: 10.1007/s11103-007-9167-6

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco cultures. *Physiol. Plant.*15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Ono, N. N., Qin, X. Q., Wilson, A. E., Li, G., and Tian, L. (2016). Two UGT84 family glycosyltransferases catalyze a critical reaction of hydrolyzable Tannin Biosynthesis in pomegranate (*Punica granatum*). *PLoS ONE* 11, 1–25. doi: 10.1371/journal.pone.0156319

Piscopo, M., Ricciardiello, M., Palumbo, G., and Troisi, J. (2016). Selectivity of metal bioaccumulation and its relationship with glutathione S-transferase levels in gonadal and gill tissues of *Mytilus galloprovincialis* exposed to Ni (II), Cu (II) and Cd (II). *Rend. Lincei Sci. Fis.* 27, 737–748. doi: 10.1007/s12210-016-0564-0

Seidlová-Wuttke, D., Hesse, O., Jarry, H., Rimoldi, G., Thelen, P., Christoffel, V., et al. (2004). *Belamcanda chinensis* and the thereof purified tectorigenin have selective estrogen receptor modulator activities. *Phytomedicine* 11, 392–403. doi: 10.1016/j.phymed.2004.01.003

Steinegger, M., Meier, M., Mirdita, M., Vöhringer, H., Haunsberger, S. J., and Söding, J. (2019). HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* 20:473. doi: 10.1186/s12859-019-3019-7

Tian, M., Zhang, X., Zhu, Y., Xie, G. Y., and Qin, M. J. (2018). Global transcriptome analyses reveal differentially expressed genes of six organs and putative genes involved in (Iso)flavonoid biosynthesis in *Belamcanda chinensis*. *Front. Plant Sci.* 9:1160. doi: 10.3389/fpls.2018.01160

Tomoyoshi, A., Masayuki, I., Toshio, A., and Shin-ichi, A. (2005). Isoflavonoid production by adventitious-root cultures of *Iris germanica*. *Plant Biotechnol.* 22, 207–215. doi: 10.5511/plantbiotechnology.22.207

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

Udomsuk, L., Jarukamjorn, K., Tanaka, H., and Putalum, W. (2010). Improved isoflavonoid production in *Pueraria candollei* hairy root cultures using elicitation. *Biotechnol. Lett.* 33, 369–374. doi: 10.1007/s10529-010-0417-3

Utsunomiya, Y., Jarukamjorn, K., Tanaka, H., and Putalum, W. (2000). Purification and inactivation by substrate of an allene oxide synthase (CYP74) from corn (*Zea mays* L.) seeds. *Phytochemistry* 53, 319–323. doi: 10.1016/S0031-9422(99)00534-8

Wang, Q., Cheng, X. L., Zhang, D. Y., Gao, X. J. Zhou, L., Qin,X. Y., et al. (2013). Tectorigenin attenuates palmitate-induced endothelial insulin resistance via targeting ROS-associated inflammation and IRS-1 pathway. *PLoS ONE* 8:e66417. doi: 10.1371/journal.pone.0066417

Wang, X., Fan, R. Y., Li, J., Li, C. F., and Zhang, Y. S. (2016). Molecular cloning and functional characterization of a novel (Iso)flavone 4′,7-O-diglucoside glucosyltransferase from *Pueraria lobata*. *Front. Plant Sci.* 7:387. doi: 10.3389/fpls.2016.00387

Wang, X., Li, C. F., Zhou, C., Li, J., and Zhang, Y. S. (2017). Molecular characterization of the C-glucosylation for puerarin biosynthesis in *Pueraria lobata*. *Plant J.* 90, 535–546. doi: 10.1111/tpj.13510

Wang, X., Li, C. F., Zhou, Z. L., and Zhang, Y. S. (2019). Identification of three (Iso)flavonoid glucosyltransferases from *Pueraria lobate*. *Front. Plant Sci.* 10:28. doi: 10.3389/fpls.2019.00028

Wang, Z. L., Wang, S., Xu, Z., Li, M. W., Chen, K., Zhang, Y. Q., et al. (2019). Highly promiscuous flavonoid 3-O-glycosyltransferase from *Scutellaria baicalensis*. *Org. Lett.* 21, 2241–2245. doi: 10.1021/acs.orglett.9b00524

Wani, W., Masoodi, K. Z., Zaid, A., Wani, S. H., Shah, F., Meena, V. S., et al. (2018). Engineering plants for heavy metal stress tolerance. *Rend. Lincei.Sci. Fis.* 29, 709–723. doi: 10.1007/s12210-018-0702-y

Wilson, A. E., Wu, S., and Tian, L. (2019). PgUGT95B2 preferentially metabolizes flavones/flavonols and has evolved independently from flavone/flavonol UGTs identified in *Arabidopsis thaliana*. *Phytochemistry* 157, 184–193. doi: 10.1016/j.phytochem.2018.10.025

Wojtaszek, P., and Stobiecki, M. (1997). Differential secretion and accumulation of isoflavonoids in *Lupinus albus* in response to fungal elicitor and $CuCl_2$. *Plant Physiol Bioch* 35, 129–135.

Wozniak, D., Janda, B., Kapusta, I., Oleszek, W., and Matkowski, A. (2010). Antimutagenic and anti-oxidant activities of isoflavonoids from *Belamcanda chinensis* (L.) DC. *Mutat. Res. Gen. Toxicol. Environ. Mut.* 696, 148–153. doi: 10.1016/j.mrgentox.2010.01.004

Xie, G. Y., Zhu, Y., Shu, P., Qin, X. Y., Wu, G., Wang, Q., et al. (2014). Phenolic metabolite profiles and antioxidants assay of three Iridaceae medicinal plants for traditional Chinese medicine "She-gan" by on-line HPLC–DAD coupled with chemiluminescence (CL) and ESI-Q-TOF-MS/MS. *J. Pharmaceut. Biomed.* 98, 40–51. doi: 10.1016/j.jpba.2014.05.008

Zhang, C. Z., Wang, S. X., Zhang, Y., Chen, J. P., and Liang, X. M. (2005). *In vitro* estrogenic activities of Chinese medicinal plants traditionally used for

the management of menopausal symptoms. *J. Ethnopharmacol.* 98, 295–300. doi: 10.1016/j.jep.2005.01.033

Zhao, X. Q., Wang, P. Q., Li, M. Z., Wang, Y. R., Jiang, X. L., Cui, L. L., et al. (2017). Functional characterization of a new tea (*Camellia sinensis*) flavonoid glycosyltransferase. *J. Agric. Food Chem.* 65, 2074–2083. doi: 10.1021/acs.jafc.6b05619

Zhu, Y., Chen, Y., Zhang, X., Xie, G. Y., and Qin, M. J. (2020). Copper stress-induced changes in biomass accumulation, antioxidant activity and flavonoid contents in *Belamcanda chinensis* calli. *Plant Cell Tiss. Org.* 142, 299–311. doi: 10.1007/s11240-020-01863-w

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 17 March 2021
doi: 10.3389/fpls.2021.633333

The Light-Induced WD40-Repeat Transcription Factor DcTTG1 Regulates Anthocyanin Biosynthesis in *Dendrobium candidum*

Ning Jia[1,2], Jingjing Wang[1], Yajuan Wang[1,2], Wei Ye[3], Jiameng Liu[1,2], Jinlan Jiang[3], Jing Sun[1,2], Peipei Yan[3], Peiyu Wang[3], Fengzhong Wang[1,2] and Bei Fan[1,2]**

[1]*Institute of Food Science and Technology, Chinese Academy of Agricultural Sciences, Beijing, China,* [2]*Laboratory of Quality & Safety Risk Assessment on Agro-products Processing, Ministry of Agriculture and Rural Affairs, Beijing, China,* [3]*Institute of Medicinal Plant Sciences, Sanming Academy of Agricultural Sciences, Sanming, China*

OPEN ACCESS

Edited by:
Cristina Garcia-Viguera, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
Shenghui Jiang, Qingdao Agricultural University, China
Kate Warpeha, University of Illinois at Chicago, United States

***Correspondence:**
Bei Fan
fanbei517@163.com
Fengzhong Wang
wangfengzhong@sina.com

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

Received: *16 December 2020*
Accepted: *22 February 2021*
Published: *17 March 2021*

Citation:
Jia N, Wang J, Wang Y, Ye W, Liu J, Jiang J, Sun J, Yan P, Wang P, Wang F and Fan B (2021) The Light-Induced WD40-Repeat Transcription Factor DcTTG1 Regulates Anthocyanin Biosynthesis in Dendrobium candidum. Front. Plant Sci. 12:633333. doi: 10.3389/fpls.2021.633333

Dendrobium candidum is used as a traditional Chinese medicine and as a raw material in functional foods. *D. candidum* stems are green or red, and red stems are richer in anthocyanins. Light is an important environmental factor that induces anthocyanin accumulation in *D. candidum*. However, the underlying molecular mechanisms have not been fully unraveled. In this study, we exposed *D. candidum* seedlings to two different light intensities and found that strong light increased the anthocyanin content and the expression of genes involved in anthocyanin biosynthesis. Through transcriptome profiling and expression analysis, we identified a WD40-repeat transcription factor, DcTTG1, whose expression is induced by light. Yeast one-hybrid assays showed that DcTTG1 binds to the promoters of *DcCHS2*, *DcCHI*, *DcF3H*, and *DcF3′H*, and a transient GUS activity assay indicated that DcTTG1 can induce their expression. In addition, DcTTG1 complemented the anthocyanin deficiency phenotype of the *Arabidopsis thaliana ttg1-13* mutant. Collectively, our results suggest that light promotes anthocyanin accumulation in *D. candidum* seedlings *via* the upregulation of DcTTG1, which induces anthocyanin synthesis-related gene expression.

Keywords: *Dendrobium candidum*, DcTTG1, anthocyanin biosynthesis, transcriptional regulation, transcription factor

INTRODUCTION

Dendrobium candidum, a perennial herb of the orchid family, has been used for thousands of years in China as a traditional Chinese medicine and in functional foods (Ng et al., 2012), including teas, juices, wines, stews, and soups (Wei et al., 2016). *D. candidum* stems are green or red, with the latter having a higher anthocyanin content. Anthocyanins have a high nutritional value and are beneficial for human health (de Pascual-Teresa and Sanchez-Ballesta, 2008). Further, they have a broad spectrum of medicinal effects, including protection against liver injury, reduction of blood pressure, improvement of eyesight, and strong anti-inflammatory activity and anticancer properties (Konczak and Zhang, 2004). In addition, anthocyanins function to attract insect pollinators and protect plants against environmental stresses. However, the

gene transcription regulatory mechanism underlying color formation in *D. candidum* stems has not been fully unraveled.

Anthocyanin biosynthesis is catalyzed by a series of enzymes, including chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3′-hydroxylase (F3′H), flavonoid 3′,5′-hydroxylase (F3′5′H), dihydroflavonol 4-reductase (DFR), and anthocyanin synthase (ANS). These enzymes are encoded by structural genes, the expression of which is regulated by transcription factors (TFs; Broun, 2005). Key anthocyanin biosynthetic enzymes have been extensively studied (Zhang et al., 2014). Anthocyanin biosynthesis starts from the flavonoid synthesis pathway with the formation of chalcone from 4-coumaroyl CoA and malonyl CoA by chalcone synthase (CHS). Anthocyanidins undergo glucose-flavonoid 3-*O*-glucosyl transferase (UFGT) modification to increase their stability (Holton and Cornish, 1995; Zhang et al., 2014). Anthocyanin aglycones include delphinidin, cyanidin, pelargonidin, peonidin, malvidin, and petunidin (Cabrita et al., 2000).

Numerous TFs regulating plant anthocyanin biosynthesis have been identified to date. They include three main types: MYB TFs, basic helix-loop-helix (bHLH) factors, and WD40-repeat proteins. MYB-bHLH-WD40 (MBW) complexes play an important role in anthocyanin biosynthesis in various plant species (Xu et al., 2015). Environmental biotic and abiotic factors play important roles in plant secondary metabolism, including anthocyanin, flavonoid, and polysaccharide metabolism. In *Arabidopsis thaliana*, environmental stress factors, such as light, cold, and drought stresses and pathogen invasion, enhance anthocyanin accumulation (Chalker-Scott, 1999). Light is one of the most important environmental factors regulating anthocyanin biosynthesis in plants. Light-regulated anthocyanin synthesis is mostly achieved through the regulation of specific TFs (Allan et al., 2008; Bai et al., 2019; Ni et al., 2019). In many fruit-bearing plants, such as pear, apple, Chinese bayberry, litchi, and eggplant, anthocyanin accumulation involves the light-mediated TF R2R3 MYB (Takos et al., 2006; Niu et al., 2010; Lai et al., 2014; Jiang et al., 2016; Bai et al., 2017). The basic leucine zipper TF HY5, which is regulated by light, directly binds to the gene promoters of R2R3 MYB regulators of anthocyanin biosynthesis to regulate anthocyanin biosynthesis (Shin et al., 2013; Jiang et al., 2016; An et al., 2017; Qiu et al., 2019). In *Arabidopsis*, the expression of anthocyanin biosynthesis-related genes, including production of anthocyanin pigment 1 (*PAP1*), *PAP2*, *TT8*, *GL3*, and *EGL3*, is induced by light (Cominelli et al., 2008). In *Arabidopsis*, the WD40-repeat protein TTG1 forms a complex with bHLH (GL3, EGL3, or TT8) and R2R3 MYB (PAP1, PAP2, MYB113, and MYB114) TFs to regulate anthocyanin synthesis (Zhang et al., 2003; Gonzalez et al., 2008). Low temperature is another important environmental factor that affects anthocyanin biosynthesis. MdbHLH3 binds to the MYC-binding site in the *MdMYB1* promoter to induce anthocyanin accumulation in apple under low temperature (Xie et al., 2012).

The effects of environmental factors on the molecular mechanism of anthocyanin synthesis in *D. candidum* have not been fully elucidated. This study aimed to clarify the molecular mechanism of light-induced anthocyanin biosynthesis regulation in *D. candidum*. To this end, we grew plants under two different light intensities. The WD40-repeat TF DcTTG1 was found to respond to light signals and to regulate anthocyanin biosynthesis. These results provide new insights in the molecular basis of anthocyanin biosynthesis in *D. candidum*.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Dendrobium candidum (cultivar JXH-1) tissues were cultured at the Institute of Medicinal Plants, Sanming Academy of Agricultural Sciences (Sanming, Fujian, China). Plant tissues were cultured in Murashige & Skoog (MS) medium supplemented with 6-BA (2.0 mg/L), sucrose (20 g/L), and agar (6 g/L), pH 5.8, under a 16-h light/8-h dark cycle, at 24°C. For light treatments, 3-cm-high seedlings were transferred to ½ MS medium and cultured under a 16-h light/8-h dark cycle at 24°C. The light intensity of fluorescent lamps (NH-GP-1200R, Nonghui, Shanghai, China) was set to 1,000 lx or 4,000 lx. Seedlings were collected after 30 days of light exposure. Tobacco (*Nicotiana benthamiana*) and wild-type (WT) and *ttg1-13* mutant *A. thaliana* (ecotype Columbia) plants were grown in a mixed matrix of vermiculite/perlite/peat soil (1:1:1, v/v/v) or on MS medium in a greenhouse under a 16-h light/8-h dark cycle and a relative humidity of 40–60%, at 22°C.

Quantification of Anthocyanins

The total anthocyanin content of *Arabidopsis* leaves was extracted and determined using previously described methods (Zhao and Dixon, 2009; Zhao et al., 2011), using cyanidin 3-*O*-glucoside as a standard. Three biological replicates were analyzed.

Dendrobium candidum samples were ground into a fine powder in liquid nitrogen immediately after collection. One hundred milligram of the powder was extracted with 80% methanol aqueous solution containing 0.5% formic acid in a centrifuge tube under shaking for 2 min followed by ultrasonication at room temperature for 30 min. After the extraction, the mixture was centrifuged at 12000 rpm for 10 min. The supernatant liquid was transferred into a new tube and subjected to UPLC-MS/MS, using a 30A UPLC system with a phenylethyl chromatographic column, column temperature of 40°C, flow rate of 0.25 ml/min, and a total injection volume of 30 μl. Three biological replicates were analyzed for each sample.

Transcriptome Sequencing

RNA was isolated from plant tissue samples using an RNA isolation kit (Aidlab, Beijing, China). RNA quantity and integrity were assessed using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, United States) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, United States), respectively. cDNA libraries were constructed with the NEB Next® Ultra™ RNA Library Prep Kit (New England Biolabs, Ipswich, MA, United States). The cDNA libraries were sequenced on the HiSeq4000 platform (Illumina, San Diego, CA, United States) at Biomarker Technology (Beijing, China). Three biological replicates were analyzed for each sample

type. Gene expression levels were analyzed according to the "fragments per kilobase of transcript per million fragments mapped" method. The DESeq R package was used to identify differentially expressed genes (DEGs) in RDc4000 vs. RDc1000. Gene Ontology (GO) enrichment of the DEGs was analyzed using the GOseq R package (Young et al., 2010). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment was investigated using the KOBAS tool (Xie et al., 2012). The sequencing data have been submitted to NCBI Sequence Read Archive database under accession numbers SRR13577108, SRR13577107, SRR13577106, SRR13577105, SRR13577104, and SRR13577103.[1]

Phylogenetic Analysis and Subcellular Localization

Sequences of TTG1-like proteins were downloaded from the National Center for Biotechnology Information database. MEGA X was used for phylogenetic analysis of the TTG1-like proteins. *Dendrobium candidum* DNA was extracted using a DNA isolation kit (Aidlab). Using primers listed in **Supplementary Table S1**, a *TTG1* sequence was cloned and a fusion vector DcTTG1::YFP was constructed. DcTTG1::YFP was transformed into *Agrobacterium* GV3101 for injection into tobacco (*N. benthamiana*) leaf epidermal cells. 4,6-Diamidino-2-phenylindole (DAPI) was used to stain the nuclei. Fluorescence signals were acquired using a confocal laser scanning microscope (model SP8; Leica Microsystems).

RNA Extraction, cDNA Synthesis, and Quantitative Reverse-Transcription PCR

RNA was extracted from *D. candidum* tissue samples using an RNA isolation kit (Aidlab). The RevertAid Premium First Strand cDNA Synthesis Kit (Fermentas/Thermo Fisher Scientific, Rochester, NY, United States) was used for cDNA synthesis. Quantitative reverse transcription PCRs (RT-qPCRs) were run on a 7500 Fast Quantitative PCR system (Applied Biosystems) using TB Green® *Premix Ex Taq*™ (Takara). *DcACT* was selected as a reference gene to normalize target gene expression according to the $2^{-\Delta\Delta CT}$ method. Three biological replicates were analyzed. The primers used for RT-qPCR are listed in **Supplementary Table S2**.

Y1H Assay

The coding region of *DcTTG1* was ligated into the pGADT7 vector and promoter fragments of several target genes related to anthocyanin biosynthesis were ligated into the pHIS2 vector (Clontech, Palo Alto, CA, United States). Fusion products of the pGADT7 and pHIS2 vectors were transformed into yeast strain Y187 cells (Clontech, Palo Alto, CA, United States). The cells were cultured in synthetic defined (SD) medium –Leu/–Trp/–His with different concentrations of 3-amino-1,2,4-triazole (3-AT, mM) to prevent background histidine leakage of the pHIS2 vector at 28°C for 3 days. Empty pGADT7 vector was used as a control. The primers used for vector construction are listed in **Supplementary Table S3**.

[1]https://ncbiinsights.ncbi.nlm.nih.gov/tag/sra/

Transcriptional Activity Assay

Transcriptional activity was analyzed according to previous reports (Liu et al., 2016; Jia et al., 2018). The coding sequence of *DcTTG1* was fused into the 35S-LUC-GUS vector driven by the 35S promoter. Promoter fragments of *DcCHs1*, *DcCHS2*, *DcCHI*, *DcF3H*, *DcF3'H*, *DcDFR1*, *DcDFR2*, *DcANS1*, and *DcUFGT* were cloned into the 35S-LUC-GUS vector driven by the GUS reporter. The internal standard was LUC. The fusion vectors were transformed into *Agrobacterium* GV3101 cells. Single clones were cultured in liquid medium (10 mM $MgCl_2$, 10 mM MES, and 100 μM acetosyringone) under shaking until an optical density at 600 nm of 0.5 was reached. Then, the *Agrobacterium* cells were injected into tobacco leaves. GUS and LUC activities were measured after 4 days. At least five biological replicates were included in each assay. The primers used for vector construction are listed in **Supplementary Table S4**.

RESULTS

Red Pigmentation in RDc4000 Is Caused by Anthocyanin Accumulation

The phenotypes of *D. candidum* seedlings grown under 1,000 lx (RDc1000) and 4,000 lx (RDc4000) are shown in **Figure 1A**. RDc1000 stems had red pigmentation. The stems of RDc4000 were darker than those of RDc1000, and the leaves were dark red. In general, RDc4000 had more red pigment than RDc1000, indicating that higher-intensity light stimulates the accumulation of red pigment. Analysis of the anthocyanin contents to reveal the type and quantity of red pigmentation under the two light conditions revealed that *D. candidum* contained delphindin, delphindin-3-*O*-B-D-glucoside, delphindin-3,5-diglucoside, cyanin, and cyanidin under both conditions, but the contents of these five compounds were different. Delphindin-3-*O*-B-D-glucoside was the most abundant in RDc1000 and delphindin-3,5-diglucoside was the most abundant in RDc4000; however, the content of delphindin did not differ between the two light conditions. Notably, the contents of delphindin-3-*O*-B-D-glucoside, delphindin-3,5-diglucoside, cyanin, and cyanidin were significantly higher in RDc4000 than in RDc1000. These results indicated that a higher light intensity promotes the production of red pigment, mainly of the anthocyanin compounds delphindin-3-*O*-B-D-glucoside, delphindin-3,5-diglucoside, cyanin, and cyanidin, in *D. candidum*.

Transcriptome Analysis of RDc1000 and RDc4000

Transcript-level differences in *D. candidum* under the different light intensities were analyzed by transcriptome sequencing of RDc1000 and RDc4000. The original sequencing data were transformed into raw reads *via* base calling. The total number of bases in the six libraries was 43.62 G, and the total number of bases in each sample was greater than 6.5 Gb. The Q20 and Q30 scores for each sample were above 97 and 94%, and the GC content was relatively consistent, around 46% (**Supplementary Table S5**). The mapping rate, i.e., the percentage of mapped reads in clean reads, the most directly reflects

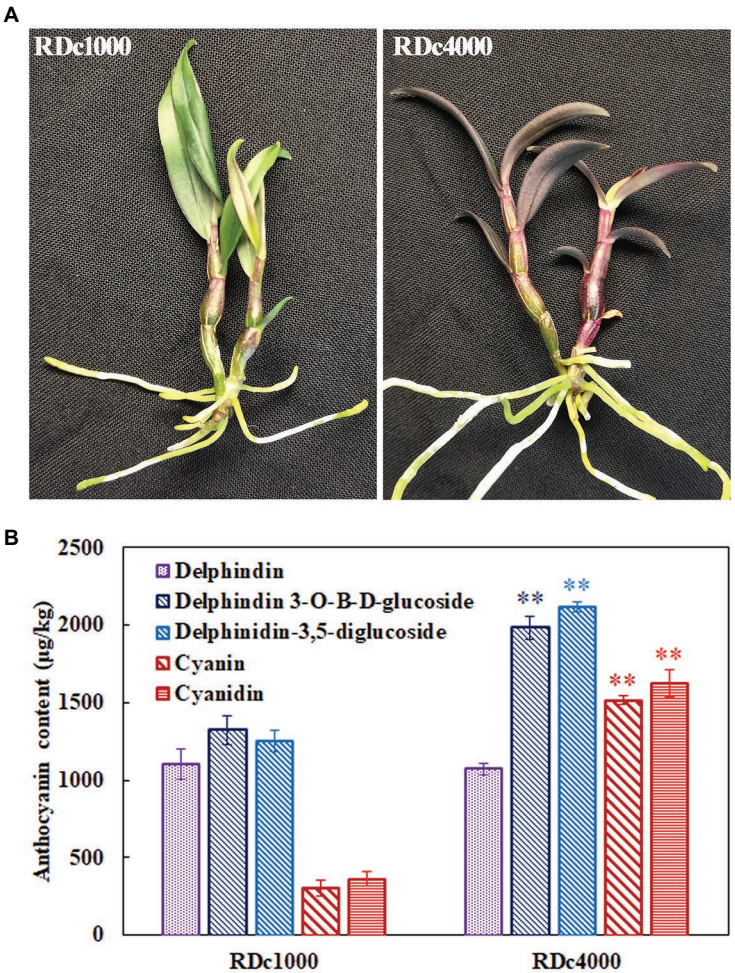

FIGURE 1 | Photographs and anthocyanin contents of *Dendrobium candidum* under two different light treatments. **(A)** Photographs of RDc1000 (red *D. candidum* seedlings under 1,000 lx) and RDc4000 (red *D. candidum* seedlings under 4,000 lx). **(B)** Anthocyanin composition in RDc1000 and RDc4000, including delphindin, delphindin-3-*O*-B-d-glucoside, delphindin-3,5-diglucoside, cyanin, and cyanidin, as assessed by UPLC-MS/MS. All data represent the mean ± SE of three biological replicates. $^{**}p < 0.01$, Student's *t*-test.

sequencing data utility. The mapping rate of the samples ranged from 90.5 to 91.08%. The percentage of uniquely and multi-mapped reads in clean reads ranged from 87.75 to 88.36% and from 2.38 to 2.74%, respectively (**Supplementary Table S6**). These findings indicated that the data met the quality requirements for analysis. In total, 130 unigenes could be assigned to 51 Gene Ontology (GO) terms in the cellular component, molecular function, and biological process categories. In the cellular component category, DEGs in RDc4000 vs. RDc1000 were mainly related to the plasma membrane, cell, cell parts, and organelles. In the molecular function category, most DEGs were related to catalytic activity and binding. In the biological process category, DEGs were mainly involved in metabolic processes, single-organism processes, and cellular processes (**Figure 2A**).

In total, 258 DEGs were found, 152 of which were upregulated and 106 of which were downregulated in RDc4000 vs. RDc1000. Among these, there were 22 TF genes, comprising eight upregulated and 14 downregulated genes, including *MYB* (two genes), *bHLH* (two genes), *WRKY* (two genes), *AP2/ERF* (two genes), *bZIP* (two genes), *NAC* (two genes), and the WD40-repeat TF gene, *DcTTG1* (**Supplementary Table S7**). Among the 22 differentially expressed TF genes, *DcTTG1* was upregulated the most strongly in RDc4000 vs. RDc1000 (**Supplementary Table S7**). Therefore, we speculated that DcTTG1 may be involved in the synthesis of anthocyanins. To identify the metabolic pathways involved in light-induced anthocyanin biosynthesis, 60 DEGs were subjected to KEGG pathway enrichment analysis, and the 20 most enriched pathways are shown in **Figure 2B**. In these pathways, we found eight DEGs involved in flavonoid biosynthesis, including genes encoding dihydroflavonol-4-reductase, naringenin, 2-oxoglutarate 3-dioxygenase, flavanone 3-dioxygenase, and leucoanthocyanidin dioxygenase, which are key genes in anthocyanin biosynthesis (**Supplementary Table S8**). These results indicated that these eight genes are the most important genes in the pathway of light-induced anthocyanin biosynthesis.

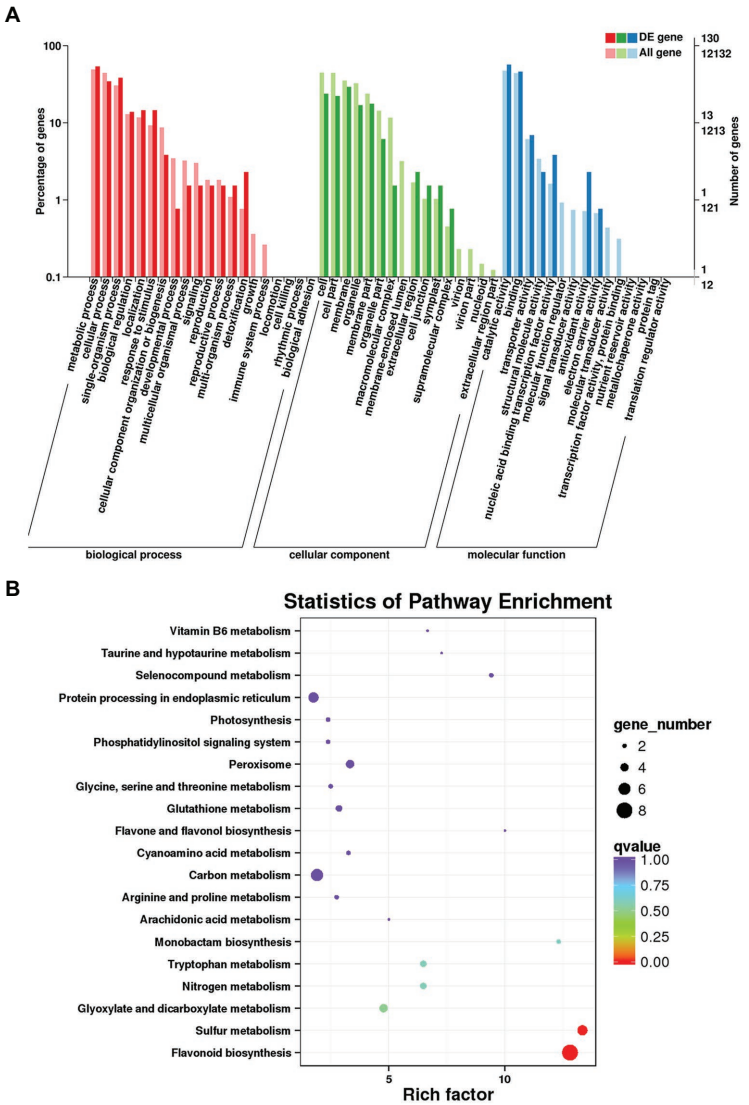

FIGURE 2 | Transcriptome analysis of *D. candidum* stems. **(A)** Gene Ontology (GO) enrichment analysis results for differentially expressed genes (DEGs) between RDc1000 and RDc4000. **(B)** Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of the DEGs.

Expression Analysis and Characterization of DcTTG1

To better understand the role of this TF in anthocyanin production in *D. candidum* under the different light regimens, it was studied in depth. The *DcTTG1* sequence was cloned from *D. candidum* cDNA. The relative expression of *DcTTG1* in RDc1000 and RDc4000 was analyzed by RT-qPCR. The results are shown in **Figure 3A**. *DcTTG1* expression in RDc4000 was 14 times higher than that in RDc1000, indicating that it is induced with increasing light intensity (**Figure 3A**). To study its regulation in anthocyanin biosynthesis, we conducted subcellular localization analysis using a Pro35S::DcTTG1-YFP vector construct that was transfected into *Agrobacterium* cells, which were injected into tobacco leaf epidermal cells. The fluorescence signal of Pro35S::DcTTG1-YFP was observed in the nucleus and overlapped with that of the nuclear dye, DAPI, indicating that DcTTG1 is located in the nucleus in *D. candidum* (**Figure 3B**). Next, we conducted phylogenetic analysis of DcTTG1 and TTG1-like proteins of other species (**Figure 3C**). DcTTG1 and PeTTG1 from *Phalaenopsis equestris* were in the same branch and were closely related. Together with the transcriptional expression differences after the different light treatments, this suggested that DcTTG1 plays an important role in light-regulated anthocyanin biosynthesis.

Expression Analysis of Key Anthocyanin Biosynthetic Genes

Anthocyanin accumulation affects plant pigmentation, resulting in different plant colors. To determine the effect of the different light intensity regimens on anthocyanin metabolism, the

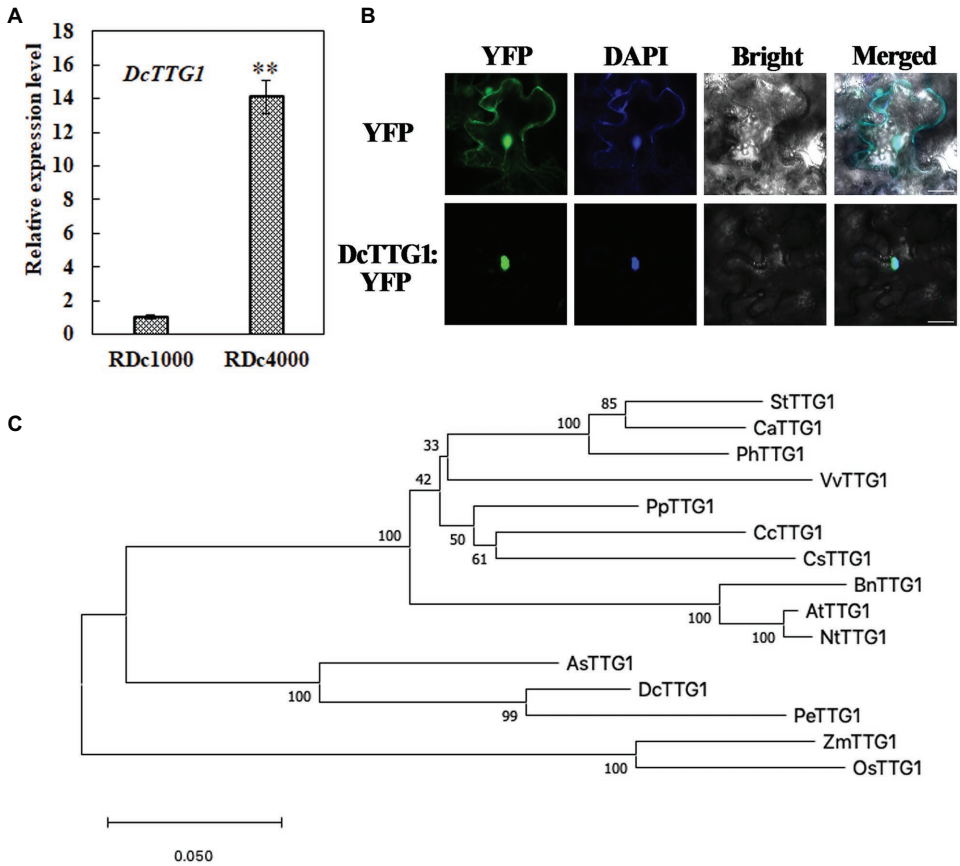

FIGURE 3 | Phylogenetic, expression, and subcellular localization analyses of DcTTG1. **(A)** Relative expression of *DcTTG1* in RDc1000 and RDc4000 as determined by quantitative reverse transcription PCR (RT-qPCR). DcACT was used as reference gene for normalization. Data represent the mean ± SE of three biological replicates. $^{**}p < 0.01$, Student's *t*-test. **(B)** Subcellular localization of DcTTG1. 4,6-Diamidino-2-phenylindole (DAPI) was used as a nuclear marker. The merged image shows the colocalization of the YFP fluorescence signal and the DAPI fluorescence signal. Bar = 20 μm. **(C)** Phylogenetic analysis of TTG1-like proteins. The following proteins were used to construct the phylogenetic tree: StTTG1 (NP_001305551.1), CaTTG1 (XP_016564215.1), PhTTG1 (AAC18914.1), VvTTG1 (CAN67365.1), PpTTG1 (ACQ65867.1), CcTTG1 (AMQ26245.1), CsTTG1 (NP_001306987.1), BnTTG1 (NP_001303154.1), AtTTG1 (CAC10524.1), NtTTG1 (ACJ06978.1), AsTTG1 (PKA51013.1), PeTTG1 (XP_020583423.1), ZmTTG1 (NP_001310302.1), and OsTTG1 (KAB8088430.1).

expression of 10 key anthocyanin biosynthetic genes, including *DcCHS1*, *DcCHS2*, *DcCHI*, *DcF3H*, *DcF3'H*, *DcF3'5'H*, *DcDFR1*, *DcDFR2*, *DcANS1*, and *DcUFGT*, was analyzed by RT-qPCR in RDc1000 and RDc4000 plants. The mRNA levels of *DcCHS1*, *DcCHS2*, *DcF3H*, *DcF3'H*, *DcDFR1*, *DcDFR2*, and *DcANS1* were significantly higher in RDc4000 than in RDc1000. *DcDFR1* was the most strongly induced, by approximately 27 times, followed by *DcF3H*. These results indicated that, the stronger the light intensity, the higher the expression of DFR, which greatly promotes anthocyanin accumulation in plants. These seven genes included early as well as late anthocyanin biosynthetic genes, indicating that high-intensity light has a significant effect on both the early and late stages of anthocyanin synthesis.

DcTTG1 May Regulate the Expression of *DcCHS2*, *DcCHI1*, *DcF3H*, and *DcF3'H* by Binding to Their Promoters

TTG1 is a WD40-repeat TF that can bind to the promoter regions of target genes to regulate their transcription. A Y1H assay was used to analyze the regulation of target genes by DcTTG1. The promoter sequences of nine anthocyanin biosynthetic genes, *DcCHs1*, *DcCHS2*, *DcCHI*, *DcF3H*, *DcF3'H*, *DcDFR1*, *DcDFR2*, *DcANS1*, and *DcUFGT*, were cloned separately to verify whether DcTTG1 has a regulatory effect on these genes. An appropriate concentration of 3-AT in the yeast growth medium (SD/–Leu/–Trp/–His) inhibited the self-activation of all target genes. The assay results suggested that DcTTG1 can bind to the promoters of *DcCHS2*, *DcCHI*, *DcF3H*, and *DcF3'H* (**Figure 4**).

To verify the Y1H assay results, we used LUC/GUS analysis to detect the effect of DcTTG1 on the activation of key anthocyanin biosynthetic gene promoters. We cloned anthocyanin biosynthetic gene promoters, which we used to drive the GUS reporter gene, and DcTTG1 was driven by the CaMV35S promoter (**Figure 5A**). Transient expression of the reporter system in tobacco revealed that DcTTG1 activated reporter gene expression driven by the promoters of *DcCHS2*, *DcCHI*, *DcF3H*, and *DcF3'H* (**Figure 5B**), corroborating that DcTTG1 regulates the expression of these four genes. Thus, DcTTG1 modulates anthocyanin

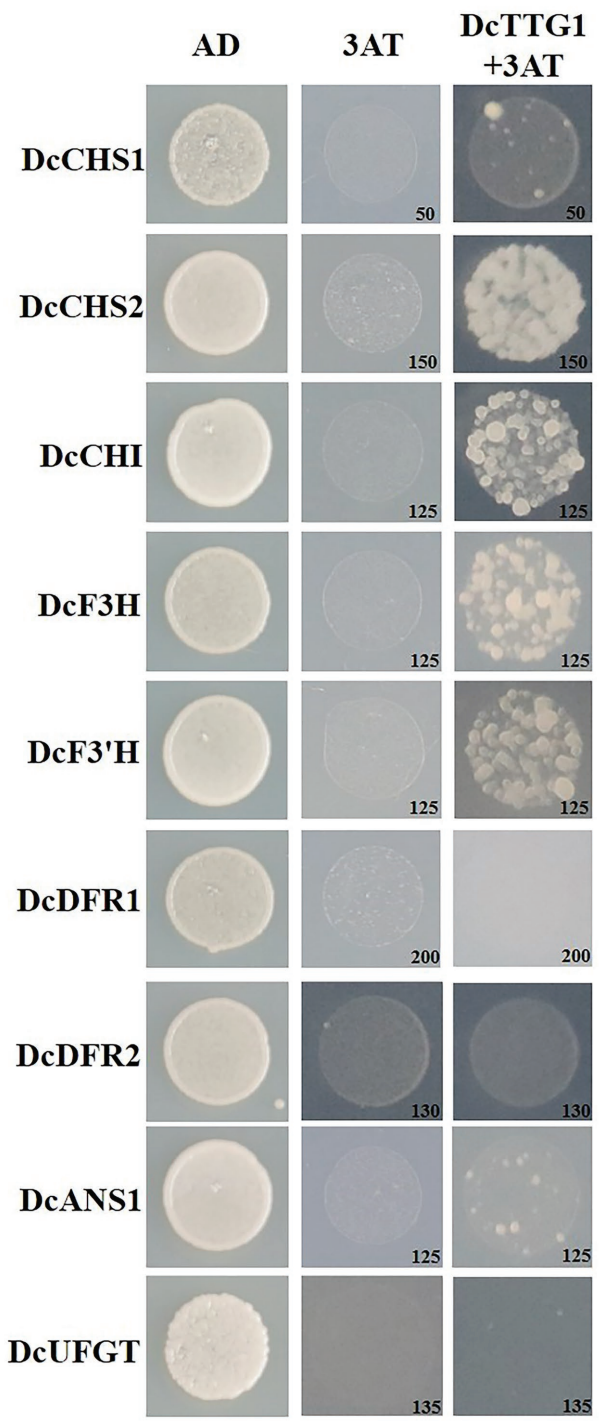

FIGURE 4 | Y1H assays of DcTTG1 and anthocyanin biosynthesis-related gene promoters. The *DcTTG1* sequence and promoter fragments of nine anthocyanin biosynthetic genes were cloned into pGADT7 and pHIS2 vectors, respectively. The plasmids were transformed into Y187 yeast cells (Clontech Laboratories, Mountain View, CA, United States) that were grown on SD (–Leu/–Trp/–His) medium at 28°C for 3 days. The optimal 3-AT concentration (in mM) to suppress background histidine leakiness after addition to the (–Leu/–Trp/–His) medium is indicated in the lower-right corner of each picture.

synthesis by regulating the expression of anthocyanin synthesis-related target genes under different light intensities.

Phenotype Rescue of the *Arabidopsis ttg1-13* Mutant

Establishing transgenic *D. candidum* lines is generally difficult and time-consuming. Therefore, we used the model plant *Arabidopsis* to verify the function of DcTTG1 in the regulation of anthocyanin biosynthesis. To further study the possible function of *DcTTG1* as a regulator of anthocyanin biosynthesis, we overexpressed *DcTTG1* driven by the CaMV35S promoter in the *Arabidopsis ttg1-13* mutant, which is defect in seed coat anthocyanin pigmentation. We obtained 18 *DcTTG1* overexpression lines and analyzed lines *DcTTG1/ttg1-13#1* and *DcTTG1/ttg1-13#2*. The *DcTTG1* expression levels in these lines were significantly higher than those in WT plants (**Figure 6A**), indicating successful overexpression. *DcTTG1* overexpression restored the *ttg1-13* mutant phenotype (**Figure 6B**; **Supplementary Figure S1**) as well as leaf anthocyanin pigmentation (**Figure 6C**), suggesting that DcTTG1 has a similar function as *Arabidopsis* TTG1.

DISCUSSION

This study revealed that the contents of delphindin-3-*O*-B-D-glucoside, delphindin-3,5-diglucoside, cyanin, and cyanidin were significantly increased in RDc4000 compared with RDc1000. Transcriptome sequencing and RT-qPCR showed that the WD40-repeat TF DcTTG1 was significantly more strongly expressed in RDc4000 than in RDc1000. Phylogenetic analysis and genetic complementation revealed the functional conservation of DcTTG1 and its orthologs in other plant species. Y1H and transcriptional activation assays indicated that DcTTG1 regulates *DcCHS2*, *DcCHI*, *DcF3H*, and *DcF3'H* expression by binding to their promoters. Thus, DcTTG1 may regulate anthocyanin biosynthesis in *D. candidum* under different light intensities.

Anthocyanin compositional and content analyses showed that the content of delphindin did not significantly differ between RDc1000 and RDc4000 (**Figure 1B**). In the anthocyanin biosynthetic pathway, the formation of delphindin requires F3'5'H catalysis (Matus et al., 2009). *DcF3'5'H* expression did not significantly differ between RDc1000 and RDc4000, which may explain the lack of a significant difference in delphindin content. However, the contents of delphindin-3-*O*-B-D-glucoside and delphindin-3,5-diglucoside were significantly different between RDc1000 and RDc4000, which may be due to significant increases in *DcDFR1*, *DcDFR2*, and *DcANS1* expression in RDc4000 (**Figure 7**). The cyanin and cyanidin contents were significantly higher in RDc4000 than in RDc1000 (**Figure 7**). The formation of cyanin and cyanidin requires F3'H catalysis (Rinaldo et al., 2015), and DcF3'H expression was significantly higher in RDc4000 than in RDc1000 (**Figure 7**). Light provides energy for plant growth and development (Lau and Deng, 2010). It regulates plant secondary metabolism, including anthocyanin and flavonoid synthesis (Zhang et al., 2018; An et al., 2020). Different light intensities have different effects on anthocyanin accumulation, which is positively correlated with light intensity in many plant species (Jaakola, 2013; Zhang et al., 2018).

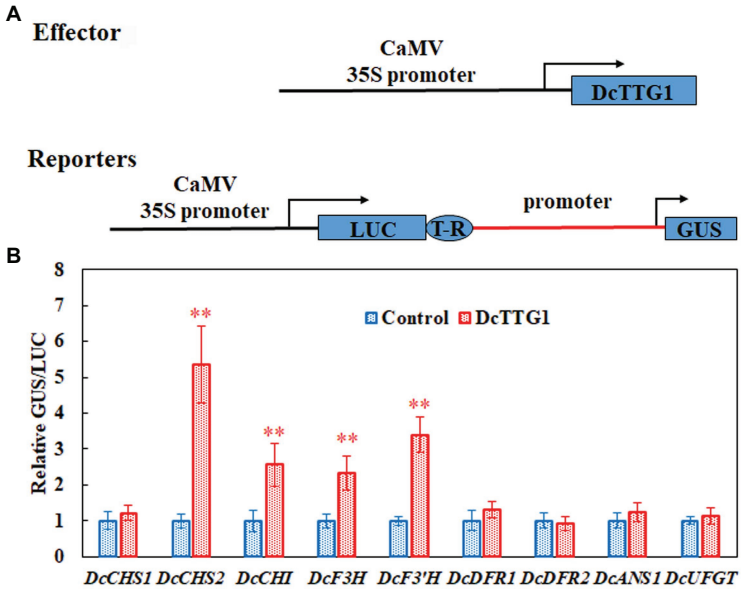

FIGURE 5 | DcTTG1 can directly bind to the promoters of anthocyanin biosynthetic genes. **(A)** Effector-reporter systems were used to determine the ability of DcTT1 to activate the promoters of anthocyanin biosynthetic genes in tobacco leaves. The reporter systems contained the *GUS* reporter gene driven by the promoters of anthocyanin biosynthetic genes, as well as *LUC* driven by CaMV35S for standardization. The effector systems contained *DcTTG1* driven by CaMV35S. Boxes, various DNA sequences. T-R, terminator. **(B)** Transcriptional activity of DcTTG1 on the promoters of anthocyanin biosynthetic genes as indicated by the reporter assay. Empty vector was used as a control, and the GUS/LUC ratio in the transformed leaves was set to 1. Data represent the mean ± SE of five biological replicates. $^{**}p < 0.01$, Student's *t*-test.

In *Arabidopsis*, *AtTTG1* is expressed in the dark, but its expression increases immediately after illumination and then remains relatively stable over time (Cominelli et al., 2008). In our study, *DcTTG1* expression increased significantly with increasing light intensity (**Figure 3A**). In Brassica, anthocyanins accumulate more strongly under LED irradiance with a photosynthetic photon flux of 220 μmol m^{-2} s^{-1} than under a flux of 330 or 440 μmol m^{-2} s^{-1} (Samuoliene et al., 2013). Within a certain range, anthocyanin accumulation gradually increases with increasing light intensity, but after a certain intensity, the amount of anthocyanin does not further increase due to photoprotection (Logan et al., 2015). Anthocyanin levels in lychee peels decreased rapidly after shading treatment, but anthocyanin biosynthesis resumed soon after the fruits were exposed to light again (Zhang et al., 2016). This shows that light can induce anthocyanin formation and is an important environmental factor affecting anthocyanin biosynthesis. Similar findings have been reported for Chinese bayberry (Niu et al., 2010), apple (Li et al., 2012), and grape berry (Azuma et al., 2012).

Transcriptome sequencing analysis was performed on RDc1000 and RDc4000 to study the relationship between light intensity and anthocyanin accumulation. Among the TFs identified by transcriptome sequencing, MYB, bHLH, and WD40 all showed significant differences in expression, indicating that MBW may be involved in anthocyanin biosynthesis induced by high light intensity in *D. candidum*. Transcriptome analysis of light-regulated anthocyanin biosynthesis in the pericarp of Litchi identified 76 TFs that regulate anthocyanin synthesis in response to light. Among these TFs, SQUAMOSA promoter binding protein-like constituted the largest population (21.1%), followed by MYB (19.7%), homologous domain leucine zipper protein (ATHB, 17.1%), and WD40 (13.2%; Zhang et al., 2016). Thus, light intensity is an important environmental factor in regulating anthocyanin biosynthesis.

The expression levels of *DcCHI*, *DcF3'5'H*, and *DcUFGT* did not differ significantly between the two light treatments, indicating that these genes are less responsive to different light intensities in *D. candidum*. Interestingly, our and previous findings indicate that the feedback regulation of anthocyanin synthetic genes in response to light varies among plant species. Leaves of *Lactuca sativa* L. var. capitata were green under 40 mmol m^{-2} s^{-1}, but red under 100 mmol m^{-2} s^{-1}, and the expression of *LsCHI* was upregulated (Zhang et al., 2018). However, in a study of light-induced regulation of anthocyanin biosynthesis in litchi fruit, the activation and expression of *LcUFGT* promoted the accumulation of anthocyanins and enhanced fruit redness (Zhang et al., 2016). Anthocyanin biosynthetic genes of different species differentially respond to light, which emphasizes the importance of studying the effects of different light treatments on anthocyanin synthesis in *D. candidum*.

DcTTG1 modulates anthocyanin synthesis by regulating the expression of anthocyanin synthesis-related target genes under different light intensities. In this study, no direct binding of TTG1 to late anthocyanin biosynthetic gene promoters was observed, which indicates that DcTTG1 has multiple modes of regulating key genes in anthocyanin biosynthesis. Gene regulatory networks in higher plants require the coordinated action of many gene interactions. Among them, the most widely studied is the regulation of anthocyanin biosynthesis by the MBW complex. TTG1 forms complexes with the R2R3 MYB TFs PAP1, PAP2,

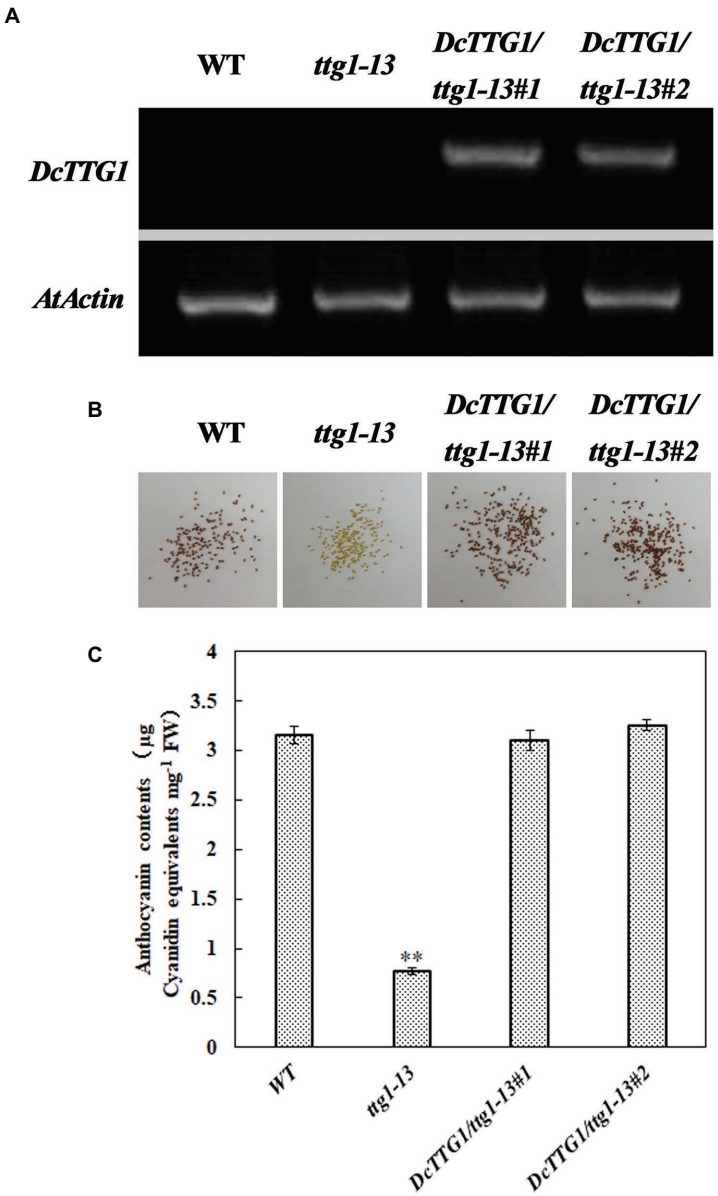

FIGURE 6 | Genetic complementation of the *Arabidopsis ttg1-13* mutant by DcTTG1. **(A)** *DcTTG1* transcripts in *Arabidopsis* young seeds. **(B)** Restored seed pigments in *ttg1-13* after overexpression of *DcTTG1*. **(C)** Quantification of anthocyanins in the leaves of WT, *ttg1*, *DcTTG1/ttg1-13#1*, and *DcTTG1/ttg1-13#2*. Data represent the mean ± SE of three biological replicates. $^{**}p < 0.01$, Student's *t*-test.

MYB113, MYB114, or TT2 and the bHLH TFs GL3, EGL3, or TT8, to form MBW (Xu et al., 2015). Research in *Arabidopsis* has shown that the TTG1-dependent MBW complex can directly bind to the promoters of *TTG2*, *TT8*, *F3'H*, *DFR*, *ANS*, *UGT79B1*, *UGT75C1*, *5MAT*, and *BLT* (Wei et al., 2019). Therefore, it can be inferred that DcTTG1 alone cannot regulate the late anthocyanin biosynthetic genes, whereas different TTG1-dependent MBW complexes can simultaneously regulate the early and late anthocyanin biosynthetic genes. The different MBW complexes formed by TTG1 have diverse regulatory roles in anthocyanin biosynthesis. The TTG1-TT8/GL3-PAP1/PAP2/MYB113/MYB114 complex can activate the expression of late biosynthetic genes, including dihydroflavonol 4-reductase (*DFR*) and phthalate synthase (*ANS*), to affect anthocyanin biosynthesis (Shi and Xie, 2014). There is no relevant research on the characteristic roles of DcTTG1 and the DcTTG1-dependent MBW complex in the regulation of anthocyanin biosynthetic structural gene expression, especially under different light conditions. Our findings may contribute to improving plant production through the regulation of certain desired metabolic compounds.

In conclusion, we identified and isolated a WD40-repeat TF, DcTTG1, as a novel regulator of light-regulated anthocyanin

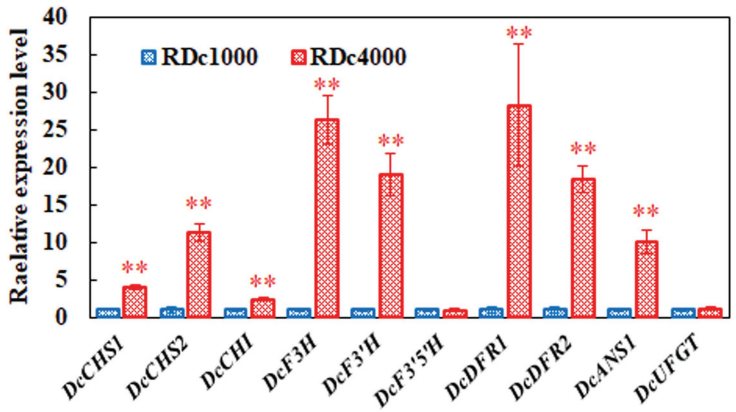

FIGURE 7 | Expression of key anthocyanin biosynthetic genes. *DcACTIN* was used as reference gene for normalization. All gene expression levels in RDc1000 were set to 1. Error bars represent the mean ± SE of three biological replicates. $^{**}p < 0.01$, Student's *t*-test.

accumulation in *D. candidum*. DcTTG1 regulates anthocyanin biosynthesis in various ways. The TF binds to the promoters of *DcCHS2*, *DcCHI*, *DcF3H*, and *DcF3'H*, which are anthocyanin biosynthesis-related genes. DcTTG1 restored the *Arabidopsis ttg1-13* mutant phenotype, which is defect in seed coat anthocyanin pigmentation, suggesting that it has a similar function as *Arabidopsis* TTG1. We believe that the findings of this study will help improve the understanding of the mechanism of light-induced anthocyanin synthesis and accumulation.

DATA AVAILABILITY STATEMENT

The transcriptome data of this study have been uploaded to NCBI Sequence Read Archive database (https://ncbiinsights.ncbi.nlm.nih.gov/tag/sra/), the accession number is PRJNA694810.

AUTHOR CONTRIBUTIONS

NJ, FW, and BF designed the study. NJ, JW, YW, JL, JJ, JS, PY, and PW performed the experiments. NJ, JW, FW, and BF wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the special fund for Supervision of Quality and Safety of Agricultural Products (Risk Assessment) from the Chinese Ministry of Agriculture (grant GJFP2019019).

ACKNOWLEDGMENTS

We thank Qinhuangdao General Spectrometry Technology Co. Ltd. for providing testing technical services (anthocyanins).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.633333/full#supplementary-material

REFERENCES

Allan, A. C., Hellens, R. P., and Laing, W. A. (2008). MYB transcription factors that colour our fruit. *Trends Plant Sci.* 13, 99–102. doi: 10.1016/j.tplants.2007.11.012

An, J. P., Liu, Y. J., Zhang, X. W., Bi, S. Q., Wang, X. F., You, C. X., et al. (2020). Dynamic regulation of anthocyanin biosynthesis at different light intensities by the BT2-TCP46-MYB1 module in apple. *J. Exp. Bot.* 71, 3094–3109. doi: 10.1093/jxb/eraa056

An, J. P., Qu, F. J., Yao, J. F., Wang, X. N., You, C. X., Wang, X. F., et al. (2017). The bZIP transcription factor MdHY5 regulates anthocyanin accumulation and nitrate assimilation in apple. *Hortic. Res.* 4:17023. doi: 10.1038/hortres.2017.23

Azuma, A., Yakushiji, H., Koshita, Y., and Kobayashi, S. (2012). Flavonoid biosynthesis-related genes in grape skin are differentially regulated by temperature and light conditions. *Planta* 236, 1067–1080. doi: 10.1007/s00425-012-1650-x

Bai, S., Sun, Y., Qian, M., Yang, F., Ni, J., Tao, R., et al. (2017). Transcriptome analysis of bagging-treated red Chinese sand pear peels reveals light-responsive pathway functions in anthocyanin accumulation. *Sci. Rep.* 7:63. doi: 10.1038/s41598-017-00069-z

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

Broun, P. (2005). Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in *Arabidopsis*. *Curr. Opin. Plant Biol.* 8, 272–279. doi: 10.1016/j.pbi.2005.03.006

Cabrita, L., Fossen, T., and Andersen, Ø. M. (2000). Colour and stability of the six common anthocyanidin 3-glucosides in aqueous solutions. *Food Chem.* 68, 101–107. doi: 10.1016/S0308-8146(99)00170-3

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress responses. *Photochem. Photobiol.* 70, 1–9. doi: 10.1111/j.1751-1097.1999.tb01944.x

Cominelli, E., Gusmaroli, G., Allegra, D., Galbiati, M., Wade, H. K., Jenkins, G. I., et al. (2008). Expression analysis of anthocyanin regulatory genes in response to different light qualities in *Arabidopsis thaliana*. *J. Plant Physiol.* 165, 886–894. doi: 10.1016/j.jplph.2007.06.010

de Pascual-Teresa, S., and Sanchez-Ballesta, M. T. (2008). Anthocyanins: from plant to health. *Phytochem. Rev.* 7, 281–299. doi: 10.1007/s11101-007-9074-0

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

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7, 1071–1083. doi: 10.2307/3870058

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

Jia, N., Liu, J. Q., Sun, Y. F., Tan, P. H., Cao, H., Xie, Y. Y., et al. (2018). Citrus sinensis MYB transcription factors CsMYB330 and CsMYB308 regulate fruit juice sac lignification through fine-tuning expression of the Cs4CL1 gene. *Plant Sci.* 277, 334–343. doi: 10.1016/j.plantsci.2018.10.006

Jiang, M., Ren, L., Lian, H., Liu, Y., and Chen, H. (2016). Novel insight into the mechanism underlying light-controlled anthocyanin accumulation in eggplant (*Solanum melongena* L.). *Plant Sci.* 249, 46–58. doi: 10.1016/j.plantsci.2016.04.001

Konczak, I., and Zhang, W. (2004). Anthocyanins—more than nature's colours. *J. Biomed. Biotechnol.* 2004:307613. doi: 10.1155/S1110724304407013

Lai, B., Li, X. J., Hu, B., Qin, Y. H., Huang, X. M., Wang, H. C., et al. (2014). LcMYB1 is a key determinant of differential anthocyanin accumulation among genotypes, tissues, developmental phases and ABA and light stimuli in *Litchi chinensis*. *PLoS One* 9:e86293. doi: 10.1371/journal.pone.0086293

Lau, O. S., and Deng, X. W. (2010). Plant hormone signaling lightens up: integrators of light and hormones. *Curr. Opin. Plant Biol.* 13, 571–577. doi: 10.1016/j.pbi.2010.07.001

Li, Y. Y., Mao, K., Zhao, C., Zhao, X. Y., Zhang, H. L., Shu, H. R., et al. (2012). MdCOP1 ubiquitin E3 ligases interact with MdMYB1 to regulate light-induced anthocyanin biosynthesis and red fruit coloration in apple. *Plant Physiol.* 160, 1011–1022. doi: 10.1104/pp.112.199703

Liu, J. Q., Chen, X. J., Liang, X. X., Zhou, X. G., Yang, F., Liu, J., et al. (2016). Alternative splicing of rice WRKY62 and WRKY76 transcription factor genes in pathogen defense. *Plant Physiol.* 171, 1427–1442. doi: 10.1104/pp.15.01921

Logan, B. A., Stafstrom, W. C., Walsh, M. J. L., Reblin, J. S., and Gould, K. S. (2015). Examining the photoprotection hypothesis for adaxial foliar anthocyanin accumulation by revisiting comparisons of green- and red-leafed varieties of coleus (*Solenostemon scutellarioides*). *Photosynth. Res.* 124, 267–274. doi: 10.1007/s11120-015-0130-0

Matus, J. T., Loyola, R., Vega, A., Pena-Neira, A., Bordeu, E., Arce-Johnson, P., et al. (2009). Post-veraison sunlight exposure induces MYB-mediated transcriptional regulation of anthocyanin and flavonol synthesis in berry skins of *Vitis vinifera*. *J. Exp. Bot.* 60, 853–867. doi: 10.1093/jxb/ern336

Ng, T. B., Liu, J., Wong, J. H., Ye, X., Wing Sze, S. C., Tong, Y., et al. (2012). Review of research on *Dendrobium*, a prized folk medicine. *Appl. Microbiol. Biotechnol.* 93, 1795–1803. doi: 10.1007/s00253-011-3829-7

Ni, J. B., Bai, S. L., Zhao, Y., Qian, M. J., Tao, R. Y., Yin, L., et al. (2019). Ethylene response factors Pp4ERF24 and Pp12ERF96 regulate blue light-induced anthocyanin biosynthesis in "Red Zaosu" pear fruits by interacting with MYB114. *Plant Mol. Biol.* 99, 67–78. doi: 10.1007/s11103-018-0802-1

Niu, S. S., Xu, C. J., Zhang, W. S., Zhang, B., Li, X., Lin-Wang, K., et al. (2010). Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (*Myrica rubra*) fruit by a R2R3 MYB transcription factor. *Planta* 231, 887–899. doi: 10.1007/s00425-009-1095-z

Qiu, Z., Wang, H., Li, D., Yu, B., Hui, Q., Yan, S., et al. (2019). Identification of candidate HY5-dependent and -independent regulators of anthocyanin biosynthesis in tomato. *Plant Cell Physiol.* 60, 643–656. doi: 10.1093/pcp/pcy236

Rinaldo, A. R., Cavallini, E., Jia, Y., Moss, S. M. A., McDavid, D. A. J., 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

Samuoliene, G., Brazaityte, A., Jankauskiene, J., Virsile, A., Sirtautas, R., Novickovas, A., et al. (2013). LED irradiance level affects growth and nutritional quality of *Brassica* microgreens. *Cent. Eur. J. Biol.* 8, 1241–1249. doi: 10.2478/s11535-013-0246-1

Shi, M. Z., and Xie, D. Y. (2014). Biosynthesis and metabolic engineering of anthocyanins in *Arabidopsis thaliana*. *Recent Pat. Biotechnol.* 8, 47–60. doi: 10.2174/18722083076661312181235 38

Shin, D. H., Choi, M., Kim, K., Bang, G., Cho, M., Choi, S. B., et al. (2013). HY5 regulates anthocyanin biosynthesis by inducing the transcriptional activation of the MYB75/PAP1 transcription factor in *Arabidopsis*. *FEBS Lett.* 587, 1543–1547. doi: 10.1016/j.febslet.2013.03.037

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

Wei, Z., Cheng, Y., Zhou, C., Li, D., Gao, X., Zhang, S., et al. (2019). Genome-wide identification of direct targets of the TTG1–bHLH–MYB complex in regulating trichome formation and flavonoid accumulation in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 20:5014. doi: 10.3390/ijms20205014

Wei, W., Feng, L., Bao, W. -R., Ma, D. -L., Leung, C. -H., Nie, S. -P., et al. (2016). Structure characterization and immunomodulating effects of polysaccharides isolated from dendrobium officinale. *J. Agric. Food Chem.* 64, 881–889. doi: 10.1021/acs.jafc.5b05180

Xie, X. B., Li, S., Zhang, R. F., Zhao, J., Chen, Y. C., Zhao, Q., et al. (2012). The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. *Plant Cell Environ.* 35, 1884–1897. doi: 10.1111/j.1365-3040.2012.02523.x

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

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

Zhang, F., Gonzalez, A., Zhao, M., Payne, C. T., and Lloyd, A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. *Development* 130, 4859–4869. doi: 10.1242/dev.00681

Zhang, H. N., Li, W. C., Wang, H. C., Shi, S. Y., Shu, B., Liu, L. Q., et al. (2016). Transcriptome profiling of light-regulated anthocyanin biosynthesis in the pericarp of litchi. *Front. Plant Sci.* 7:963. doi: 10.3389/fpls.2016.00963

Zhang, Y., Xu, S., Cheng, Y., Peng, Z., and Han, J. (2018). Transcriptome profiling of anthocyanin-related genes reveals effects of light intensity on anthocyanin biosynthesis in red leaf lettuce. *PeerJ* 6:e4607. doi: 10.7717/peerj.4607

Zhao, J., and Dixon, R. A. (2009). MATE transporters facilitate vacuolar uptake of epicatechin 3′-*O*-glucoside for proanthocyanidin biosynthesis in *Medicago truncatula* and *Arabidopsis*. *Plant Cell* 21, 2323–2340. doi: 10.1105/tpc.109.067819

Zhao, J., Huhman, D., Shadle, G., He, X. Z., Sumner, L. W., Tang, Y., et al. (2011). MATE2 mediates vacuolar sequestration of flavonoid glycosides and glycoside malonates in *Medicago truncatula*. *Plant Cell* 23, 1536–1555. doi: 10.1105/tpc.110.080804

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.

frontiers in Plant Science

REVIEW
published: 06 April 2021
doi: 10.3389/fpls.2021.642581

Diversity of Chemical Structures and Biosynthesis of Polyphenols in Nut-Bearing Species

*Chaiwat Aneklaphakij[1,2], Tomoki Saigo[2], Mutsumi Watanabe[2], Thomas Naake[3], Alisdair R. Fernie[3], Somnuk Bunsupa[1], Veena Satitpatipan[1] and Takayuki Tohge[2]**

[1] Department of Pharmacognosy, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand, [2] Graduate School of Biological Science, Nara Institute of Science and Technology, Ikoma, Japan, [3] Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany

Nuts, such as peanut, almond, and chestnut, are valuable food crops for humans being important sources of fatty acids, vitamins, minerals, and polyphenols. Polyphenols, such as flavonoids, stilbenoids, and hydroxycinnamates, represent a group of plant-specialized (secondary) metabolites which are characterized as health-beneficial antioxidants within the human diet as well as physiological stress protectants within the plant. In food chemistry research, a multitude of polyphenols contained in culinary nuts have been studied leading to the identification of their chemical properties and bioactivities. Although functional elucidation of the biosynthetic genes of polyphenols in nut species is crucially important for crop improvement in the creation of higher-quality nuts and stress-tolerant cultivars, the chemical diversity of nut polyphenols and the key biosynthetic genes responsible for their production are still largely uncharacterized. However, current technical advances in whole-genome sequencing have facilitated that nut plant species became model plants for omics-based approaches. Here, we review the chemical diversity of seed polyphenols in majorly consumed nut species coupled to insights into their biological activities. Furthermore, we present an example of the annotation of key genes involved in polyphenolic biosynthesis in peanut using comparative genomics as a case study outlining how we are approaching omics-based approaches of the nut plant species.

Keywords: flavonoids, chemical diversity, nuts, comparative genomics, polyphenols, health benefits

OPEN ACCESS

Edited by:
Pedro Mena,
University of Parma, Italy

Reviewed by:
Wajid Waheed Bhat,
Michigan State University, United States
Joong-Hoon Ahn,
Konkuk University, South Korea

****Correspondence:***
Takayuki Tohge
tohge@bs.naist.jp

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

Received: *16 December 2020*
Accepted: *25 February 2021*
Published: *06 April 2021*

Citation:
Aneklaphakij C, Saigo T, Watanabe M, Naake T, Fernie AR, Bunsupa S, Satitpatipan V and Tohge T (2021) Diversity of Chemical Structures and Biosynthesis of Polyphenols in Nut-Bearing Species. Front. Plant Sci. 12:642581. doi: 10.3389/fpls.2021.642581

INTRODUCTION

Nuts, such as chestnut and hazelnut, are oil-rich seeds comprising of an edible fruit with a hard outer shell attached to a cupule. Additionally, drupe seeds such as almond, peanut, pistachio, walnut, macadamia, pecan, and cashew nut also contain a hard shell and are thus referred to as "nuts." Since nuts contain precious types of phytonutrients which exhibit beneficial health-promoting properties, they are regarded as one of the most valuable culinary crops (Ros, 2010; Bodoira and Maestri, 2020). Indeed, consuming nuts provides a rich source of nutritional components, including fatty acids, minerals, vitamins, proteins, and fibers (Ros, 2010; Chen C. Y. O. et al., 2019). In addition, nuts were also found as a rich source of "plant-specialized (secondary) metabolites" which are a vast array of bioactive compounds. Plant-specialized compounds tend to

act as stress protectants against biotic and abiotic stresses from the external environment (Tohge et al., 2013b, 2016; Kessler and Kalske, 2018; Pais et al., 2018; Scossa et al., 2019). These metabolites broadly correspond to important physiological and ecological functions, for example, attracting insects for pollination by volatiles and color; as antifeedants and allelochemicals against herbivores; for visible pigmentation; and for lightening from stress conditions, e.g., ultraviolet radiation, elicitors, temperature, and water deficiency (Bodoira and Maestri, 2020; Corso et al., 2020; Yuan and Grotewold, 2020).

Plant-specialized metabolites are allocated to three main categories, namely, terpenoids, alkaloids, and polyphenols, classified on their chemical core skeletons and biosynthetic pathways (Tohge et al., 2013a,b, 2014; Chen C. Y. O. et al., 2019; Bodoira and Maestri, 2020; Yuan and Grotewold, 2020). Polyphenols are ubiquitously present among fruits, vegetables, and seeds, including nuts, with multiple claims of these compounds presenting human health beneficial effects on human health (Lepiniec et al., 2006; Tohge and Fernie, 2017; Hano and Tungmunnithum, 2020; Alseekh et al., 2020a). Numerous research publications investigating polyphenols in nuts have been published (Bolling et al., 2011; Alasalvar and Bolling, 2015; Bodoira and Maestri, 2020). According to our current update, polyphenols in nut plant species are classified into six major groups, including phenols, flavonoids, tannins, stilbenoids, lignans, and coumarins (**Figure 1**). Phenols including phenolic acids (C6–C1 skeleton) and hydroxycinnamic acids (C6–C3) are the compounds containing an aromatic ring attached with at least one hydroxyl moiety (**Figure 1A**; Bodoira and Maestri, 2020; Hano and Tungmunnithum, 2020). Flavonoids (C6–C3–C6), one of the largest classes of specialized metabolites, are distributed extensively in the plant kingdom (**Figure 1B**; Tohge et al., 2013a,b; Mathesius, 2018). Flavonoids reported in seeds of ten major nut plant species are subdivided into eight subclasses, i.e., flavanols, flavones, flavonols, flavanones, flavanonols, isoflavones, anthocyanins, and proanthocyanidins (**Figures 1B,C**; Tohge et al., 2013b; Saigo et al., 2020). Tannins, comprising of hydrolyzable and condensed tannins, are astringent and bitter plant substances being particularly numerous in nuts seeds (Landete, 2011). Stilbene compounds are comprised of two aromatic rings connecting with an ethylene bridge (C6–C2–C6 backbone) and commonly found as monomers and oligomers in both aglycone and glycoside forms (**Figure 1D**; Akinwumi et al., 2018; Hano and Tungmunnithum, 2020). Oxidative dimerization of two or more phenylpropanoid molecules produces lignans for lignin synthesis, and as such important precursors for the development of plant vessels and cell walls (**Figure 1E**; Barker, 2019). Coumarins are plant benzo-α-pyrone compounds generated by the reaction of pyran-benzene ring condensation (**Figure 1F**; Küpeli Akkol et al., 2020).

With respect to biological or pharmaceutical activities, polyphenols have been found to be beneficial components in both human health promotion and disease prevention. They were shown to exhibit antioxidative, anticancer, cardio-protective, antibacterial, anti-inflammatory, and immune system-promoting properties and to exert protection for skin against UV radiation, against neurodegenerative diseases, chronic diseases, obesity, and diabetes, and against the current pandemic coronavirus disease (COVID-19), being reported (Nayak et al., 2015; Tohge and Fernie, 2017; Puksasook et al., 2017a,b; Cory et al., 2018; Tungmunnithum et al., 2018; Renaud and Martinoli, 2019; Ngwa et al., 2020). Several review articles focus on the health benefits of nut consumption (Ros, 2010; Bolling et al., 2011; Alasalvar and Bolling, 2015), with the major groups of polyphenols present in nut seeds being characterized (Bolling et al., 2011; Alasalvar and Bolling, 2015; Bodoira and Maestri, 2020). Currently, the chemical diversity of specialized metabolites and metabolic polymorphisms have specifically highlighted the decoration of polyphenols, with such decorations being found to be a key factor in the enhancement of bioactivity of specialized

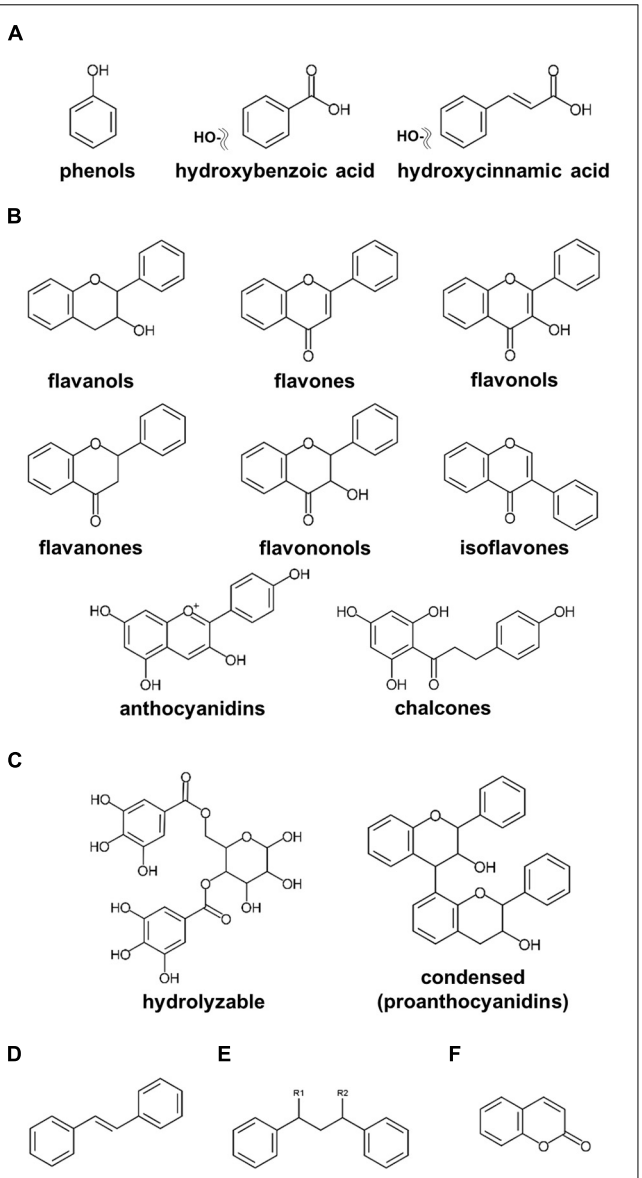

FIGURE 1 | Core structures of polyphenolics found in major nut plant species. **(A)** phenols and phenolic acids. **(B)** flavonoids. **(C)** tannins. **(D)** stilbenoids. **(E)** lignans. **(F)** coumarins.

compounds (Tohge et al., 2016; Peng et al., 2017). In fact, most of the aforementioned biological activities of polyphenols are derived from not only the aglycone form but also the decorated form. The capacity of biological activities such as antioxidant capacity significantly depends on the chemical structures, since these dramatically affect the bioavailability and especially absorption of the compounds (Cipolletti et al., 2018; Gulcin, 2020). Although the core biosynthetic pathways of many polyphenols are conversed among genetically and taxonomically distant plant species, these species often accumulate polyphenols in a tissue-specific manner. In the case of seed-specific specialized metabolites, polyphenols are assumed to be involved in environmental stress protection during seed desiccation and dormancy. As such, the elucidation and understanding of physiological functions of such tissue-specific specialized metabolites are highly valuable.

Current technological and theoretical development of omics-based approaches has enabled that the genome-wide characterization of biosynthetic genes can be carried out, representing an important route by which phenol and polyphenol production could be enhanced in plants (Butelli et al., 2008; Tohge and Fernie, 2010; Zhang et al., 2015; Fernie and Tohge, 2017; Alseekh et al., 2020b). Notably, almost all genes encoding enzymes responsible for structure decoration remain ambiguous. However, as yet, genomic data is only available for peanut and almond; therefore, studies concerning nut plant polyphenolics are not as extensive as they could have been. As such, it is important to update and synthesize the collective information concerning chemical diversity in nut plant species. Given the presence of some species-specific polyphenolics in nuts, it is likely that such a compendium will prove a useful resource for biological activity investigations.

In this review, the current knowledge of polyphenolic compounds in major nut plant species is summarized in terms of their biological activities, chemical diversity, and biosynthetic genes. Ten eminently consumed nut plant species, namely, groundnut/peanut (*Arachis hypogaea*), almond (*Prunus dulcis*), pistachio (*Pistacia vera*), Japanese chestnut (*Castanea crenata*), Chinese chestnut (*Castanea mollissima*), walnut (*Juglans regia*), hazelnut (*Corylus avellana*), macadamia (*Macadamia integrifolia*), pecan (*Carya illinoinensis*), and cashew nut (*Anacardium occidentale*), are presented (**Figure 2** and **Table 1**). We additionally propose a future perspective for generating an integrative omics approach for functional genomics utilizing polyphenolic biosynthesis in nut plant species as a case study.

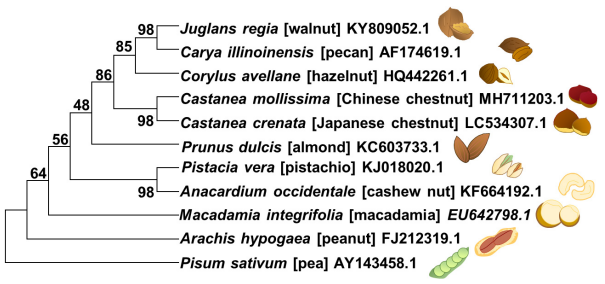

FIGURE 2 | Phylogenetic relationship of ten nut plant species presented in this review. The phylogenetic tree was constructed by MEGA X (Kumar et al., 2018) by using the sequences of ITS (internal transcribed spacer) gene of each species which were retrieved from the NCBI database. *P. sativum* was considered as an outgroup. The parameters for maximum-likelihood analysis were shown as followed: Kimura 2-parameter model, complete deletion, and bootstrap (1000 replicates). Values presented on the branches demonstrated bootstrap support in percentage.

POLYPHENOLICS IN NUTS AND THEIR BIOACTIVITIES ASSOCIATED WITH HEALTH-PROMOTING BENEFITS

Several popular nut plant species are consumed as snacks and food supplements, since they are rich in phytonutrients especially fatty acids, protein, minerals, and polyphenolics (Ros, 2010; Vinson and Cai, 2012). To date, research in food chemistry has suggested several biological activities of nut extracts with studies on the antioxidant activity being particularly prominent. From a health beneficial perspective, antioxidants are responsible for the elimination of reactive oxygen species (ROS) or free radical molecules such as superoxide, nitric oxide, and hydrogen peroxide radicals from the human body in order to prevent the generation of hazardous substances which underlie many chronic diseases (Gulcin, 2020). Polyphenols are one of the best-known and major sources of natural antioxidants due to their effective scavenging activity resulting from the presence of several hydroxyl groups present on the structures, especially those on the *ortho-* and *para-* positions of the aromatic ring (Shahidi and Ambigaipalan, 2015; Gulcin, 2020). On the basis of their total polyphenolic quantity, the highest antioxidant activities were found in raw walnut and roasted almond (Vinson and Cai, 2012). The biological activities of nut polyphenol antioxidants against major stress such as oxidative stress, aging, and age-related disease prevention were summarized in previous reviews (Ros, 2010; Bodoira and Maestri, 2020; Hano and Tungmunnithum, 2020). Nut consumption has been suggested to play a key role in cardio-protection by reducing cardiovascular risk factors, including coronary heart disease, hypertension, and blood cholesterol levels (Ros, 2010; De Souza et al., 2017). In a recent study, almond skin extract was found to show antimicrobial and antiviral activities against *Staphylococcus aureus* and herpes simplex virus type I, respectively (Musarra-Pizzo et al., 2019). Due to such indication of nut consumption and health-promoting benefits, polyphenols contained in culinary nuts have been focused and studied with quantification of known health-promoting polyphenols in food chemistry research. Bioactivities of major and specific polyphenolics in nut plant species are summarized in **Table 2**. Given that health beneficial components such as resveratrols, chlorogenic acids, catechins, and rutin are detected in nuts, health-promoting benefits of culinary nuts are considered with bioactivities and concentrations of these polyphenols. Additionally, nut-specific polyphenolics such as cardanols and anacardic acid were found as cashew nut-specific antioxidant compounds

TABLE 1 | Nut plant species presented in this article.

Name	Species name	Genome sequencing	BioProject ID (NCBI)
Groundnut/peanut	*Arachis hypogaea*	Bertioli et al., 2019	PRJNA419393,PRJNA480120, PRJNA680825
Almond	*Prunus dulcis*	Sánchez-Pérez et al., 2019	PRJDB7547, PRJNA497779
Pistachio	*Pistacia vera*	Zeng et al., 2019	PRJNA578116
Japanese chestnut	*Castanea crenata*		
Chinese chestnut	*Castanea mollissima*		PRJNA559042
Walnut	*Juglans regia*	Martínez-García et al., 2016	PRJNA350852, PRJNA445704
Hazelnut	*Corylus avellana*	Lucas et al., 2020	PRJEB31933
Macadamia	*Macadamia integrifolia*		
Pecan	*Carya illinoinensis*		
Cashew nut	*Anacardium occidentale*		

(**Table 2**). Ellagic acid, which is present in several nuts including almond, walnut, pecan, Japanese chestnut, and hazelnut, has been reported as an inhibitor of inflammatory mediator molecules such as cyclooxygenase and nuclear factor κB, providing anti-inflammatory activity (El-Shitany et al., 2014). Anacardic acid from cashew nut, ellagic acid from walnut and pecan, genistein from peanut and hazelnut, and resveratrol from peanut have demonstrated anticancer properties with numerous molecular targets (Falasca et al., 2014). Resveratrol and its prenylated derivatives in peanut have been reported to mitigate against neurodegenerative diseases such as Alzheimer's and Parkinson's disease via their antioxidant, anti-β-amyloid aggregation, anti-β-secretase, neuroprotective, and neuritogenicity properties (Puksasook et al., 2017a,b; Navarro et al., 2018). Captivatingly, myricetin which is found in pistachio and hazelnut as well as resveratrol in peanut, almond, and pistachio are recently claimed as potential phytochemical compounds that could counteract the current COVID-19 pandemic (Han et al., 2020; Ngwa et al., 2020).

CHEMICAL STRUCTURAL DIVERSITY OF POLYPHENOLICS AMONG SEEDS OF NUT PLANT SPECIES

The chemo-diversity of plant metabolism is a highly important factor affecting plant ecological processes and plant metabolic evolution (Kessler and Kalske, 2018). Furthermore, the various characteristics of phytochemical structures show diverse modes of action with regard to the prevention and treatment of human diseases given differences in their physicochemical properties. Recent overviews focusing on plant structural diversity in anthocyanins/proanthocyanidins (Saigo et al., 2020), glucosinolates (Blažević et al., 2020), and diterpenoid alkaloids (Shen et al., 2020) have been published. With this regard, for example, some decorations such as oxidation for enhancing acidity, methylation and acylation for reducing polarity, and glycosylation for stability and solubility are considered as the potential factors corresponding to diversification of biological functions. Moreover, the updated plant chemo-diversity database is a powerful tool for enthusing new pharmaceutical drug discovery (Lautié et al., 2020). To illustrate the diversity of chemical structures of polyphenolics found in seeds of nut plant species, both raw and processed nut seeds were included in our chemical diversity analysis from the renowned literature-based phytochemical database KNApSAcK (http://kanaya.naist.jp/KNApSAcK/, searched by plant scientific names in July, 2020; Afendi et al., 2011). Furthermore, several current phytochemical reports were included to illustrate the structural diversity of nut polyphenols. A list of all 214 polyphenols is provided in **Supplementary Table 1**, while the structural diversity of nut polyphenolic compounds is presented in **Figure 3**.

Phenols and Phenolic Acids

Several types of phenols have been found in nut plant species (**Table 2** and **Figure 3A**). Four chemical isomers, *alpha-*, *beta-*, *gamma-*, and *delta-* of tocopherol, were detected in pistachio kernel, walnut kernel, and whole cashew nut (Horvath et al., 2006; Ballistreri et al., 2009; Trox et al., 2011). Cardanols and their derivatives were reported only in the kernel of the cashew nut (Trevisan et al., 2006; Alvarenga et al., 2016; Bodoira and Maestri, 2020). In pecan and walnut kernel, many glycosylated and methylated ellagic acid derivatives are present at high abundance (De La Rosa et al., 2011; Grace et al., 2014; Regueiro et al., 2014; Robbins et al., 2014; Jia et al., 2018). Gallic, *p*-coumaric, chlorogenic, and *p*-hydroxybenzoic acids are the most abundant phenolic acids found in the kernel and skin of seeds of ten nut plant species. In addition, several phenolic acids were specifically presented in some of the nut plant species. Anacardic acid and derivatives were present in cashew nut and pistachio kernel (Trevisan et al., 2006; Alvarenga et al., 2016; Bodoira et al., 2019; Salehi et al., 2019; Bodoira and Maestri, 2020). Coutaric and fertaric acids are constituents in peanut and hazelnut skin (Ma et al., 2014; Pelvan et al., 2018). Chinese chestnut skin contains high levels of gentisic acid and 2,3,4-trihydroxybenzoic acid (Xu et al., 2020). In fact, Bodoira and Maestri have mentioned that phenolic acids of peanut are found only in the skin (Bodoira and Maestri, 2020). Nevertheless, nut phenolic acids, such as phloretic acid and dihydroxybenzoic acid, were found not only in the skin but also in the kernel of peanut (Bisby, 1994). Apart from common phenolic acids, numerous derivatives of gallic, hydroxybenzoic, and hydroxycinnamic acids, such as esterification with other phenolic acids, hydroxylated and

TABLE 2 | Bioactivities of major and specific polyphenolics in nut plant species.

Compound name	Nut species	Bioactivities
Phenols		
Cardanols*	Cashew nut	Antioxidant, antimutagenic, and antitumoral activity (Maia et al., 2015; Schneider et al., 2016)
Ellagic acid	Almond, walnut, Japanese chestnut, pecan, hazelnut	An inhibitor of inflammatory mediators (El-Shitany et al., 2014)
Gallic acid	Almond, cashew nut, Chinese chestnut, hazelnut, Japanese chestnut, pecan, pistachio, walnut	Antioxidant, anti-inflammatory, anticancer, antimicrobial, cardiovascular, and gastrointestinal treatment, protective effect on neuropsychological diseases (Zanwar et al., 2014; Kahkeshani et al., 2019)
p-Hydroxy benzoic acid	Peanut, almond, walnut, Japanese chestnut, Chinese chestnut, hazelnut, pecan	Osteogenic activity, antimicrobial activity, antifungal, estrogenic, and antimutagenic properties (Cho et al., 1998; Pugazhendhi et al., 2005; Heleno et al., 2015)
Chlorogenic acid	Almond, Chinese chestnut, hazelnut, peanut, pecan, pistachio, walnut	Antioxidant, anti-hepatitis B virus, antidiabetic effect, DNA protective effect, neuroprotective effect, protection from cardiovascular diseases (Sato et al., 2011; Zuo et al., 2015; Gil and Wianowska, 2017)
p-Coumaric acid	Peanut, almond, walnut, Japanese chestnut, Chinese chestnut, cashew nut, hazelnut	Antioxidant, hyperlipidemia treatment, antimicrobial, antiviral, anti-inflammatory, anticancer, antidiabetic (Lee et al., 2019; Shen et al., 2019)
Anacardic acid and its derivatives*	Cashew nut	Antioxidant, antibacterial, cytotoxicity against *A. salina*, acetylcholinesterase inhibition (Muroi and Kubo, 1996; Maia et al., 2015; Morais et al., 2017)
Stilbenoids		
Resveratrol	Peanut, almond, pistachio	Antioxidant,cancer chemopreventive, anti-β-amyloid aggregation, anti-β-secretase activity, neuroprotective, neuritogenicity, cardiovascular protective, anti-inflammatory, blood glucose-lowering, anticancer, anti-obesity (Kuršvietienė et al., 2016; Puksasook et al., 2017a,b)
Flavonoids: flavonol		
(+)-Catechin	Peanut, almond, pistachio, walnut, pecan, Chinese chestnut, cashew nut, hazelnut	Antioxidant, antimicrobial, antiviral, anti-inflammatory, anti-allergenic, anticancer, prevention of cardiovascular diseases, and neurodegenerative diseases (Isemura, 2019; Bae et al., 2020)
(–)-Epicatechin	Peanut, almond, pistachio, pecan, cashew nut, hazelnut, walnut	Antioxidant, antidiabetes, anticancer, anti-inflammatory, antihypertensive, antidyslipidemic (Abdulkhaleq et al., 2017; Bernatova, 2018)
Flavonoids: flavone		
Luteolin	Peanut, almond, pistachio, Chinese chestnut	Antioxidant, cardioprotective effects, anti-inflammatory, antidiabetic, antimicrobial, anticancer (Lin et al., 2008; Dong et al., 2017; Luo et al., 2017)
Flavonoids: flavonol		
Quercetin	Peanut, almond, pistachio, Chinese chestnut, hazelnut, walnut	Antioxidant, anti-inflammatory, cardiovascular disease prevention, neurodegenerative disorders treatment, anticancer, antibacterial, antiviral (Anand David et al., 2016; Xu et al., 2019)
Rutin	Almond, pistachio, walnut, Chinese chestnut, hazelnut	Antioxidant, neuroprotective, hepatoprotective, cardioprotective, antifungal, antimalarial, antibacterial, anticancer (Gao et al., 2002; Ganeshpurkar and Saluja, 2017)
Isoquercitrin	Almond, pistachio, walnut, hazelnut	Antioxidant, neurological disorders, anti-allergic, antidiabetic, anti-inflammatory (Palazzolo et al., 2012; Valentová et al., 2014)
Flavonoids: flavanone		
Eriodictyol	Peanut, almond, pistachio, Chinese chestnut, hazelnut	Antioxidant, cardioprotective, skin protection, antitumor, antidiabetic, anti-inflammatory, cytoprotective, hepatoprotective, neuroprotective (Deng et al., 2020; Li et al., 2020)
Lignans		
(+)-Lariciresinol	Almond, cashew nut, chestnut, hazelnut, peanut, pecan, pistachio, walnut	Antifungal, antibacterial (Hwang et al., 2011; Bajpai et al., 2017)
(–)-Matairesinol	Almond, cashew nut, chestnut, hazelnut, peanut, pecan, pistachio, walnut	Antioxidant, anti-osteoclastogenic, anti-angiogenic, anticancer, antifungal, IgE-suppressive activity (Yamauchi et al., 2006; Kawahara et al., 2010; Choi et al., 2014)

**Specific present in cashew nut.*

methylated derivatives, are also conspicuously reported in nut seeds (Bodoira and Maestri, 2020). Polymerization of tartaric acid with two other molecules of phenolic acids is produced only in peanut skin reported by Ma et al. (2014).

Flavonoids

A multitude of flavonoids including flavanols, flavones, flavonols, flavanones, flavanonols, isoflavones, anthocyanins, and proanthocyanidins have been found nut plant species (**Table 2**

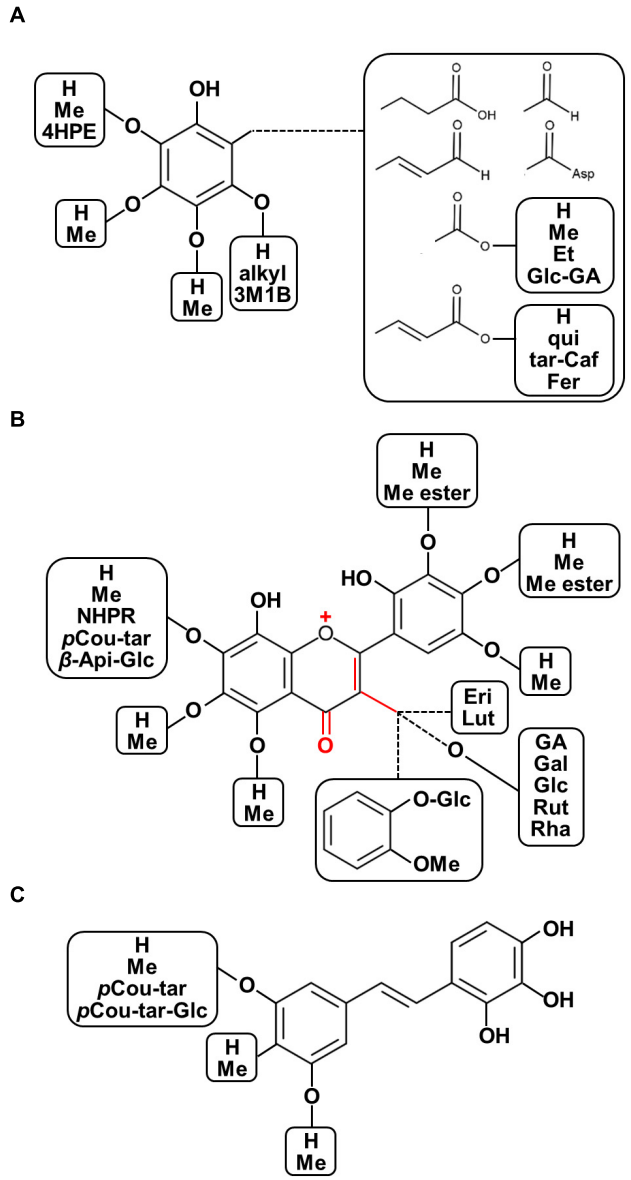

FIGURE 3 | Diversity of chemical structures of polyphenolics found in seeds of nut plant species. **(A)** Phenols and phenolic acid derivatives, **(B)** flavonoids, and **(C)** stilbenoids. Abbreviations: OMe, methoxy; 4HPE, 4-(hydroxypheny)ethylene; Asp, aspartate; Me, methyl; Et, ethyl; qui, quinic acid; tar, tartaric acid; fer, ferulic acid; 3M1B, 3-methyl-1-butenyl; *p*Cou-tar, *p*-coumaroyl-tartarate; *p*Cou-tar-Glc, *p*-coumaroyl-tartarate-glucoside; Pre, isoprenyl group; Me ester, methyl ester; NHPR, neohesperidose; Api-Glc, -apiosyl-glucose; Eri, eriodictyol; Lut, luteolin; GA, gallic acid; Gal, galactose; Glc, glucose; Rut, rutinose; and Rha, rhamnose.

and **Figures 1B**, **3B**). When the hydroxyl group is connected to carbon position three of the C ring, the molecule belongs to the flavanol subclass. Flavonols represent the principal subgroups of flavonoids since they display a rich diversity of derivatives and are also a basic structure of proanthocyanidins (Tohge et al., 2017; Saigo et al., 2020). Catechin and epicatechin are visibly rich in seeds of almost all of the ten major nut plant species. Moreover, their derivatives via esterification with gallic acid, epicatechin-gallate, gallocatechin-gallate, epigallocatechin, and epigallocatechin gallate, were also determined to be abundant in these species being found in whole almond seed (Bolling, 2017), kernel and skin of cashew nut (Salehi et al., 2019), pecan, walnut, Chinese chestnut kernel (Regueiro et al., 2014; Jia et al., 2018; Zhang Y. et al., 2020), and hazelnut skin (Del Rio et al., 2011; Pelvan et al., 2018). Luteolin and apigenin are typical flavones found in peanut skin (Bodoira et al., 2017), almond kernel (Čolić et al., 2017), kernel and the skin of pistachio (Tomaino et al., 2010; Fabani et al., 2013), Chinese chestnut kernel (Zhang Y. et al., 2020), and Japanese chestnut skin (Tuyen et al., 2017). Predominantly among studied nut plant species, whole almond seeds comprise the most diverse types of flavonols, including quercetin, kaempferol, isorhamnetin, and their *O*-glycoside derivatives (Milbury et al., 2006; Monagas et al., 2007; Bolling et al., 2010; Valdés et al., 2015; Bolling, 2017; Čolić et al., 2017). Flavonol derivatives are frequently glycosylated at the hydroxyl group at position three. Aside from major flavonols, in Chinese and Japanese chestnut kernel, minor flavonols such as rhamnetin and morin have been detected (Tuyen et al., 2017; Zhang Y. et al., 2020). Metabolites containing the saturated C ring belong to the flavanone subgroup of flavonoids. Eriodictyol is the main flavanone detected among nut seeds. As for flavonols, several types of flavanones and their derivatives are characterized in whole almond seeds (Milbury et al., 2006; Čolić et al., 2017; Bodoira and Maestri, 2020) with glycosylated derivatives generally displaying glycosylation on the hydroxyl group at position seven. Particularly, naringin (naringenin-7-*O*-neohesperidoside) is a major flavanone derivative in kernels of almond, walnut, and Chinese chestnut as well as in pistachio kernel and skin (Tomaino et al., 2010; Čolić et al., 2017; Vu et al., 2018; Zhang Y. et al., 2020). By contrast, flavanonols and hydroxylated derivatives of flavanones are only minor constituents of nuts; for example, aromadendrin and taxifolin are reported in whole almond seeds (Monagas et al., 2007; Bolling, 2017; Vu et al., 2018). The core structure of isoflavones differs from other flavonoids by linkage of the phenyl ring to position three of ring C supplemented with the ketone group at position four. Isoflavones are not broadly found in nut seeds. However, genistein and its glucosides, daidzein, and daidzin, are present in peanut kernel and skin, whole almond, and pistachio kernel (Ballistreri et al., 2009; Tomaino et al., 2010; Bolling, 2017; Bodoira and Maestri, 2020). Anthocyanidins (aglycone form) and anthocyanins (glycoside form) are extensively known as plant pigments in seeds, flowers, and fruits (Saigo et al., 2020). Cyanidin and its glycoside derivatives, i.e., glucoside and galactoside, were reported as constituents in whole almond seeds (Bolling, 2017) and pistachio kernel and skin (Ballistreri et al., 2009; Tomaino et al., 2010; Fabani et al., 2013; Bodoira and Maestri, 2020). Chalcones or the so-called open-chain flavonoids are a very unique flavonoid subclass in the plant kingdom. Phloretin and its glucoside derivative, phlorizin, are the two chalcones found in the almond kernel and hazelnut skin (Del Rio et al., 2011; Čolić et al., 2017).

Tannins

Both tannin subgroups, i.e., hydrolyzable and condensed tannins, are abundant in peanut skin, whole almond, and walnut kernel. Hydrolyzable tannins are further allocated into two subgroups gallotannins and ellagitannins (Soares et al., 2020). The latter including strictinin, pedunculagin, tellimagrandin, glasrin, rugosin, casuarinin, and praecoxin (Fukuda et al., 2003; Grace et al., 2014; Regueiro et al., 2014; Robbins et al., 2014; Jia et al., 2018) are more prominent than the former. Brown or non-visible colors of plant seed, peel, and bark are caused by proanthocyanidins or condensed tannins (Saigo et al., 2020). As mentioned above, proanthocyanidin is synthesized by combining flavonol molecules, i.e., catechin, epicatechin, or their galloylated derivatives; polymerization started from dimers by several types of inter flavan linkage (Soares et al., 2020). Proanthocyanidins have been characterized in seed testa of nut plant species, especially in peanut, almond, pistachio, and hazelnut (Lou et al., 1999, 2004; Yu et al., 2006; Monagas et al., 2007; Del Rio et al., 2011; Sarnoski et al., 2012; Fabani et al., 2013; Grace et al., 2014; Slatnar et al., 2015; Bodoira et al., 2017, 2019; Bolling, 2017; Pelvan et al., 2018; Bodoira and Maestri, 2020). As stated in the current review article (Saigo et al., 2020), proanthocyanidins and tannins are arduous to study, since their condensed structures are very complex due to high molecular weight and various types of chemical bond configuration.

Stilbenoids

Typically, stilbene compounds are rarely found in the seeds of nut plant species. Resveratrol, one of the most well-known stilbenes ubiquitously found in grape, *Vitis vinifera* L. (Salehi et al., 2018), was detected in peanut skin, whole almond seeds, and pistachio kernel (Ballistreri et al., 2009; Ballard et al., 2010; Xie and Bolling, 2014; Čolić et al., 2017; Bodoira and Maestri, 2020). Moreover, various types of stilbene derivatives were also found. Two prenylated resveratrols, arachidin I and II, were found in the peanut kernel (Bisby, 1994). In whole almond seed, the glycosylated resveratrol named polydatin is the most prominently detected stilbenoid along with small amounts of a methylated and two hydroxylated resveratrols called pterostilbene, piceatannol, and oxyresveratrol, respectively (Xie and Bolling, 2014).

Lignans and Coumarins

Although lignans are not well-investigated in nut plant species, several major plant lignans and their hydroxylated derivatives have been reported in the nut species that we are reviewing in this article, such as lariciresinol, matairesinol, secoisolariresinol, cyclolariciresinol, and 7-hydroxymatairesinol, whereas cashew nut contains the highest total lignan contents (Bolling, 2017; Rodríguez-García et al., 2019). The same is true of lignans in nut species; coumarins in seeds of nut plant species have not been characterized well. Some simple and pyrone-substituted coumarin compounds, e.g., aesculin, aesculetin, umbelliferone, and coumestrol, are mostly found in whole almond, Japanese chestnut skin, and Chinese chestnut kernel (Bolling, 2017; Čolić et al., 2017; Tuyen et al., 2017; Chang et al., 2020).

RECENT UPDATES OF THE POLYPHENOLIC BIOSYNTHETIC FRAMEWORK IN SEEDS OF NUT PLANT SPECIES

An overview of the known polyphenolic biosynthetic framework is summarized in **Figure 4** (Lushchak and Semchuk, 2012; Cheynier et al., 2013; Anantharaju et al., 2016; Valanciene et al., 2020). Even though the enzymatic genes regarding polyphenolic biosynthesis are elucidated and well characterized in model plants and crop species, such as *Arabidopsis thaliana*, *Zea mays*, and *Camellia sinensis* (Falcone Ferreyra et al., 2012; Jiang et al., 2013), numerous key genes are largely uncharacterized in seeds of nut plant species, probably due to the lower sequence similarity of genes between major model plants and model nut plant species. Currently, the information of biosynthetic genes has been reported only in peanut and pecan. Peanut chalcone isomerase (CHI) types I and II have been identified (Wang et al., 2012; Liu et al., 2015), and stilbene synthase (*AhSTS*), the most vital enzyme for resveratrol biosynthesis, is functionalized (Condori et al., 2009). In the pecan kernel, three chalcone synthases (CHS) were isolated and properly characterized (Zhang et al., 2019b). Moreover, RNA-Seq also provided several other gene candidates of flavonoid biosynthesis, including PAL, C4H, 4CL, CHI, F3H, F3′H, DFR, ANS, LAR, ANR, and MYB transcription factor (Zhang et al., 2019a). However, those candidates have been well identified by Huang and colleagues (Huang et al., 2019). In 2014, Chen et al. (2014) performed the first MYB gene family in peanut investigation and found that AhMYB15 is related to flavonol biosynthesis. Aside from seeds, gallate glucosyltransferase (GGT) I and II are important for β-glucogallin, the intermediate compound for hydrolyzable tannin biosynthesis, which were identified in walnut leaves (Martínez-García et al., 2016). Stilbenoid prenyltransferases named AhR4DT-1 and AhR3′DT-1 involved in prenylated resveratrol production in peanut kernel were explicated in peanut hairy root (Yang et al., 2018). According to current knowledge, it is apparent that the responsible key genes for polyphenolics in the seed of nut plant species remain largely undetermined, especially in the case of decoration enzymatic genes, which are largely responsible for producing chemical structure diversification. Identification of these genes thus represents an important priority for future research.

MODEL NUTS: THE STATE OF THE ART

At present, nut plant species, including peanut (Bertioli et al., 2019), almond (Sánchez-Pérez et al., 2019), pistachio (Zeng et al., 2019), walnut (Martínez-García et al., 2016), and hazelnut (Lucas et al., 2020), have been genome sequenced (**Table 1**). Furthermore, draft genome sequences of some chestnut cultivars have been currently deposited in the genome database. Such information provides the advantages of those plant species which became "model nut" species for elucidation of physiological and biological functions of key metabolic genes via omics-based approaches. The genome sequence of the cultivated peanut,

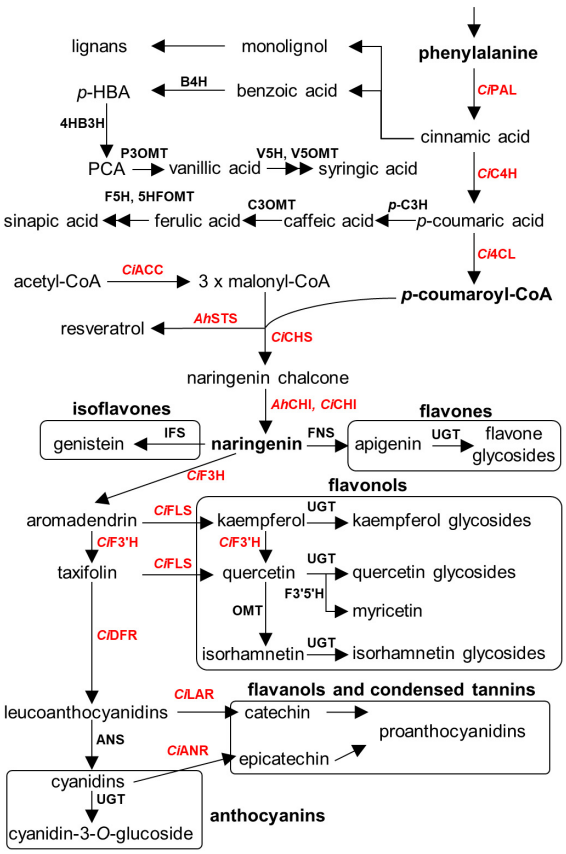

FIGURE 4 | Overview of polyphenolics biosynthesis in major nut plant species. Genes involved in each biosynthetic step are displayed. Characterized genes are shown in red color. Abbreviations used. E4P, erythrose-4-phosphate; PEP, phosphoenolpyruvate; DAHPS, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; DHQS, 3-dehydroquinate synthase; DHQ, 3-dehydroquinate; DHD, 3-dehydroquinate dehydratase; DHS, 3-dehydroshikimate; *Jr*GGT, gallate 1-β-glucosyltransferase from walnut; SDH, shikimate dehydrogenase; SK, shikimate kinase; EPSPS, 5-enolpyruvylshikimate 3-phosphate synthase; CS, chorismate synthase; CM, chorismate mutase; HPPD, *p*-hydroxyphenylpyruvate dioxygenase; PAT, prephenate aminotransferase; AD, arogenate dehydratase; *Ci*PAL, phenylalanine ammonia lyase of pecan; B4H, benzoic acid 4-hydroxylase; 4HB3H, 4-hydroxybenzoic acid 3-hydroxylase; P3OMT, protocatechuic acid-3-*O*-methyltransferase; V5H, vanillic acid 5-hydroxylase; V5OMT, vanillic acid 5-*O*-methyltransferase; *Ci*C4H, cinnamic acid-4-hydroxylase from pecan; *p*-C3H, *p*-coumaric acid 3-hydroxylase; C3OMT, caffeic acid 3-*O*-methyltransferase; F5H, ferulic acid 5-hydroxylase; 5HFOMT, 5-hydroxyferulic acid *O*-methyltransferase; *Ci*4CL, 4-coumaroyl-CoA ligase from pecan; *Ci*ACC, acetyl coenzyme A carboxylase from pecan; *Ah*STS, stilbene synthase from peanut; CiCHS, chalcone synthase from pecan; *Ah*CHI, *Ci*CHI, chalcone isomerase from peanut and pecan; IFS, isoflavone synthase; FNS, flavone synthase; *Ci*F3H, flavanone 3-hydroxylase from pecan; *Ci*FLS, flavonol synthase from pecan; *Ci*F3′H, flavonoid 3′-hydroxylase from pecan; *Ci*F3′5′H, flavanoid 3′5′-hydroxylase from pecan; *Ci*DFR, dihydroflavonol reductase from pecan; ANS, anthocyanidin synthase; *Ci*LAR, leucoanthocyanidin reductase; *Ci*ANR, anthocyanidin reductase from pecan; and UGT, uridine diphosphate glycosyltransferase.

known as allotetraploids, was elucidated after sequencing its diploid ancestors, *A. duranensis* and *A. ipaensis*, providing helpful hints for peanut domestication (Bertioli et al., 2016). Currently, using peanut genomics data, lipid metabolism (Chen X. et al., 2019) and the genes involved in size and lipid content in seeds, leaf disease resistance, and nitrogen fixing capacity (Zhuang et al., 2019) were investigated and annotated. Importantly, PacBio and chromosome conformation capture (Hi-C) technologies were performed in order to improve data reading quality and a complete peanut genome sequence was lately reported (Bertioli et al., 2019), resulting in a very high-quality genome.

The walnut genome which was first reported in 2016 revealed some of the genes involved in polyphenolic transformations (Martínez-García et al., 2016). The first walnut reference genome was used to generate high-density 700-K single-nucleotide polymorphism (SNP) arrays (Marrano et al., 2019). Importantly, this tool was used to identify gene candidates responsible for flowering process disclosure (Bernard et al., 2020). However, given that the early walnut genome remained at the scaffold level, multi-omics studies for unraveling biological function and regulation were obstructed. For this reason, the *de novo* assembly of the complete walnut genome was attempted by various techniques. The hybridization of walnut species sequencing by single-molecule or PacBio long-read sequencing and optical genome mapping technologies demonstrated a high quality of parental genome sequence (Wu and Gmitter, 2019; Zhu et al., 2019). Recently, nanopore long-read sequencing supplemented with Hi-C technology has been utilized for a high-quality walnut chromosome level genome assembly (Marrano et al., 2020). Illumina sequencing coupled with Hi-C data is also found to provide high-quality genome data and uncover differences (Zhang J. et al., 2020). Almond is also recently described to have a complete genome sequence. The complete genome sequence of almond was initially reported to be coupled with forty-six kilobases of the gene cluster encoding five basic helix–loop–helix (bHLH) transcription factors (Sánchez-Pérez et al., 2019). Fascinatingly, bHLH2 is identified to be involved in amygdalin biosynthesis (Sánchez-Pérez et al., 2019). Alioto and colleagues similarly performed almond genome sequencing and determined transposon elements related to amygdalin biosynthesis and diversification in peach (Alioto et al., 2020). With an ever-increasing number of high-quality nut plant genomes, gene conservation of functional genes is one of the important topics deserving further investigation of plant metabolism. Hazelnut is the most recently reported complete genome sequence by a hybrid sequencing strategy combining short reads, long reads, and proximity ligation methods (Lucas et al., 2020). The European hazelnut (Corylus avellana L. cv. Tombul) was sequenced focusing on gene families encoding hazelnut allergens and the pathogen-resistance locus proteins that are an important for crop improvement in *C. avellana*.

BIOSYNTHETIC GENE CONSERVATION IN NUT GENOMES

In spite of the fact that flavonoids are highly diverse in seeds of nut plant species, the enzymatic gene involved in flavonoid biosynthesis named CHS is frequently mentioned as one of the most conserved key enzymes. In addition, stilbene synthase

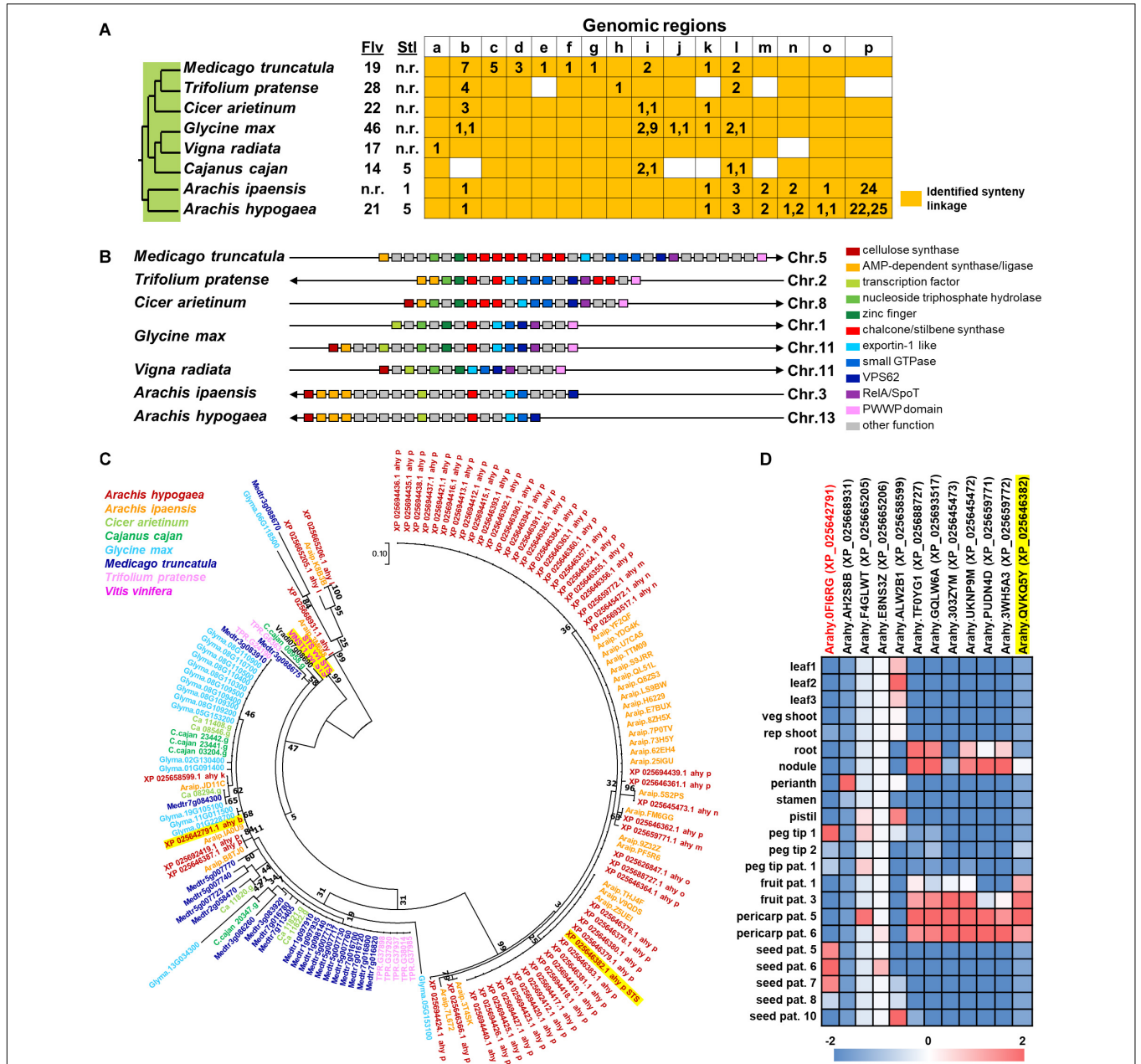

FIGURE 5 | Comparative genomics of CHS and STS in eight legume plant species. **(A)** Genomic synteny analysis of CHS and STS. Gene lists of six legume plant species were retrieved from Plaza database (Dicot 4.5; http://bioinformatics.psb.ugent.be/plaza/). Almond and peanut genome sequences were obtained from the NCBI database. Numbers in each region indicate the number of tandem duplicated genes. Numbers in each region indicate the number of tandem duplicated genes and intra synteny genes. Abbreviations used: Flv, flavonoids; and Stl, stilbenoids. **(B)** Genomic structure of synteny regions b presented in **Figure 1A**. **(C)** Phylogenetic relationship of CHS and STS located in genome synteny analysis of thirteen plant species. Amino acid sequences were attained from the Plaza database (Dicot 4.5; http://bioinformatics.psb.ugent.be/plaza/) coupled with the NCBI database. The phylogenetic trees were constructed with aligned protein sequences by MEGA7 (Kumar et al., 2016) using the neighbor-joining method with the following parameters: Poisson correction, complete deletion, and bootstrap (1000 replicates, random seed). The protein sequences were aligned by MUSCLE implemented in MEGA. Values on the branches indicate bootstrap support in percentages. The tree with the highest log likelihood (–1287.37) is shown. Colors of circle indicate plant species. **(D)** Gene expression profile of peanut tissues. Each gene is connected with the syntenic region. Gene expression data; part and stage of peanuts are described by previous work of Clevenger and colleagues (Clevenger et al., 2016).

(STS), the key enzyme of stilbene biosynthesis which resulted in resveratrol as a first product, is discovered as a homolog of CHS since they play a similar function and contain a conserved cysteine residue; hence, STS is described as belonging to the CHS family (Schröder and Schröder, 1990; Lanz et al., 1991). In the nut species presented in this article, *AhSTS1*

(arahy.QVKQ5Y, XM_025790597.1, ABY86219.1) in *A. hypogaea* has been characterized (Condori et al., 2009). Both CHS and STS belong to the type III polyketide synthase (PKS) family and have been occurred during functional diversification of PKS (Tohge et al., 2013b). In current informatics research, the phylogenomic synteny network combined with phylogenetic analyses of whole-genomic data of 126 plant species has been developed and focused on the macroevolution of diversification of the PKS family (Naake et al., 2020). In this review, we performed the comparative genomics of CHS and STS between nuts and closely taxonomic-related species by phylogenomic synteny analysis. The total diverged genomic synteny regions of CHS and STS in eight legume plant species are comprised in twenty regions (**Figures 5A,B**). Three regions (b, k, and l) are commonly found in all studied plant species, whereas m, n, o, and p regions are specific to peanut species. Additionally, the numerous tandem gene duplication is prominently presented in regions b, c, d, I, l, m, n, and p. Even though CHS and STS genes are generally noticed in every mentioned plant species, the flavonoid and/or stilbenoid are still not reported providing a research gap for further metabolite investigation. Based on CHS and STS syntenic regions, a phylogenetic analysis demonstrates the clearly visible separation of region p in wild and cultivated peanut (*Arachis ipaensis* and *A. hypogaea*, respectively), soybean (*Glycine max*), red clover (*Trifolium pretense*), and medicago (*Medicago truncatula*) from the others; it suggests that those uncharacterized genes may function as CHSs (**Figure 5C**). Interestingly, two characterized grapevine STSs show a separated relationship from *PhSTS* with high bootstrap support. This suggests that these two enzymes evolved from different ancestors may be caused by neo-functionalization of species-specific tandem gene duplications. Nevertheless, a major part of the two displays low bootstrap percentage, which suggests that the CHS and STS phylogeny of the current studied plant species is highly assorted and complex; as such, further investigation is imperative. Apart from genomics investigations, transcriptomic studies are also important for metabolomics studies and crop improvement. The gene expression profiles of various peanut tissues have been revealed (Clevenger et al., 2016). Thus, we attempted to link our synteny analysis with this previously published gene expression data. *phSTS* (XP_025646382.1, ABY86219.1) is expressed in fruits of peanuts. Obviously, XP_025642791 (gene ID. Arahy.0FI6RG) which is pinpointed on region b is mainly expressed during the development of peanut seeds (**Figure 5D**). Furthermore, a co-expression network analysis of these genes is presented in **Figure 6**. *phSTS* shows a highly correlated expression with many tandem gene-duplicated CHSs, genes encoding CHI, and *O*-methyltransferase (OMT). On the other hand, Arahy.0FI6RG is correlated with genes related to several types of enzymes, including hydrolases, oxidoreductases, epimerases, and kinases. Moreover, some of these genes are involved in the regulation of transcription factors and transporters. The functions of these candidate genes need to be evaluated experimentally; however, our example of the *in silico* omics-based approach provides insights for future researches into the elucidation of key genes involved in nut polyphenols for metabolomics-assisted breeding approaches aimed at enhancing health-beneficial components.

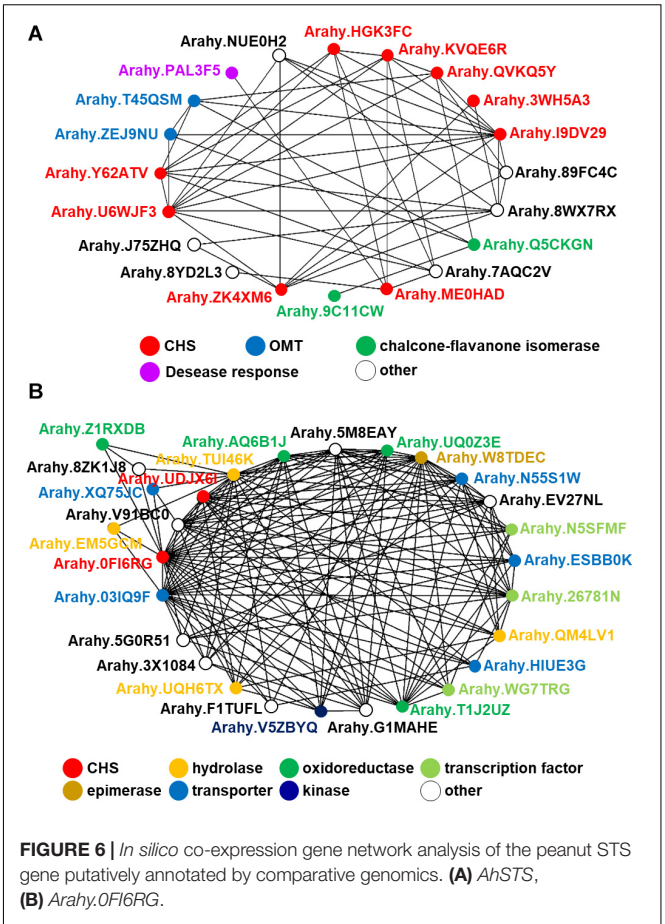

FIGURE 6 | *In silico* co-expression gene network analysis of the peanut STS gene putatively annotated by comparative genomics. **(A)** *AhSTS*, **(B)** *Arahy.0FI6RG*.

SUMMARY AND FUTURE PERSPECTIVE

Nuts are regarded as treasured food crops due to their high contents of potential bioactive components which are able to promote human health benefit. In our summary of the chemical diversity of nut polyphenols, flavonoids are found as the major structurally diversified polyphenols in both aglycone and decorated forms among seeds of nut plant species. With regard to the latter, the glycoside is the main category of polyphenolic derivatives. Diversification of chemical structures results in different effectiveness of biological activities, particularly antioxidants. Although the polyphenolic biosynthetic pathway is widely known, genes of nut plant species encoding enzymes responsible for each step remain uncharacterized. Genome synteny analysis of CHS and STS provides a strategic example for understanding the evolution and conservation of these two enzymes in seeds of nut plant species. Notably, there are several research gaps for nut plant species since much of our knowledge is fragmentary and considerable further investigation is required. Deciphering the multi-omics (genomics, transcriptomics, proteomics, and metabolomics) of nut plant species will provide fundamental data for their physiological function and potential for crop improvement, including increasing crop yield, stress, and disease tolerance,

as well as enhance the production of human health beneficial specialized metabolites.

AUTHOR CONTRIBUTIONS

CA and TT wrote the manuscript. CA, MW, and TT conceived, designed, and conceptualized the outline of the manuscript. TS, CA, and TN performed comparative genome analysis. VS, SB, MW, AF, and TT supervised and edited the manuscript. All authors have read and approved the final manuscript.

FUNDING

Research activity of TT was additionally supported by the JSPS KAKENHI Grant-in-Aid for Scientific Research B (19H03249) and C (19K06723).

ACKNOWLEDGMENTS

CA would like to specially thank Mahidol Medical Scholars Program (MSP), Mahidol University, Thailand, for the opportunity to work on this review article at Graduate School of Biological Science, Nara Institute of Science and Technology (NAIST), Ikoma, Japan. SB and VS thank Faculty of Pharmacy, Mahidol University. TN and AF would like to thank the funding from the Max-Planck-Society. TT, MW, and TS gratefully acknowledge the partial support by the NAIST.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.642581/full#supplementary-material

REFERENCES

Abdulkhaleq, L. A., Assi, M. A., Noor, M. H. M., Abdullah, R., Saad, M. Z., and Taufiq-Yap, Y. H. (2017). Therapeutic uses of epicatechin in diabetes and cancer. *Vet. World* 10, 869–872. doi: 10.14202/vetworld.2017.869-872

Afendi, F. M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., et al. (2011). KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.* 53:e1. doi: 10.1093/pcp/pcr165

Akinwumi, B., Bordun, K.-A., and Anderson, H. (2018). Biological activities of stilbenoids. *Int. J. Mol. Sci.* 19:792. doi: 10.3390/ijms19030792

Alasalvar, C., and Bolling, B. W. (2015). Review of nut phytochemicals, fat-soluble bioactives, antioxidant components and health effects. *Br. J. Nutr.* 113, S68–S78. doi: 10.1017/s0007114514003729

Alioto, T., Alexiou, K. G., Bardil, A., Barteri, F., Castanera, R., Cruz, F., et al. (2020). Transposons played a major role in the diversification between the closely related almond and peach genomes: results from the almond genome sequence. *Plant J.* 101, 455–472. doi: 10.1111/tpj.14538

Alseekh, S., Ofner, I., Liu, Z. Y., Osorio, S., Vallarino, J., Last, R. L., et al. (2020a). Quantitative trait loci analysis of seed-specialized metabolites reveals seed-specific flavonols and differential regulation of glycoalkaloid content in tomato. *Plant J.* 103, 2007–2024. doi: 10.1111/tpj.14879

Alseekh, S., Perez de Souza, L., Benina, M., and Fernie, A. R. (2020b). The style and substance of plant flavonoid decoration; towards defining both structure and function. *Phytochemistry* 174:112347. doi: 10.1016/j.phytochem.2020.112347

Alvarenga, T. A., De Oliveira, P. F., De Souza, J. M., Tavares, D. C., Andrade, E., Silva, M. L., et al. (2016). Schistosomicidal activity of alkyl-phenols from the cashew anacardium occidentale against *Schistosoma mansoni* adult worms. *J. Agric. Food Chem.* 64, 8821–8827. doi: 10.1021/acs.jafc.6b04200

Anand David, A. V., Arulmoli, R., and Parasuraman, S. (2016). Overviews of biological importance of Quercetin: a bioactive Flavonoid. *Pharmacogn. Rev.* 10, 84–89. doi: 10.4103/0973-7847.194044

Anantharaju, P. G., Gowda, P. C., Vimalambike, M. G., and Madhunapantula, S. V. (2016). An overview on the role of dietary phenolics for the treatment of cancers. *Nutr. J.* 15:99. doi: 10.1186/s12937-016-0217-2

Bae, J., Kim, N., Shin, Y., Kim, S.-Y., and Kim, Y.-J. (2020). Activity of catechins and their applications. *Biomed. Dermatol.* 4:8. doi: 10.1186/s41702-020-0057-8

Bajpai, V. K., Shukla, S., Paek, W. K., Lim, J., Kumar, P., Kumar, P., et al. (2017). Efficacy of (+)-Lariciresinol to control bacterial growth of *Staphylococcus aureus* and *Escherichia coli* O157:H7. *Front. Microbiol.* 8:804. doi: 10.3389/fmicb.2017.00804

Ballard, T. S., Mallikarjunan, P., Zhou, K., and O'Keefe, S. (2010). Microwave-assisted extraction of phenolic antioxidant compounds from peanut skins. *Food Chem.* 120, 1185–1192. doi: 10.1016/j.foodchem.2009.11.063

Ballistreri, G., Arena, E., and Fallico, B. (2009). Influence of ripeness and drying process on the Polyphenols and Tocopherols of *Pistacia vera* L. *Molecules* 14, 4358–4369. doi: 10.3390/molecules14114358

Barker, D. (2019). Lignans. *Molecules* 24:1424. doi: 10.3390/molecules24071424

Bernard, A., Marrano, A., Donkpegan, A., Brown, P. J., Leslie, C. A., Neale, D. B., et al. (2020). Association and linkage mapping to unravel genetic architecture of phenological traits and lateral bearing in Persian walnut (*Juglans regia* L.). *BMC Genom.* 21:203. doi: 10.1186/s12864-020-6616-y

Bernatova, I. (2018). Biological activities of (–)-epicatechin and (–)-epicatechin-containing foods: focus on cardiovascular and neuropsychological health. *Biotechnol. Adv.* 36, 666–681. doi: 10.1016/j.biotechadv.2018.01.009

Bertioli, D. J., Cannon, S. B., Froenicke, L., Huang, G., Farmer, A. D., Cannon, E. K. S., et al. (2016). The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat. Genet.* 48, 438–446. doi: 10.1038/ng.3517

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., et al. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nat. Genet.* 51, 877–884. doi: 10.1038/s41588-019-0405-z

Bisby, F. (1994). *Phytochemical Dictionary of the Leguminosae*. New York, NY: Chapman & Hall.

Blažević, I., Montaut, S., Burčul, F., Olsen, C. E., Burow, M., Rollin, P., et al. (2020). Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry* 169:112100. doi: 10.1016/j.phytochem.2019.112100

Bodoira, R., and Maestri, D. (2020). Phenolic compounds from nuts: extraction, chemical profiles, and bioactivity. *J. Agric. Food Chem.* 68, 927–942. doi: 10.1021/acs.jafc.9b07160

Bodoira, R., Rossi, Y., Montenegro, M., Maestri, D., and Velez, A. (2017). Extraction of antioxidant polyphenolic compounds from peanut skin using water-ethanol at high pressure and temperature conditions. *J. Supercrit. Fluids* 128, 57–65. doi: 10.1016/j.supflu.2017.05.011

Bodoira, R., Velez, A., Rovetto, L., Ribotta, P., Maestri, D., and Martínez, M. (2019). Subcritical fluid extraction of antioxidant phenolic compounds from pistachio (*Pistacia vera* L.) Nuts: experiments, modeling, and optimization. *J. Food Sci.* 84, 963–970. doi: 10.1111/1750-3841.14507

Bolling, B. W. (2017). Almond polyphenols: methods of analysis, contribution to food quality, and health promotion. *Compr. Rev. Food Sci. Food Saf.* 16, 346–368. doi: 10.1111/1541-4337.12260

Bolling, B. W., Chen, C. Y. O., McKay, D. L., and Blumberg, J. B. (2011). Tree nut phytochemicals: composition, antioxidant capacity, bioactivity, impact factors. A systematic review of almonds, Brazils, cashews, hazelnuts, macadamias, pecans, pine nuts, pistachios and walnuts. *Nutr. Res. Rev.* 24, 244–275. doi: 10.1017/s095442241100014x

Bolling, B. W., Dolnikowski, G., Blumberg, J. B., and Chen, C. Y. O. (2010). Polyphenol content and antioxidant activity of California almonds depend on

cultivar and harvest year. *Food Chem.* 122, 819–825. doi: 10.1016/j.foodchem.2010.03.068

Butelli, E., Titta, L., Giorgio, M., Mock, H. P., Matros, A., Peterek, S., et al. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.* 26, 1301–1308. doi: 10.1038/nbt.1506

Chang, X., Liu, F., Lin, Z., Qiu, J., Peng, C., Lu, Y., et al. (2020). Phytochemical profiles and cellular antioxidant activities in chestnut (*Castanea mollissima* BL.) kernels of five different cultivars. *Molecules* 25:178. doi: 10.3390/molecules25010178

Chen, C. Y. O., Milbury, P. E., and Blumberg, J. B. (2019). Polyphenols in almond skins after blanching modulate plasma biomarkers of oxidative stress in healthy humans. *Antioxidants* 8:95. doi: 10.3390/antiox8040095

Chen, X., Lu, Q., Liu, H., Zhang, J., Hong, Y., Lan, H., et al. (2019). Sequencing of cultivated peanut, *Arachis hypogaea*, yields insights into genome evolution and oil improvement. *Mol. Plant* 12, 920–934. doi: 10.1016/j.molp.2019.03.005

Chen, N., Yang, Q., Pan, L., Chi, X., Chen, M., Hu, D., et al. (2014). Identification of 30 MYB transcription factor genes and analysis of their expression during abiotic stress in peanut (*Arachis hypogaea* L.). *Gene* 533, 332–345. doi: 10.1016/j.gene.2013.08.092

Cheynier, V., Comte, G., Davies, K. M., Lattanzio, V., and Martens, S. (2013). Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiol. Biochem.* 72, 1–20. doi: 10.1016/j.plaphy.2013.05.009

Cho, J.-Y., Moon, J.-H., Seong, K.-Y., and Park, K.-H. (1998). Antimicrobial activity of 4-hydroxybenzoic acid and trans 4-hydroxycinnamic acid isolated and identified from rice hull. *Biotechnol. Biochem.* 62, 2273–2276. doi: 10.1271/bbb.62.2273

Choi, S.-W., Park, K.-I., Yeon, J.-T., Ryu, B. J., Kim, K.-J., and Kim, S. H. (2014). Anti-osteoclastogenic activity of matairesinol via suppression of p38/ERK-NFATc1 signaling axis. *BMC Complement. Altern. Med.* 14:35. doi: 10.1186/1472-6882-14-35

Cipolletti, M., Solar Fernandez, V., Montalesi, E., Marino, M., and Fiocchetti, M. (2018). Beyond the antioxidant activity of dietary polyphenols in cancer: the modulation of estrogen receptors (ERs) signaling. *Int. J. Mol. Sci.* 19:2624. doi: 10.3390/ijms19092624

Clevenger, J., Chu, Y., Scheffler, B., and Ozias-Akins, P. (2016). A developmental Transcriptome map for Allotetraploid *Arachis hypogaea*. *Front. Plant Sci.* 7:1446. doi: 10.3389/fpls.2016.01446

Čolić, S. D., Fotirić Akšić, M. M., Lazarević, K. B., Zec, G. N., Gašić, U. M., Dabić Zagorac, D. Č, et al. (2017). Fatty acid and phenolic profiles of almond grown in Serbia. *Food Chem.* 234, 455–463. doi: 10.1016/j.foodchem.2017.05.006

Condori, J., Medrano, G., Sivakumar, G., Nair, V., Cramer, C., and Medina-Bolivar, F. (2009). Functional characterization of a stilbene synthase gene using a transient expression system in planta. *Plant Cell Rep.* 28, 589–599. doi: 10.1007/s00299-008-0664-0

Corso, M., Perreau, F., Mouille, G., and Lepiniec, L. (2020). Specialized phenolic compounds in seeds: structures, functions, and regulations. *Plant Sci.* 296:110471. doi: 10.1016/j.plantsci.2020.110471

Cory, H., Passarelli, S., Szeto, J., Tamez, M., and Mattei, J. (2018). The role of polyphenols in human health and food systems: a mini-review. *Front. Nutr.* 5:87. doi: 10.3389/fnut.2018.00087

De La Rosa, L. A., Alvarez-Parrilla, E., and Shahidi, F. (2011). Phenolic compounds and antioxidant activity of kernels and shells of mexican pecan (*Carya illinoinensis*). *J. Agric. Food Chem.* 59, 152–162. doi: 10.1021/jf1034306

De Souza, R. G. M., Schincaglia, R. M., Pimentel, G. D., and Mota, J. F. (2017). Nuts and human health outcomes: a systematic review. *Nutrients* 9:1311. doi: 10.3390/nu9121311

Del Rio, D., Calani, L., Dall'Asta, M., and Brighenti, F. (2011). Polyphenolic composition of hazelnut skin. *J. Agric. Food Chem.* 59, 9935–9941. doi: 10.1021/jf202449z

Deng, Z., Hassan, S., Rafiq, M., Li, H., He, Y., Cai, Y., et al. (2020). Pharmacological activity of Eriodictyol: the major natural Polyphenolic Flavanone. *Evid. Based Complement. Altern. Med.* 2020:6681352. doi: 10.1155/2020/6681352

Dong, H., Yang, X., He, J., Cai, S., Xiao, K., and Zhu, L. (2017). Enhanced antioxidant activity, antibacterial activity and hypoglycemic effect of luteolin by complexation with manganese(ii) and its inhibition kinetics on xanthine oxidase. *RSC Adv.* 7, 53385–53395. doi: 10.1039/C7RA11036G

El-Shitany, N. A., El-Bastawissy, E. A., and El-Desoky, K. (2014). Ellagic acid protects against carrageenan-induced acute inflammation through inhibition of nuclear factor kappa B, inducible cyclooxygenase and proinflammatory cytokines and enhancement of interleukin-10 via an antioxidant mechanism. *Int. Immunopharmacol.* 19, 290–299. doi: 10.1016/j.intimp.2014.02.004

Fabani, M. P., Luna, L., Baroni, M. V., Monferran, M. V., Ighani, M., Tapia, A., et al. (2013). Pistachio (*Pistacia vera* var Kerman) from argentinean cultivars. A natural product with potential to improve human health. *J. Funct. Foods* 5, 1347–1356. doi: 10.1016/j.jff.2013.05.002

Falasca, M., Casari, I., and Maffucci, T. (2014). Cancer chemoprevention with nuts. *J. Natl. Cancer Inst.* 106:dju238. doi: 10.1093/jnci/dju238

Falcone Ferreyra, M. L., Rius, S. P., and Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3:222. doi: 10.3389/fpls.2012.00222

Fernie, A. R., and Tohge, T. (2017). The genetics of plant metabolism. *Annu. Rev. Genet.* 51, 287–310. doi: 10.1146/annurev-genet-120116-024640

Fukuda, T., Ito, H., and Yoshida, T. (2003). Antioxidative polyphenols from walnuts (*Juglans regia* L.). *Phytochemistry* 63, 795–801. doi: 10.1016/s0031-9422(03)00333-9

Ganeshpurkar, A., and Saluja, A. K. (2017). The pharmacological potential of Rutin. *Saudi Pharm. J. SPJ Off. Publ. Saudi Pharm. Soc.* 25, 149–164. doi: 10.1016/j.jsps.2016.04.025

Gao, Z., Xu, H., and Huang, K. (2002). Effects of rutin supplementation on antioxidant status and iron, copper, and zinc contents in mouse liver and brain. *Biol. Trace Elem. Res.* 88, 271–279.

Gil, M., and Wianowska, D. (2017). Chlorogenic acids - their properties, occurrence and analysis. *Ann. UMCS Chem.* 72:61. doi: 10.17951/aa.2017.72.1.61

Grace, M. H., Warlick, C. W., Neff, S. A., and Lila, M. A. (2014). Efficient preparative isolation and identification of walnut bioactive components using high-speed counter-current chromatography and LC-ESI-IT-TOF-MS. *Food Chem.* 158, 229–238. doi: 10.1016/j.foodchem.2014.02.117

Gulcin, İ (2020). Antioxidants and antioxidant methods: an updated overview. *Arch. Toxicol.* 94, 651–715. doi: 10.1007/s00204-020-02689-3

Han, Y.-J., Ren, Z.-G., Li, X.-X., Yan, J.-L., Ma, C.-Y., Wu, D.-D., et al. (2020). Advances and challenges in the prevention and treatment of COVID-19. *Int. J. Med. Sci.* 17, 1803–1810. doi: 10.7150/ijms.47836

Hano, C., and Tungmunnithum, D. (2020). Plant Polyphenols, more than just simple natural antioxidants: oxidative stress, aging and age-related diseases. *Medicines* 7:26. doi: 10.3390/medicines7050026

Heleno, S. A., Martins, A., Queiroz, M. J. R. P., and Ferreira, I. C. F. R. (2015). Bioactivity of phenolic acids: metabolites versus parent compounds: a review. *Food Chem.* 173, 501–513. doi: 10.1016/j.foodchem.2014.10.057

Horvath, G., Wessjohann, L., Bigirimana, J., Jansen, M., Guisez, Y., Caubergs, R., et al. (2006). Differential distribution of tocopherols and tocotrienols in photosynthetic and non-photosynthetic tissues. *Phytochemistry* 67, 1185–1195. doi: 10.1016/j.phytochem.2006.04.004

Huang, Y., Xiao, L., Zhang, Z., Zhang, R., Wang, Z., Huang, C., et al. (2019). The genomes of pecan and Chinese hickory provide insights into Carya evolution and nut nutrition. *Gigascience* 8:giz036. doi: 10.1093/gigascience/giz036

Hwang, B., Cho, J., Hwang, I., Jin, H.-G., Woo, E.-R., and Lee, D. G. (2011). Antifungal activity of lariciresinol derived from *Sambucus williamsii* and their membrane-active mechanisms in *Candida albicans*. *Biochem. Biophys. Res. Commun.* 410, 489–493. doi: 10.1016/j.bbrc.2011.06.004

Isemura, M. (2019). Catechin in human health and disease. *Molecules* 24:528. doi: 10.3390/molecules24030528

Jia, X., Luo, H., Xu, M., Zhai, M., Guo, Z., Qiao, Y., et al. (2018). Dynamic Changes in phenolics and antioxidant capacity during pecan (*Carya illinoinensis*) kernel ripening and its phenolics profiles. *Molecules* 23:435. doi: 10.3390/molecules23020435

Jiang, X., Liu, Y., Li, W., Zhao, L., Meng, F., Wang, Y., et al. (2013). Tissue-specific, development-dependent phenolic compounds accumulation profile and gene expression pattern in tea plant [*Camellia sinensis*]. *PLoS One* 8:e62315. doi: 10.1371/journal.pone.0062315

Kahkeshani, N., Farzaei, F., Fotouhi, M., Alavi, S. S., Bahramsoltani, R., Naseri, R., et al. (2019). Pharmacological effects of gallic acid in health and diseases: a mechanistic review. *Iran. J. Basic Med. Sci.* 22, 225–237. doi: 10.22038/ijbms.2019.32806.7897

Kawahara, S., Iwata, I., Fujita, E., Yamawaki, M., Nishiwaki, H., Sugahara, T., et al. (2010). IgE-suppressive activity of (À)-matairesinol and its structure-activity relationship. *Biosci. Biotechnol. Biochem.* 74, 1878–1883. doi: 10.1271/bbb.100275

Kessler, A., and Kalske, A. (2018). Plant secondary metabolite diversity and species interactions. *Annu. Rev. Ecol. Evol. Syst.* 49, 115–138. doi: 10.1146/annurev-ecolsys-110617-062406

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Küpeli Akkol, E., Genç, Y., Karpuz, B., Sobarzo-Sánchez, E., and Capasso, R. (2020). Coumarins and coumarin-related compounds in pharmacotherapy of cancer. *Cancers* 12:1959. doi: 10.3390/cancers12071959

Kuršvietienė, L., Stanevičienė, I., Mongirdienė, A., and Bernatonienė, J. (2016). Multiplicity of effects and health benefits of resveratrol. *Medicina* 52, 148–155. doi: 10.1016/j.medici.2016.03.003

Landete, J. M. (2011). Ellagitannins, ellagic acid and their derived metabolites: a review about source, metabolism, functions and health. *Food Res. Int.* 44, 1150–1160. doi: 10.1016/j.foodres.2011.04.027

Lanz, T., Tropf, S., Marner, F. J., Schröder, J., and Schröder, G. (1991). The role of cysteines in polyketide synthases. Site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. *J. Biol. Chem.* 266, 9971–9976.

Lautié, E., Russo, O., Ducrot, P., and Boutin, J. A. (2020). Unraveling plant natural chemical diversity for drug discovery purposes. *Front. Pharmacol.* 11:397. doi: 10.3389/fphar.2020.00397

Lee, M., Rho, H. S., and Choi, K. (2019). Anti-inflammatory effects of a p-coumaric acid and Kojic acid derivative in LPS-stimulated RAW264.7 macrophage cells. *Biotechnol. Bioprocess Eng.* 24, 653–657. doi: 10.1007/s12257-018-0492-1

Lepiniec, L., Debeaujon, I., Routaboul, J.-M., Baudry, A., Pourcel, L., Nesi, N., et al. (2006). Genetics and biochemistry of seed flavonoids. *Annu. Rev. Plant Biol.* 57, 405–430. doi: 10.1146/annurev.arplant.57.032905.105252

Li, W., Du, Q., Li, X., Zheng, X., Lv, F., Xi, X., et al. (2020). Eriodictyol inhibits proliferation, metastasis and induces apoptosis of glioma cells via PI3K/Akt/NF-κB signaling pathway. *Front. Pharmacol.* 11:114. doi: 10.3389/fphar.2020.00114

Lin, Y., Shi, R., Wang, X., and Shen, H.-M. (2008). Luteolin, a flavonoid with potential for cancer prevention and therapy. *Curr. Cancer Drug Targets* 8, 634–646. doi: 10.2174/156800908786241050

Liu, Y., Zhao, S., Wang, J., Zhao, C., Guan, H., Hou, L., et al. (2015). Molecular cloning, expression, and evolution analysis of type II CHI gene from peanut (*Arachis hypogaea* L.). *Dev. Genes Evol.* 225, 1–10. doi: 10.1007/s00427-015-0489-0

Lou, H., Yamazaki, Y., Sasaki, T., Uchida, M., Tanaka, H., and Oka, A. (1999). A-type proanthocyanidins from peanut skins. *Phytochemistry* 51, 297–308. doi: 10.1016/S0031-9422(98)00736-5

Lou, H., Yuan, H., Ma, B., Ren, D., Ji, M., and Oka, S. (2004). Polyphenols from peanut skins and their free radical-scavenging effects. *Phytochemistry* 65, 2391–2399. doi: 10.1016/j.phytochem.2004.06.026

Lucas, S. J., Kahraman, K., Avşar, B., Buggs, R. J. A., and Bilge, I. (2020). A chromosome-scale genome assembly of European hazel (*Corylus avellana* L.) reveals targets for crop improvement. *Plant J.* 105, 1413–1430. doi: 10.1111/tpj.15099

Luo, Y., Shang, P., and Li, D. (2017). Luteolin: a flavonoid that has multiple cardio-protective effects and its molecular mechanisms. *Front. Pharmacol.* 8:692. doi: 10.3389/fphar.2017.00692

Lushchak, V. I., and Semchuk, N. M. (2012). Tocopherol biosynthesis: chemistry, regulation and effects of environmental factors. *Acta Physiol. Plant* 34, 1607–1628. doi: 10.1007/s11738-012-0988-9

Ma, Y., Kosińska-Cagnazzo, A., Kerr, W. L., Amarowicz, R., Swanson, R. B., and Pegg, R. B. (2014). Separation and characterization of phenolic compounds from dry-blanched peanut skins by liquid chromatography-electrospray ionization mass spectrometry. *J. Chromatogr.* 1356, 64–81. doi: 10.1016/j.chroma.2014.06.027

Maia, F. J. N., Ribeiro, F. W. P., Rangel, J. H. G., Lomonaco, D., Luna, F. M. T., de Lima-Neto, P., et al. (2015). Evaluation of antioxidant action by electrochemical and accelerated oxidation experiments of phenolic compounds derived from cashew nut shell liquid. *Ind. Crops Prod.* 67, 281–286. doi: 10.1016/j.indcrop.2015.01.034

Marrano, A., Britton, M., Zaini, P. A., Zimin, A. V., Workman, R. E., Puiu, D., et al. (2020). High-quality chromosome-scale assembly of the walnut (*Juglans regia* L.) reference genome. *Gigascience* 9:giaa050. doi: 10.1093/gigascience/giaa050

Marrano, A., Martínez-García, P. J., Bianco, L., Sideli, G. M., Di Pierro, E. A., Leslie, C. A., et al. (2019). A new genomic tool for walnut (*Juglans regia* L.): development and validation of the high-density Axiom™ J. regia 700K SNP genotyping array. *Plant Biotechnol. J.* 17, 1027–1036. doi: 10.1111/pbi.13034

Martínez-García, P. J., Crepeau, M. W., Puiu, D., Gonzalez-Ibeas, D., Whalen, J., Stevens, K. A., et al. (2016). The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of non-structural polyphenols. *Plant J.* 87, 507–532. doi: 10.1111/tpj.13207

Mathesius, U. (2018). Flavonoid functions in plants and their interactions with other organisms. *Plants* 7:30. doi: 10.3390/plants7020030

Milbury, P. E., Chen, C.-Y., Dolnikowski, G. G., and Blumberg, J. B. (2006). Determination of Flavonoids and Phenolics and their distribution in almonds. *J. Agric. Food Chem.* 54, 5027–5033. doi: 10.1021/jf0603937

Monagas, M., Garrido, I., Lebrón-Aguilar, R., Bartolome, B., and Gómez-Cordovés, C. (2007). Almond (*Prunus dulcis*(Mill.) D.A. Webb) skins as a potential source of bioactive polyphenols. *J. Agric. Food Chem.* 55, 8498–8507. doi: 10.1021/jf071780z

Morais, S. M., Silva, K. A., Araujo, H., Vieira, I. G. P., Alves, D. R., Fontenelle, R. O. S., et al. (2017). Anacardic acid constituents from cashew nut shell liquid: nmr characterization and the effect of unsaturation on its biological activities. *Pharmaceuticals* 10:31.

Muroi, H., and Kubo, I. (1996). Antibacterial activity of anacardic acid and totarol, alone and in combination with methicillin, against methicillinresistant *Staphylococcus aureus*. *J. Appl. Bacteriol.* 80, 387–394. doi: 10.1111/j.1365-2672.1996.tb03233.x

Musarra-Pizzo, M., Ginestra, G., Smeriglio, A., Pennisi, R., Sciortino, M. T., and Mandalari, G. (2019). The antimicrobial and antiviral activity of polyphenols from almond (*Prunus dulcis* L.) Skin. *Nutrients* 11:2355. doi: 10.3390/nu11102355

Naake, T., Maeda, H. A., Proost, S., Tohge, T., and Fernie, A. R. (2020). Kingdom-wide analysis of the evolution of the plant type III polyketide synthase superfamily. *Plant Physiol.* kiaa086. doi: 10.1093/plphys/kiaa086

Navarro, G., Martínez-Pinilla, E., Ortiz, R., Noé, V., Ciudad, C. J., and Franco, R. (2018). Resveratrol and related Stilbenoids, Nutraceutical/dietary complements with health-promoting actions: industrial production, safety, and the search for mode of action. *Compr. Rev. Food Sci. Food Saf.* 17, 808–826. doi: 10.1111/1541-4337.12359

Nayak, B., Liu, R. H., and Tang, J. (2015). Effect of processing on phenolic antioxidants of fruits, vegetables, and grains—a review. *Crit. Rev. Food Sci. Nutr.* 55, 887–918. doi: 10.1080/10408398.2011.654142

Ngwa, W., Kumar, R., Thompson, D., Lyerly, W., Moore, R., Reid, T.-E., et al. (2020). Potential of flavonoid-inspired Phytomedicines against COVID-19. *Molecules* 25:2707. doi: 10.3390/molecules25112707

Pais, A. L., Li, X., and Jenny Xiang, Q.-Y. (2018). Discovering variation of secondary metabolite diversity and its relationship with disease resistance in *Cornus florida* L. *Ecol. Evol.* 8, 5619–5636. doi: 10.1002/ece3.4090

Palazzolo, G., Horvath, P., and Zenobi-Wong, M. (2012). The flavonoid isoquercitrin promotes neurite elongation by reducing RhoA activity. *PLoS One* 7:e49979. doi: 10.1371/journal.pone.0049979

Pelvan, E., Olgun, E. Ö, Karadağ, A., and Alasalvar, C. (2018). Phenolic profiles and antioxidant activity of Turkish Tombul hazelnut samples (natural, roasted, and roasted hazelnut skin). *Food Chem.* 244, 102–108. doi: 10.1016/j.foodchem.2017.10.011

Peng, M., Shahzad, R., Gul, A., Subthain, H., Shen, S. Q., Lei, L., et al. (2017). Differentially evolved glucosyltransferases determine natural variation of rice flavone accumulation and UV-tolerance. *Nat. Commun.* 8:1975. doi: 10.1038/s41467-017-02168-x

Pugazhendhi, D., Pope, G. S., and Darbre, P. D. (2005). Oestrogenic activity of p-hydroxybenzoic acid (common metabolite of paraben esters) and methylparaben in human breast cancer cell lines. *J. Appl. Toxicol.* 25, 301–309. doi: 10.1002/jat.1066

Puksasook, T., Kimura, S., Tadtong, S., Jiaranaikulwanitch, J., Pratuangdejkul, J., Kitphati, W., et al. (2017a). Erratum to: semisynthesis and biological evaluation of prenylated resveratrol derivatives as multi-targeted agents for Alzheimer's disease. *J. Nat. Med.* 71, 683–684. doi: 10.1007/s11418-017-1110-9

Puksasook, T., Kimura, S., Tadtong, S., Jiaranaikulwanitch, J., Pratuangdejkul, J., Kitphati, W., et al. (2017b). Semisynthesis and biological evaluation of prenylated resveratrol derivatives as multi-targeted agents for Alzheimer's disease. *J. Nat. Med.* 71, 665–682. doi: 10.1007/s11418-017-1097-2

Regueiro, J., Sánchez-González, C., Vallverdú-Queralt, A., Simal-Gándara, J., Lamuela-Raventós, R., and Izquierdo-Pulido, M. (2014). Comprehensive identification of walnut polyphenols by liquid chromatography coupled to linear ion trap-Orbitrap mass spectrometry. *Food Chem.* 152, 340–348. doi: 10.1016/j.foodchem.2013.11.158

Renaud, J., and Martinoli, M.-G. (2019). Considerations for the use of polyphenols as therapies in neurodegenerative diseases. *Int. J. Mol. Sci.* 20:1883. doi: 10.3390/ijms20081883

Robbins, K. S., Ma, Y., Wells, M. L., Greenspan, P., and Pegg, R. B. (2014). Separation and characterization of phenolic compounds from U.S. pecans by liquid chromatography-tandem mass spectrometry. *J. Agric. Food Chem.* 62, 4332–4341. doi: 10.1021/jf500909h

Rodríguez-García, C., Sánchez-Quesada, C., Toledo, E., Delgado-Rodríguez, M., and Gaforio, J. (2019). Naturally Lignan-rich foods: a dietary tool for health promotion? *Molecules* 24:917. doi: 10.3390/molecules24050917

Ros, E. (2010). Health benefits of nut consumption. *Nutrients* 2, 652–682. doi: 10.3390/nu2070652

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

Salehi, B., Gültekin, ÖM., Kırkın, C., Özçelik, B., Morais, B., Carneiro, F. C., et al. (2019). Anacardium plants: chemical,nutritional composition and biotechnological applications. *Biomolecules* 9:465. doi: 10.3390/biom9090465

Salehi, B., Mishra, A., Nigam, M., Sener, B., Kilic, M., Sharifi-Rad, M., et al. (2018). Resveratrol: a double-edged sword in health benefits. *Biomedicines* 6:91. doi: 10.3390/biomedicines6030091

Sánchez-Pérez, R., Pavan, S., Mazzeo, R., Moldovan, C., Aiese Cigliano, R., Del Cueto, J., et al. (2019). Mutation of a bHLH transcription factor allowed almond domestication. *Science* 364, 1095–1098. doi: 10.1126/science.aav8197

Sarnoski, P. J., Johnson, J. V., Reed, K. A., Tanko, J. M., O'Keefe, S. F. (2012). Separation and characterisation of proanthocyanidins in Virginia type peanut skins by LC-MSn. *Food Chem.* 131, 927–939. doi: 10.1016/j.foodchem.2011.09.081

Sato, Y., Itagaki, S., Kurokawa, T., Ogura, J., Kobayashi, M., Hirano, T., et al. (2011). In vitro and in vivo antioxidant properties of chlorogenic acid and caffeic acid. *Int. J. Pharm.* 403, 136–138. doi: 10.1016/j.ijpharm.2010.09.035

Schneider, B. U. C., Meza, A., Beatriz, A., Pesarini, J. R., Carvalho, P. C., de Mauro, M., et al. (2016). Cardanol: toxicogenetic assessment and its effects when combined with cyclophosphamide. *Genet. Mol. Biol.* 39, 279–289.

Schröder, G., and Schröder, G. (1990). Stilbene and chalcone synthases: related enzymes with key functions in plant-specific pathways. *Z. Naturforsch. C. J. Biosci.* 45, 1–8. doi: 10.1515/znc-1990-1-202

Scossa, F., Roda, F., Tohge, T., Georgiev, M. I., and Fernie, A. R. (2019). The hot and the colorful: Understanding the metabolism, genetics and evolution of consumer preferred metabolic traits in pepper and related species. *CRC Crit. Rev. Plant Sci.* 38, 339–381. doi: 10.1080/07352689.2019.1682791

Shahidi, F., and Ambigaipalan, P. (2015). Phenolics and polyphenolics in foods, beverages and spices: antioxidant activity and health effects - a review. *J. Funct. Foods* 18, 820–897. doi: 10.1016/j.jff.2015.06.018

Shen, Y., Liang, W.-J., Shi, Y.-N., Kennelly, E. J., and Zhao, D.-K. (2020). Structural diversity, bioactivities, and biosynthesis of natural diterpenoid alkaloids. *Nat. Prod. Rep.* 37, 763–796. doi: 10.1039/d0np00002g

Shen, Y., Song, X., Li, L., Sun, J., Jaiswal, Y., Huang, J., et al. (2019). Protective effects of p-coumaric acid against oxidant and hyperlipidemia-an in vitro and in vivo evaluation. *Biomed. Pharmacother.* 111, 579–587. doi: 10.1016/j.biopha.2018.12.074

Slatnar, A., Mikulic-Petkovsek, M., Stampar, F., Veberic, R., and Solar, A. (2015). Identification and quantification of phenolic compounds in kernels, oil and bagasse pellets of common walnut (*Juglans regia* L.). *Food Res. Int.* 67, 255–263. doi: 10.1016/j.foodres.2014.11.016

Soares, S., Brandão, E., Guerreiro, C., Soares, S., Mateus, N., and De Freitas, V. (2020). Tannins in food: insights into the molecular perception of astringency and bitter taste. *Molecules* 25:2590. doi: 10.3390/molecules25112590

Tohge, T., de Souza, L. P., and Fernie, A. R. (2014). Genome-enabled plant metabolomics. *J. Chromatogr. B* 966, 7–20. doi: 10.1016/j.jchromb.2014.04.003

Tohge, T., de Souza, L. P., and Fernie, A. R. (2017). Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. *J. Exp. Bot.* 68, 4013–4028. doi: 10.1093/jxb/erx177

Tohge, T., and Fernie, A. R. (2010). Combining genetic diversity, informatics and metabolomics to facilitate annotation of plant gene function. *Nat. Protoc.* 5, 1210–1227. doi: 10.1038/nprot.2010.82

Tohge, T., and Fernie, A. R. (2017). An overview of compounds derived from the Shikimate and Phenylpropanoid pathways and their medicinal importance. *Mini. Rev. Med. Chem.* 17, 1013–1027. doi: 10.2174/1389557516666160624123425

Tohge, T., Watanabe, M., Hoefgen, R., and Fernie, A. R. (2013a). Shikimate and phenylalanine biosynthesis in the green lineage. *Front. Plant Sci.* 4:62. doi: 10.3389/fpls.2013.00062

Tohge, T., Watanabe, M., Hoefgen, R., and Fernie, A. R. (2013b). The evolution of phenylpropanoid metabolism in the green lineage. *Crit. Rev. Biochem. Mol. Biol.* 48, 123–152. doi: 10.3109/10409238.2012.758083

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

Tomaino, A., Martorana, M., Arcoraci, T., Monteleone, D., Giovinazzo, C., and Saija, A. (2010). Antioxidant activity and phenolic profile of pistachio (*Pistacia vera* L., variety Bronte) seeds and skins. *Biochimie* 92, 1115–1122. doi: 10.1016/j.biochi.2010.03.027

Trevisan, M. T. S., Pfundstein, B., Haubner, R., Würtele, G., Spiegelhalder, B., Bartsch, H., et al. (2006). Characterization of alkyl phenols in cashew (*Anacardium occidentale*) products and assay of their antioxidant capacity. *Food Chem. Toxicol.* 44, 188–197. doi: 10.1016/j.fct.2005.06.012

Trox, J., Vadivel, V., Vetter, W., Stuetz, W., Kammerer, D. R., Carle, R., et al. (2011). Catechin and epicatechin in testa and their association with bioactive compounds in kernels of cashew nut (*Anacardium occidentale* L.). *Food Chem.* 128, 1094–1099. doi: 10.1016/j.foodchem.2011.04.018

Tungmunnithum, D., Thongboonyou, A., Pholboon, A., and Yangsabai, A. (2018). Flavonoids and other Phenolic compounds from medicinal plants for pharmaceutical and medical aspects: an overview. *Medicines* 5:93. doi: 10.3390/medicines5030093

Tuyen, P., Xuan, T., Khang, D., Ahmad, A., Quan, N., Tu Anh, T., et al. (2017). Phenolic compositions and antioxidant properties in Bark, Flower, Inner Skin, Kernel and Leaf extracts of *Castanea crenata* Sieb. et Zucc. *Antioxidants* 6:31. doi: 10.3390/antiox6020031

Valanciene, E., Jonuskiene, I., Syrpas, M., Augustiniene, E., Matulis, P., Simonavicius, A., et al. (2020). Advances and prospects of phenolic acids production, biorefinery and analysis. *Biomolecules* 10:874. doi: 10.3390/biom10060874

Valdés, A., Vidal, L., Beltrán, A., Canals, A., and Garrigós, M. C. (2015). Microwave-Assisted extraction of Phenolic compounds from almond skin byproducts (*Prunus amygdalus*): a multivariate analysis approach. *J. Agric. Food Chem.* 63, 5395–5402. doi: 10.1021/acs.jafc.5b01011

Valentová, K., Vrba, J., Bancířová, M., Ulrichová, J., and Křen, V. (2014). Isoquercitrin: pharmacology, toxicology, and metabolism. *Food Chem. Toxicol.* 68, 267–282. doi: 10.1016/j.fct.2014.03.018

Vinson, J. A., and Cai, Y. (2012). Nuts, especially walnuts, have both antioxidant quantity and efficacy and exhibit significant potential health benefits. *Food Funct.* 3, 134–140. doi: 10.1039/c2fo10152a

Vu, D. C., Vo, P. H., Coggeshall, M. V., and Lin, C.-H. (2018). Identification and characterization of phenolic compounds in black walnut kernels. *J. Agric. Food Chem.* 66, 4503–4511. doi: 10.1021/acs.jafc.8b01181

Wang, X., Zhang, Y., Xia, H., Yuan, M., Zhao, C., and Li, A. Q. (2012). Cloning and expression analysis of peanut (*Arachis hypogaea* L.) CHI gene. *Electron. J. Biotechnol.* 15:5. doi: 10.2225/vol15-issue1-fulltext-6

Wu, G. A., and Gmitter, F. G. (2019). Novel assembly strategy cracks open the mysteries of walnut genome evolution. *Hortic. Res.* 6:57. doi: 10.1038/s41438-019-0143-5

Xie, L., and Bolling, B. W. (2014). Characterisation of stilbenes in California almonds (*Prunus dulcis*) by UHPLC-MS. *Food Chem.* 148, 300–306. doi: 10.1016/j.foodchem.2013.10.057

Xu, D., Hu, M.-J., Wang, Y.-Q., and Cui, Y.-L. (2019). Antioxidant activities of quercetin and its complexes for medicinal application. *Molecules* 24:1123. doi: 10.3390/molecules24061123

Xu, Z., Meenu, M., Chen, P., and Xu, B. (2020). Comparative study on Phytochemical profiles and antioxidant capacities of chestnuts produced in different geographic area in China. *Antioxidants* 9:190. doi: 10.3390/antiox9030190

Yamauchi, S., Sugihara, T., Nakashima, Y., Okada, A., Akiyama, K., Kishida, T., et al. (2006). Radical and superoxide scavenging activities of Matairesinol and oxidized Matairesinol. *Biosci. Biotechnol. Biochem.* 70, 1934–1940. doi: 10.1271/bbb.60096

Yang, T., Fang, L., Sanders, S., Jayanthi, S., Rajan, G., Podicheti, R., et al. (2018). Stilbenoid prenyltransferases define key steps in the diversification of peanut phytoalexins. *J. Biol. Chem.* 293, 28–46. doi: 10.1074/jbc.ra117.000564

Yu, J., Ahmedna, M., Goktepe, I., and Dai, J. (2006). Peanut skin procyanidins: composition and antioxidant activities as affected by processing. *J. Food Comp. Anal.* 19, 364–371. doi: 10.1016/j.jfca.2005.08.003

Yuan, L., and Grotewold, E. (2020). Plant specialized metabolism. *Plant Sci.* 2020:110579. doi: 10.1016/j.plantsci.2020.110579

Zanwar, A. A., Badole, S. L., Shende, P. S., Hegde, M. V., and Bodhankar, S. L. (2014). "Chapter 80 - Role of gallic acid in cardiovascular disorders," in *Polyphenols in Human Health and Disease*, eds R. R. Watson, V. R. Preedy, and D. Zibadi (San Diego: Academic Press), 1045–1047. doi: 10.1016/B978-0-12-398456-2.00080-3

Zeng, L., Tu, X. L., Dai, H., Han, F. M., Lu, B. S., Wang, M. S., et al. (2019). Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. *Genome Biol.* 20:79. doi: 10.1186/s13059-019-1686-3

Zhang, C., Yao, X., Ren, H., Chang, J., and Wang, K. (2019a). RNA-Seq reveals flavonoid biosynthesis-related genes in pecan (*Carya illinoinensis*) Kernels. *J. Agric. Food Chem.* 67, 148–158. doi: 10.1021/acs.jafc.8b05239

Zhang, C., Yao, X., Ren, H., Wang, K., and Chang, J. (2019b). Isolation and characterization of three Chalcone synthase genes in pecan (*Carya illinoinensis*). *Biomolecules* 9:236. doi: 10.3390/biom9060236

Zhang, J., Zhang, W., Ji, F., Qiu, J., Song, X., Bu, D., et al. (2020). A high-quality walnut genome assembly reveals extensive gene expression divergences after whole-genome duplication. *Plant Biotechnol. J.* 18, 1848–1850. doi: 10.1111/pbi.13350

Zhang, Y., Yang, Z., Liu, G., Wu, Y., and Ouyang, J. (2020). Inhibitory effect of chestnut (*Castanea mollissima* Blume) inner skin extract on the activity of α-amylase, α-glucosidase, dipeptidyl peptidase IV and in vitro digestibility of starches. *Food Chem.* 324:126847. doi: 10.1016/j.foodchem.2020.126847

Zhang, Y., Butelli, E., Alseekh, S., Tohge, T., Rallapalli, G., Luo, J., et al. (2015). Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. *Nat. Commun.* 6:8635. doi: 10.1038/ncomms9635

Zhu, T., Wang, L., You, F. M., Rodriguez, J. C., Deal, K. R., Chen, L., et al. (2019). Sequencing a *Juglans regia*×*J. microcarpa* hybrid yields high-quality genome assemblies of parental species. *Hortic. Res.* 6:8635. doi: 10.1038/s41438-019-0139-1

Zhuang, W., Chen, H., Yang, M., Wang, J., Pandey, M. K., Zhang, C., et al. (2019). The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nat. Genet.* 51, 865–876. doi: 10.1038/s41588-019-0402-2

Zuo, J., Tang, W., and Xu, Y. (2015). "Chapter 68 - Anti-hepatitis B virus activity of chlorogenic acid and its related compounds," in *Coffee in Health and Disease Prevention*, ed. V. Preedy (San Diego: Academic Press), 607–613. doi: 10.1016/B978-0-12-409517-5.00068-1

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.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 20 April 2021
doi: 10.3389/fphar.2021.643857

Xanthohumol ameliorates Diet-Induced Liver Dysfunction via Farnesoid X Receptor-Dependent and Independent Signaling

Ines L. Paraiso[1,2], Thai Q. Tran[2], Armando Alcazar Magana[1,2,3], Payel Kundu[4], Jaewoo Choi[1], Claudia S. Maier[3], Gerd Bobe[1,5], Jacob Raber[2,4,6], Chrissa Kioussi[2] and Jan F. Stevens[1,2]**

[1]Linus Pauling Institute, Oregon State University, Corvallis, OR, United States, [2]Department of Pharmaceutical Sciences, Oregon State University, Corvallis, OR, United States, [3]Department of Chemistry, Oregon State University, Corvallis, OR, United States, [4]Department of Behavioral Neuroscience, Oregon Health and Science University, Portland, OR, United States, [5]Department of Animal and Rangeland Sciences, Oregon State University, Corvallis, OR, United States, [6]Department of Neurology, Psychiatry and Radiation Medicine, Division of Neuroscience, Oregon National Primate Research Center, Oregon Health and Science University, Portland, OR, United States

OPEN ACCESS

Edited by:
M. Carmen González-Mas,
University of Valencia, Spain

Reviewed by:
Saskia Van Mil,
University Medical Center Utrecht, Netherlands
Christina Lamers,
University of Basel, Switzerland

***Correspondence:**
Chrissa Kioussi
Chrissa.kioussi@oregonstate.edu
Jan F. Stevens
fred.stevens@oregonstate.edu

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: 18 December 2020
Accepted: 22 March 2021
Published: 20 April 2021

Citation:
Paraiso IL, Tran TQ, Magana AA, Kundu P, Choi J, Maier CS, Bobe G, Raber J, Kioussi C and Stevens JF (2021) Xanthohumol ameliorates Diet-Induced Liver Dysfunction via Farnesoid X Receptor-Dependent and Independent Signaling. Front. Pharmacol. 12:643857. doi: 10.3389/fphar.2021.643857

The farnesoid X receptor (FXR) plays a critical role in the regulation of lipid and bile acid (BA) homeostasis. Hepatic FXR loss results in lipid and BA accumulation, and progression from hepatic steatosis to nonalcoholic steatohepatitis (NASH). This study aimed to evaluate the effects of xanthohumol (XN), a hop-derived compound mitigating metabolic syndrome, on liver damage induced by diet and FXR deficiency in mice. Wild-type (WT) and liver-specific FXR-null mice ($FXR^{Liver-/-}$) were fed a high-fat diet (HFD) containing XN or the vehicle formation followed by histological characterization, lipid, BA and gene profiling. HFD supplemented with XN resulted in amelioration of hepatic steatosis and decreased BA concentrations in $FXR^{Liver-/-}$ mice, the effect being stronger in male mice. XN induced the constitutive androstane receptor (CAR), pregnane X receptor (PXR) and glucocorticoid receptor (GR) gene expression in the liver of $FXR^{Liver-/-}$ mice. These findings suggest that activation of BA detoxification pathways represents the predominant mechanism for controlling hydrophobic BA concentrations in $FXR^{Liver-/-}$ mice. Collectively, these data indicated sex-dependent relationship between FXR, lipids and BAs, and suggest that XN ameliorates HFD-induced liver dysfunction via FXR-dependent and independent signaling.

Keywords: nonalcoholic fatty liver disease, farnesoid X receptor, bile acids, lipid metabolism, xanthohumol

INTRODUCTION

Dyslipidemia coincides with other metabolic disorders such as obesity, hypertension, and glucose intolerance, defined as metabolic syndrome (MetS), which increase the risk to develop type 2 diabetes (T2D) and cardiovascular diseases (Porez et al., 2012). Obesity and T2D are also associated with nonalcoholic fatty liver disease (NAFLD), a spectrum of chronic liver abnormalities from simple steatosis to nonalcoholic steatohepatitis (NASH) to liver cirrhosis (Larter et al., 2010; Chiang, 2013). The growing prevalence of obesity and high-fat diet (HFD)-induced dyslipidemia represent a public health problem worldwide and the development of drugs with a combined effect on different risk factors may be more effective than the use of combinatorial therapy to manage patients' global risks.

Before the discovery of statins, hypercholesterolemia was primarily treated with bile acid (BA) sequestrants, which bind BAs in the intestine and prevent their reabsorption, thereby promoting the hepatic synthesis of BAs from cholesterol (Staels and Kuipers, 2007; Porez et al., 2012). BA synthesis in hepatocytes occurs largely through the classical pathway initiated by the rate-limiting enzyme cholesterol 7α-hydroxylase (CYP7A1). The classical pathway forms the primary BAs, cholic acid (CA) and chenodeoxycholic acid (CDCA), following a multistep enzymatic process. In complement, CYP27A1 initiates an alternative pathway of BA synthesis that also leads to CDCA synthesis (Garcia et al., 2018). Shortly after their synthesis, BAs are conjugated to glycine or taurine and stored into the gallbladder (Chiang, 2013; Garcia et al., 2018). Besides their involvement in transcriptional regulation of cholesterol metabolism (Chiang, 2002; Trauner et al., 2010), BAs regulate hepatic gluconeogenesis, glycogen synthesis and insulin sensitivity (Ma et al., 2006; Trauner et al., 2010). BAs also modulate neurotransmission, neuroendocrine responses, and neurogenesis indicating their importance in neurological functions (Schubring et al., 2012; McMillin and DeMorrow, 2016). However, BA accumulation causes inflammation, hepatic injury (Chiang, 2017) and is associated with motor and cognitive impairments (Huang et al., 2016; McMillin et al., 2016). A key regulator of maintaining lipid and BA homeostasis is the farnesoid X receptor (FXR, NR1H4), which upon activation by BAs, polyunsaturated fatty acids and farnesylated proteins (Forman et al., 1995; Makishima et al., 1999; Zhao et al., 2004), regulates the expression of target genes involved in various physiological processes (Chawla et al., 2001; Sun et al., 2021). An increase of intracellular BAs also activates the constitutive androstane receptor (CAR) and pregnane X receptor (PXR). They modulate transcriptional regulation of their targets including genes encoding hepatic BA metabolizing enzymes and BA/organic anion transporters (Guo et al., 2003; Uppal et al., 2005; Lee et al., 2006). Subsequently, FXR, CAR and PXR have emerged as promising targets for the treatment of metabolic disorders associated with MetS (Gao and Xie, 2012; Porez et al., 2012).

Xanthohumol (XN) is a hop-derived flavonoid, which mitigates obesity-related metabolic impairments by improving dysfunctional glucose and lipid metabolism in HFD-fed animals (Miranda et al., 2016; Miranda et al., 2018). Treatment of HFD-fed C57BL/6J mice with a diet containing XN decreases their plasma low-density lipoprotein cholesterol (LDL-c), IL-6, Homeostatic Model Assessment of Insulin Resistance (HOMA-IR) and leptin concentrations (Miranda et al., 2016). XN enhances fatty acid oxidation as a result of mild mitochondrial uncoupling (Kirkwood et al., 2013) and decreases adipocyte markers such as PPARγ, C/EBPα and DGAT1 (Yang et al., 2007). This effect might be at least partly mediated by FXR, since XN is a ligand of FXR (Yang et al., 2016) that modulates FXR downstream gene expression in a manner similar to selective bile acid receptor modulators (SBARM) (Nozawa, 2005; Paraiso et al., 2020). However, the extent to which FXR signaling mediates the *in vivo* effects of XN is unknown. Both activation of hepatic FXR and inhibition of intestinal FXR have beneficial effects in obesity-related metabolic diseases (Sun et al., 2021) due to differential effects on metabolic regulation (Kim et al., 2007; Schmitt et al., 2015). These effects are further emphasized by the observation that intestine-specific FXR knockout mice are resistant to HFD-induced obesity, while HFD-fed liver-specific FXR knockout mice develop NAFLD (Li et al., 2013; Schmitt et al., 2015). Therefore, tissue-specific mouse models are necessary to dissect the complex effects of FXR on dyslipidemia. In the current study, we used liver-specific FXR-null mice ($FXR^{Liver-/-}$) to investigate the effect of XN on dyslipidemia and BA accumulation. Our findings demonstrate that XN ameliorate HFD-induced hepatic injury and dysfunctional lipid and BA metabolism in WT and $FXR^{Liver-/-}$ mice. We also provide evidence that XN induces expression of nuclear receptors (NRs) including CAR, PXR and the glucocorticoid receptor (GR) involved in the metabolism of BAs and lipids. These findings have potentially important implications in the treatment of metabolic and cholestatic diseases.

MATERIALS AND METHODS

Animal Studies

All animal experiments were performed in accordance with institutional and National Health and Medical Research Council guidelines. The experimental protocol was approved by the Institutional Animal Care and Use Committee at Oregon State University and the studies were carried out in accordance with the approved protocol (IACUC 2019-0001). Nine-week-old WT male and female C57BL/6J mice were obtained from Jackson Laboratory (Bar Harbor, ME, United States). $FXR^{Liver-/-}$ mice were generated by crossing $FXR^{FL/FL}$ mice with mice harboring the Cre recombinase under the control of the albumin promoter (Alb^{Cre}) to produce the Alb^{Cre}:$FXR^{FL/FL}$ or $FXR^{Liver-/-}$ mice (Kong et al., 2016). All mice were in C57BL/6J genetic background for over 12 generations. Mice were housed in groups of two–3 in ventilated cages under a 12–12-h light-dark cycle and fed a HFD (Dyets Inc. Bethlehem, PA, United States) containing 60, 20 and 20% total calories from fat, carbohydrate and protein, respectively. XN (purity >99%) from Hopsteiner Inc (New York, NY, United States) was mixed into the diet as previously described (Miranda et al., 2018) to deliver a dose of 60 mg/kg body weight/day. The control diet contained an identical amount of the vehicle. 15 WT mice (8 females, 7 males) and 18 $FXR^{Liver-/-}$ mice (10 females, 8 males) were fed a control HFD, while 15 WT mice and 18 $FXR^{Liver-/-}$ mice were treated with XN for a duration of 12 weeks. Food intake and body weights were recorded weekly. At week 10, fasting glucose was measured after 6 h of fasting by using the One Touch UltraMini glucometer (LifeScan Inc. Milpitas, CA, United States). At the end of 12 weeks of feeding, fed-state mice were euthanized by cervical dislocation, their blood collected, and their liver and hippocampus were dissected for further analyses. Deletion of FXR in the liver of $FXR^{Liver-/-}$ mice was confirmed by genotyping at weaning (Kong et al., 2016). Quantitative PCR after the feeding experiment. FXR

mRNA levels were ~ 3-fold lower in the liver of mutant compared to WT mice (**Supplementary Figure S1**).

Histology

Liver biopsies from $n = 3$ male mice per genotype-diet group were fixed in 4% paraformaldehyde, embedded in OCT and 10 µm-thick sections were used for histology. Hematoxylin and Eosin (H&E) and Sudan black staining were performed as previously described (Chang et al., 2019).

Measurement of Hepatic Transaminase Activities and Plasma Leptin Concentrations

To measure ALT and AST enzymatic activities, liver samples ($n = 6$ per genotype-diet group) were homogenized in 10 ml of 100 mM Tris (pH = 7.8) per Gram of tissue. The homogenates were centrifuged at 10,000 ×g for 15 min at 4°C. The supernatants were analyzed for ALT and AST activity using colorimetric assay kits purchased from Cayman Chemical (Ann Arbor, MI, United States). Plasma leptin concentrations (n = 5–7 per genotype-diet group) were measured using the Enzyme Immunoassay kit from SPI Bio Inc.. (Sherbrooke, QC, Canada) as per manufacturer's instructions.

Liver Lipidomics

Mouse liver samples (50 mg, n = 15–18 per genotype-diet group) were spiked with SPLASH® Lipidomix® internal standards from Avanti Lipids (Alabaster, AL, United States) and homogenized with zirconium oxide beads and 1 ml of cold methylene chloride: isopropanol: methanol (25:10:65, v/v/v) + 0.1% butylated hydroxytoluene (BHT). The mixture was incubated at –20°C for 1 h and centrifuged at 13,000 rpm for 10 min 20 µL of the supernatant was diluted 1/10 in extraction solvent before MS analysis. UPLC was performed using a 1.7 µm particle, 2.1 mm × 100, CSH C18 Column (Waters, Milford, MA, United States) coupled to a quadrupole TOF mass spectrometer (AB SCIEX, TripleTOF 5600) operated in information-dependent MS/MS acquisition mode. LC and MS conditions were developed by our group and described previously by Choi et al. (Choi et al., 2015) with some adjustments. For positive ion mode LC-QToF-MS/MS, the mobile phases consisted of (A) 60:40 (v/v) acetonitrile: water with ammonium formate (10 mM) and formic acid (0.1%) and (B) 90:10 (v/v) isopropanol: acetonitrile with ammonium formate (10 mM) and formic acid (0.1% formic acid). For analyses run in the negative ion mode, ammonium acetate (10 mM) was used as the modifier. Quantification of lipid species was performed using MultiQuant Software version 3.0.2 (SCIEX), after annotation in PeakView Software Version 1.2 (SCIEX) based on accurate masses and retention times for each lipid. The library of lipid profiling for identification was introduced by Cajka et al. (Cajka et al., 2017).

Bile Acid Analysis

Plasma samples collected post-euthanasia (20 µL, $n = 15–18$ per genotype-diet group) were spiked with 0.1 ng of cholic acid-d_4 internal standard (Cayman Chemical, Ann Harbor, MI, United States) per µL of plasma. 1 ml of ice-cold acetonitrile was added, and the mixture was vortexed and centrifuged at 13,000 rpm for 10 min. The supernatant was evaporated under vacuum and reconstituted in 50% MeOH.

Liver samples without gallbladder (25 mg, $n = 15–18$ per genotype-diet group) were homogenized in 1 ml of solvent (isopropanol/water, 2:1, v/v with 0.1% formic acid) containing 1.6 ng/ml of cholic acid-d_4 internal standard. Samples were homogenized using a counter-top bullet blender for 5 min and centrifuged at 13,000 rpm for 5 min. The supernatants were filtered with OSTRO phospholipid removal plate (Waters, Milford, MA, United States), evaporated under vacuum and reconstituted in 50% MeOH.

Left and right hippocampus were pooled, ground in liquid nitrogen and freeze-dried. The samples were weighed and spiked with 1 pg of cholic acid-d_4 internal standard per mg of hippocampus. Approximately 8 mg of hippocampus (dry weight, $n = 13–17$ per genotype-diet group) were homogenized with 1:30 µL (m/v) of 50% MeOH using a counter-top bullet blender for 10 min and centrifuged at 15,000 rpm for 20 min and supernatants used for HPLC analysis.

UPLC was performed using a 1.7 µm particle, 2.1 mm × 100, CSH C18 column (Waters, Milford, MA, United States) coupled to a hybrid triple quadrupole linear ion trap mass spectrometer (AB SCIEX, 4000 QTRAP). LC and MS conditions were developed by our group and described in the Supplemental data. BAs were identified by matching their retention time, isotopic pattern, exact mass of the $[M-H]^-$ ion and fragmentation pattern with those of authentic standards (IROA Technologies, Sea Girt, NJ, United States). SRM transitions used for quantification are listed in (**Supplementary Table S1**) and additional parameters such as collision energy are listed in (**Supplementary Table S2**).

100% of the mice had hepatic and plasma BA above the detection limit and 82% (52 out of 63 mice) had hippocampal BA above the detection limit, i.e. 74% of the WT mice vs. 89% of the $FXR^{Liver-/-}$ mice.

XN and Metabolites Concentrations in Liver and Plasma

Liver and plasma extracts ($n = 15–17$ per genotype-diet group) were analyzed for XN and metabolites by LC-MS/MS using a hybrid triple quadrupole linear ion trap mass spectrometer (AB SCIEX, 4000 QTRAP). Analytes were separated by UPLC carried out using a 2.1 × 50 mm Agilent Zorbax 300 SB-C8 3.5 µm column (Agilent, Santa Clara, CA, United States). Each run lasted 6 min at a flow rate of 0.4 ml/min. The elution gradient started at 30% solvent B (0.1% formic acid in acetonitrile) in solvent A (0.1% formic acid in water) and was increased to 60% solvent over the initial 1.5 min. The gradient was held at 60% for 1 min, increased to 100% B for 0.5 min, held at 100% B from 3.0 to 3.8 min, then dropped to 30% B in 0.1 min. The column was equilibrated for 2.1 min until 6.0 min. SRM transitions for quantification were m/z 353 → 119 for XN and isoxanthohumol (IX), m/z 339 → 219 for 8-prenylnaringenin

(8 PN), m/z 355 → 249 for α,β-dihydroxanthohumol (DXN), and m/z 341 → 235 for O-desmethyl-α,β-dihydroxanthohumol (DDXN).

RNA Sequencing and Analysis

RNA was prepared from liver samples (n = 4-5 per genotype-diet group) and sequenced as previously described (Singh et al., 2018). All samples were processed and analyzed in parallel. Sequence quality was assessed by FastQC. Reads were aligned by Hisat2 (Kim et al., 2019) and Samtools (Li et al., 2009). A gene count matrix was generated byStringtie (Pertea et al., 2015). Two data sets, HFD (control) and HFD-XN (treatment) were derived from the gene count matrix. Each set was analyzed in parallel by DESeq2 (Love et al., 2014) for differential expression (DE) calculation. DE was calculated for $FXR^{Liver-/-}$ mutant over wildtype. PCA plots were generated using the DESeq2 package. Benjamini-Hochberg multiple-test correction was applied to control for the number of false positives with an adjusted 5% statistical significance threshold (Benjamini and Hochberg, 1995). A heatmap was created using the pheatmap package in R (version 3.6). Functional annotation clustering was achieved in Network Analyst v3.0 using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

Real-Time PCR

RNA samples from mouse liver (n = 4-5 per genotype-diet group) were reverse-transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Waltham, MA, United States). universal SYBR® Green Supermix (Bio-Rad, Hercules, CA, United States) was used following the manufacturer's protocol and amplifications were performed using the ABI Prism 7300 Real-Time PCR System (Applied Biosystems, Waltham, MA, United States). Each sample had two technical replicates. Gene expression was normalized to levels of Polymerase-II. Relative gene expression was calculated using the 2^{-ddCt} method. All primers were purchased from IDT technologies (Coralville, IA, United States) and are listed in (**Supplementary Table S3**).

Statistical Analysis

Statistical data were analyzed in SAS version 9.4 (SAS Ins. Inc., Cary, NC). Plasma leptin, AST, and ALT concentrations, intestinal gene expression, and liver receptor data were not normally distributed and could not be normalized through transformation. Therefore, these parameters were analyzed using the non-parametric Wilcoxon rank sum test after checking for interactions. We categorized values into elevated and normal and used Fisher's exact test to compare treatment groups. BAs and XN concentrations were not normally distributed but rather distributed logarithmically to the base 10, where 1 is equal to 10, two is equal to 100, and 3 is equal to 1000, and were analyzed on that scale. In addition, we categorized BA values into elevated and normal and compared treatment groups using Fisher's exact test. The remaining lipid data were analyzed without transformation. The effect of XN-treatment was evaluated separately for WT and $FXR^{Liver-/-}$ mice using a generalized linear model in PROC GLM with XN-treatment, sex, and their interaction, because $FXR^{Liver-/-}$ mice had larger variance estimates than WT mice. The effects of genotype and sex were evaluated in untreated mice using a generalized linear model in PROC GLM with genotype, sex, and their interactions, as XN modified the effect of genotype and sex. All statistical tests were two-sided. Significance was declared at $p \leq 0.05$. Correlations were tested by calculating non-parametric Spearman's correlation coefficient, r.

RESULTS

Sex Influences XN Metabolism in WT and $FXR^{Liver-/-}$ Mice

During the course of the study, weight gain in HFD-fed WT and $FXR^{Liver-/-}$ mice were comparable (**Table 1**). HFD-fed males gained more body weight than females (**Supplementary Figure S2**); this effect was significant in WT mice ($p < 0.0001$) but not in $FXR^{Liver-/-}$ mice ($p = 0.1$). To ensure oral bioavailability of XN in WT and $FXR^{Liver-/-}$ mice, we measured liver and plasma concentrations of XN and its metabolites in XN-treated mice. Oral bioavailability of XN was comparable in both genotypes (**Supplementary Table S4**), but IX, a product of XN isomerization, reached higher concentrations in the liver of WT mice compared to $FXR^{Liver-/-}$ mice. Moreover, we observed sex-related differences as females had significantly higher concentrations of XN and IX than males (**Table 2**). Since there was no difference in food intake, this observation is likely a result of the lower body weight in females compared to males. α,β-Dihydroxanthohumol (DXN), a bacterial metabolite of XN (Paraiso et al., 2019) was not affected by sex or genotype, while 8-prenylnaringenin (8PN) hepatic concentrations were elevated in male $FXR^{Liver-/-}$ mice. These observations suggest an influence of sex on XN pharmacokinetics, likely due to the differences in weight and volume of distribution between males and females.

XN Ameliorates HFD-Induced Liver Damage

To assess if the HFD successfully induced NAFLD, we examined liver sections from three representative male mice per treatment group. Hematoxylin and eosin (H&E) stained liver sections showed hepatic steatosis in the form of vacuoles with a clear appearance in HFD-fed WT and $FXR^{Liver-/-}$ mice (**Figures 1A,B**). We observed a reduction in number and size of these vacuoles in both genotypes in XN-treated mice (**Figures 1C,D**). Sections stained with Sudan black confirmed an increase in lipid vacuoles in $FXR^{Liver-/-}$ mice compared to WT (**Figures 1E,F**), which was reversed in XN-treated mice (**Figures 1G,H**, **Supplementary Figure S3**). Another marker of NAFLD is the proportion of liver weight (LW) over total body weight (LW%). After 12 weeks on the HFD, untreated $FXR^{Liver-/-}$ mice exhibited increased LW (p = 0.02) and LW% (p = 0.03) than WT mice (**Table 1**). In addition, males had increased LW than females in WT mice (p = 0.001) and $FXR^{Liver-/-}$ mice ($p < 0.0001$, **Supplementary Figure S4A**). Fasting glucose was also elevated in males compared to females in both genotypes (**Supplementary Figure S4B**).

To measure the extent to which the steatosis had resulted in liver tissue damage, we measured aspartate aminotransferase (AST) and alanine aminotransferase (ALT) enzymatic activities

TABLE 1 | A list of metabolic parameters measured in WT and $FXR^{Liver-/-}$ mice upon 10 weeks ([a]) or 12°weeks of HFD ± XN.

	WT	WT XN	$FXR^{Liver-/-}$	$FXR^{Liver-/-}$ XN
Initial body weight (g)	22.28 ± 0.9	22.18 ± 0.71	22.57 ± 0.73	22.32 ± 0.52
Body weight gain[a] (g)	12.22 ± 1.62	11.71 ± 1.73	16.07 ± 1.63	15.72 ± 1.02
Body weight gain (g)	15.25 ± 1.38	14.62 ± 1.45	16.22 ± 1.54	16.24 ± 1.18
Fasting glucose[a] (mg/dl)	200.73 ± 13.55	198 ± 9.59	200.5 ± 10.82	203.81 ± 6.2
Liver weight (g)	1.11 ± 0.08	1.09 ± 0.12	$1.35^{\#}$ ± 0.09	1.23 ± 0.07
% Liver weight (% body weight)	2.95 ± 0.11	2.84 ± 0.17	$3.50^{\#}$ ± 0.18	3.41 ± 0.19
AST (U/mL)	0.32 ± 0.09	0.28 ± 0.08	$1.46^{\#}$ ± 0.6	0.24^{*} ±0.08
ALT (U/mL)	0.53 ± 0.14	0.38 ± 0.1	0.83 ± 0.39	0.33 ± 0.08
Food intake (g/day)	3.05 ± 0.34	3.40 ± 0.17	2.51 ± 0.05	2.79 ± 0.06
Leptin (ng/ml)	30.14 ± 6.99	29.75 ± 4.03	66.2 ± 15.36	43.38 ± 10.52

Data displayed as mean ± SEM. Significant differences are marked as $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ *for effect of XN treatment,* $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ *for genotype comparison,* $^{\&}p < 0.05$, $^{\&\&}p < 0.01$, $^{\&\&\&}p < 0.001$ *for gender comparison.*

TABLE 2 | Concentrations of XN and metabolites (IX, 8PN, DXN) in the plasma and liver of females vs. males HFD-fed WT and $FXR^{Liver-/-}$ mice.

			Plasma (nM)	
	Female WT XN	**Male WT XN**	**Female $FXR^{Liver-/-}$ XN**	**Male $FXR^{Liver-/-}$ XN**
XN	22.61 ± 4.4	16.94 ± 3.2	30.45 ± 5.3	$14.65^{\&}$ ± 1.8
IX	16.0 ± 4.0	$8.73^{\&}$ ± 2.0	13.55 ± 3.2	8.78 ± 1.6
DXN	1.90 ± 0.4	3.02 ± 1.2	1.36 ± 0.45	3.60 ± 1.89
			Liver (nmol/g)	
	Female WT XN	**Male WT XN**	**Female $FXR^{Liver-/-}$ XN**	**Male $FXR^{Liver-/-}$ XN**
XN	0.28 ± 0.05	$0.16^{\&}$ ± 0.03	0.23 ± 0.08	$0.29^{\&}$ ± 0.05
IX	1.53 ± 0.35	$0.68^{\&}$ ± 0.08	$0.58^{\#\#}$ ± 0.09	0.65 ± 0.05
8PN	0.04 ± 0.01	0.05 ± 0.01	0.06 ± 0.03	$0.13^{\&,\&\&}$ ± 0.04

Data displayed as mean ± SEM (n = 7–10 per group). $^{\#}p < 0.05$, $^{\#\#}p < 0.01$ *for genotype comparison,* $^{\&}p < 0.05$, $^{\&\&}p < 0.01$ *for gender comparison.*

in liver homogenates. Absence of hepatic FXR might promote liver tissue damage as AST levels were increased in untreated $FXR^{Liver-/-}$ compared to WT mice (p = 0.02, **Table 1**). XN reduced AST levels in treated $FXR^{Liver-/-}$ mice (p = 0.03). Plasma leptin concentrations were elevated in $FXR^{Liver-/-}$ mice but differences in leptin and food intake among groups were not significant. These results suggest that HFD-induced NAFLD is accentuated in absence of hepatic FXR and the severity of the hepatic steatosis is attenuated by XN supplementation.

XN Ameliorates HFD-Induced Lipid Accumulation

We annotated and measured the relative abundances of 116 individual hepatic lipids including triglycerides (TG), free cholesterol, esterified cholesterol (CE), ceramides and sphingomyelins (SM; **Figure 2A**). We observed sex, genotype and XN-dependent effects on lipid composition (**Figures 2B–H**).

Hepatic TG were increased in female $FXR^{Liver-/-}$ compared to their WT counterparts, but male $FXR^{Liver-/-}$ mice displayed the most severe hepatic steatosis. Male $FXR^{Liver-/-}$ mice had elevated hepatic TG, cholesterol, CE and ceramides compared to female $FXR^{Liver-/-}$ mice and compared to male WT mice (**Figures 2B–E**). The proportion of CE over total cholesterol (%CE) and the proportion of ceramide over total sphingolipids (%ceramide) was higher in $FXR^{Liver-/-}$ males compared to WT males (**Figures 2G,H**). These data indicate that male mice are more responsive to diet-induced hepatic steatosis in the absence of FXR signaling in the liver.

In $FXR^{Liver-/-}$ mice, XN had a predominant effect in male mice, which exhibited the highest hepatic lipid accumulation. Total CE (p = 0.06), cholesterol (p = 0.07) and ceramides (p = 0.0005) were decreased in XN-treated male $FXR^{Liver-/-}$ mice compared to the untreated mice. The %ceramide, a measure of the proportion of ceramides among other sphingolipids, correlated better with histological improvements than ceramide abundances. XN treatment decreased %ceramide in male WT (p = 0.01) and male $FXR^{Liver-/-}$ mice (p = 0.02); both groups had the most elevated %ceramide among the untreated groups. SM abundances followed a trend opposite to that of other lipids and were increased in XN-treated WT males (**Figure 2F**). These data suggest that XN regulates lipid metabolism via pathways independent of hepatic FXR signaling.

XN Ameliorates HFD-Induced Dysfunctional BA Metabolism

We screened for 34 individual BAs in the plasma, liver and hippocampus of WT and $FXR^{Liver-/-}$ mice using UPLC-MS/MS.

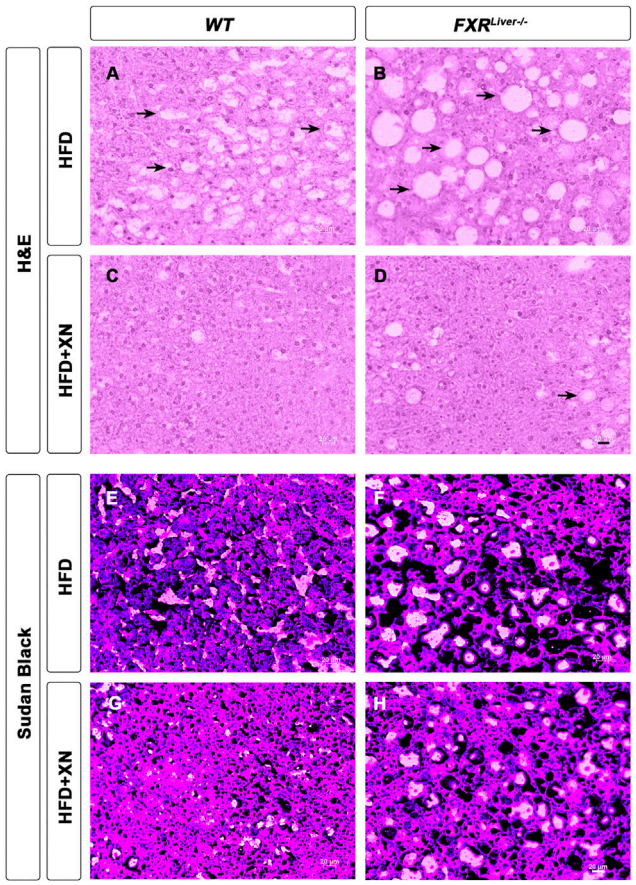

FIGURE 1 | XN prevents HFD-induced hepatic steatosis. Representative liver histology by H&E **(A-D)** and Sudan Black **(E-H)** of male WT and $FXR^{Liver-/-}$ mice fed HFD ± XN. Arrows indicate vacuoles, a characteristic structure of hepatic steatosis.

Hippocampal BAs were measured to investigate BA retention in tissues deficient in BA detoxification and export mechanisms. Fifteen BAs were detected and quantified in the liver, 12 BAs in the plasma and 7 BAs in the hippocampus.

$FXR^{Liver-/-}$ had higher BA concentrations in plasma ($p < 0.0001$, **Figure 3A**) and liver ($p = 0.01$, **Figure 3C**) than WT mice, while an increase in hippocampal BAs was observed in $FXR^{Liver-/-}$ males only ($p = 0.04$, **Supplementary Figure S5**). The most severe BA accumulation occurred in the plasma of $FXR^{Liver-/-}$ mice, with a 7-fold increase in total BAs vs. 2-fold increase in the liver. The change was driven by an increase in primary conjugated BAs (**Figures 3A–D**). Hippocampal BA retention was more pronounced than hepatic BA retention as $FXR^{Liver-/-}$ mice exhibited 9-fold increase in total hippocampal BAs compared to WT (**Figures 3E,F**) indicating passage of BAs through a possibly altered blood brain barrier (BBB). These data suggest that hepatic mechanisms of BA efflux remained more efficient than cerebral mechanisms of BA efflux in $FXR^{Liver-/-}$ mice. This is further supported by our observation that, in $FXR^{Liver-/-}$ mice, plasma BAs were more strongly correlated to hippocampal BAs ($r = 0.90$, $p < 0.0001$) than to hepatic BAs ($r = 0.49$, $p = 0.005$, **Figures 3G,H**). BA pool composition was also modified in $FXR^{Liver-/-}$ mice. The percentage of plasma primary conjugated BAs over total BAs was 50% in WT mice vs. 85% in $FXR^{Liver-/-}$ mice (**Figure 3B**). Hepatic conjugation of BAs improves their hydrophilicity and reduces their toxicity suggesting that $FXR^{Liver-/-}$ mice developed a metabolic mechanism to counter BA-mediated toxicity.

XN treatment promoted BA synthesis in WT mice but attenuated BA accumulation in $FXR^{Liver-/-}$ mice. XN effect on BA concentrations was independent of sex in WT mice. XN supplementation of WT mice resulted in increased plasma primary conjugated BAs ($p = 0.03$, **Figure 3A**) and increased hepatic primary unconjugated BAs ($p = 0.02$, **Figure 3C**). These observations are in accordance with previous reports that XN induces CYP7A1 and hepatic BA synthesis in WT mice resulting in increased BA concentrations (Paraiso et al., 2020). DCA, TCA, β-MCA and FXR antagonists, T-α-MCA and T-β-MCA, were increased in the liver and/or plasma of XN-treated WT mice (**Figures 4A,B**). By contrast, XN treatment resulted in decreased BA concentrations in $FXR^{Liver-/-}$ mice. CA and CA-derived BAs including DCA, TCA and TDCA were decreased in the liver of $FXR^{Liver-/-}$ mice in both sexes, with males exhibiting more significant changes (**Figures 4A–C**). CA, DCA and ω-MCA were decreased in the plasma of XN-treated male $FXR^{Liver-/-}$ mice.

In summary, XN reduced most individual BAs of the classical pathway in $FXR^{Liver-/-}$ mice and increased BAs from the alternative pathway of synthesis in WT mice (**Figures 4A,B**; **Supplementary Figure S6**). Since BAs did not reach pathological concentrations in WT mice, these results suggest an adaptation of XN mechanism of action to the pathophysiological conditions and the possible activation of BA receptors independent of FXR. These observations further support a genotype-specific differential modulation of metabolism by XN.

XN Ameliorates HFD-Induced Hepatic Gene Profiles

We analyzed changes in global gene expression profiles in hepatic tissue of HFD-fed WT and $FXR^{Liver-/-}$ mice. Since a smaller subset of samples was sequenced, male and female mice RNA sequencing data were pooled to increase the power of the analysis. Functional annotation clustering revealed that genes differentially affected by HFD in WT and $FXR^{Liver-/-}$ mice can be classified into two main functional groups: genes involved in metabolic processes vs. genes involved in inflammation and carcinogenic processes (**Supplementary Table S5**). The comparison between XN-treated and untreated $FXR^{Liver-/-}$ mice revealed 243 shared genes, with 759 features unique to untreated mice and 170 features unique to XN-treated mice (**Figure 5A**). Within the shared features, XN supplementation impacted several gene networks including lipid metabolism (Mgat2, Sptlc2, Smpd3), ABC transporters involved in BA transport (Abcc4, Abcc3, Abcb11), metabolism of xenobiotics with genes involved in phase I and II metabolism (Gsta1, Gstm3, Ugt1a7c, Sult2a7), PI3K-Akt signaling pathway (Tnc, Tlr2, Spp1, Thbs1, Lamb3), cytokine-cytokine receptor interactions (Ccl2, Cxcl9/10, Cd9, Tnfrsf1a) and amino acid metabolism (Sardh, Aadat, Kyat3) (**Figure 5B**).

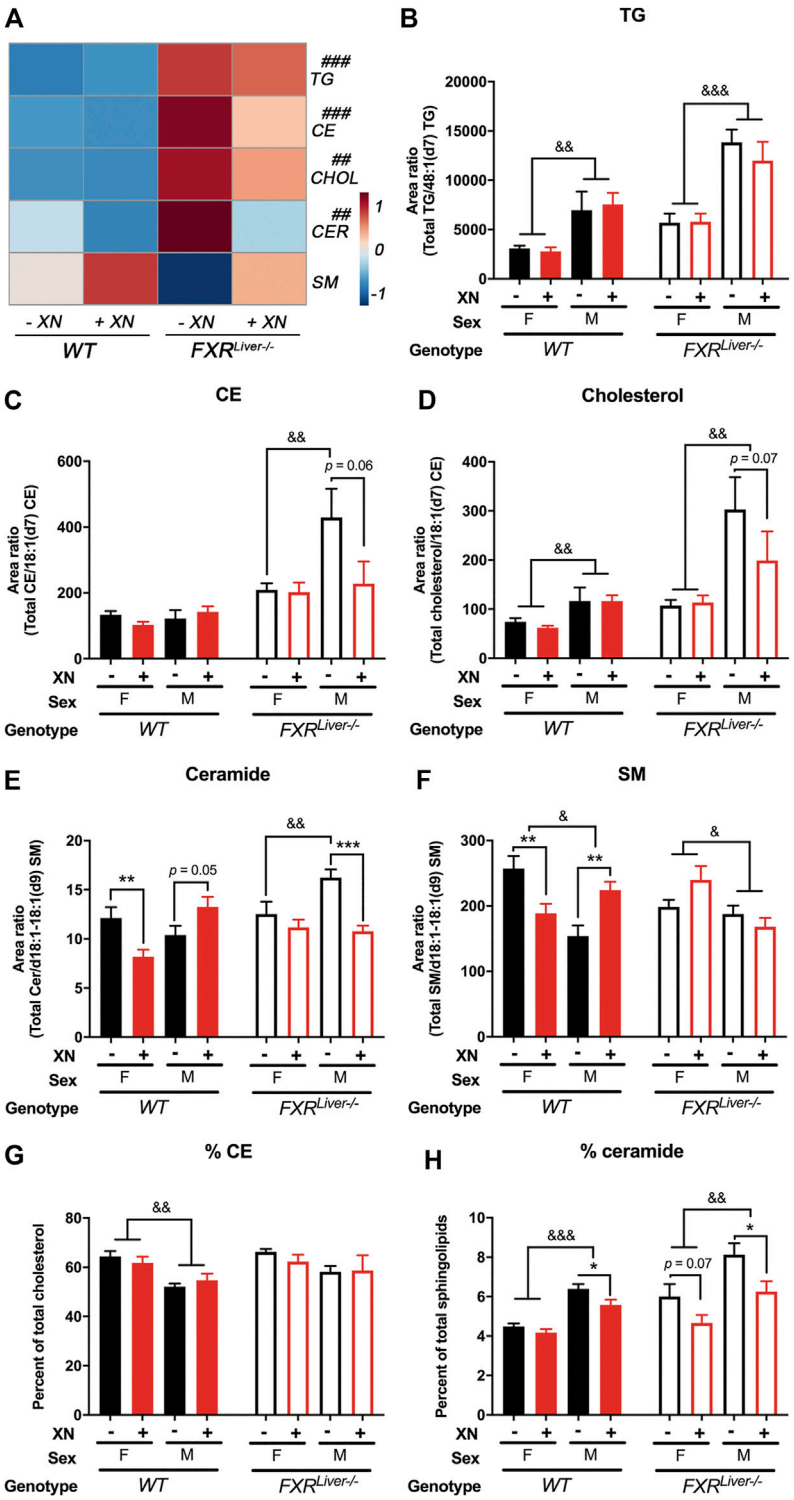

FIGURE 2 | XN prevents HFD-induced ceramide accumulation **(A)** Heatmap of hepatic triglycerides (TG), cholesterol esters (CE), free cholesterol (CHOL), ceramides (CER) and sphingomyelins (SM) in HFD-fed WT and $FXR^{Liver-/-}$ mice untreated or treated with XN. Total relative abundance of **(B)** TG **(C)** CE **(D)** free cholesterol **(E)** ceramide and **(F)** SM in the liver of WT and $FXR^{Liver-/-}$ mice untreated or treated with XN. Proportion of **(G)** CE and **(H)** ceramide in the liver of WT and $FXR^{Liver-/-}$ mice untreated or treated with XN. Values are mean ± SEM (n = 7–10 per group). $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ for effect of XN treatment; $\#p < 0.05$, $\#\#p < 0.01$, $\#\#\#p < 0.001$ for genotype comparison; $^{\&}p < 0.05$, $^{\&}p < 0.01$, $^{\&}p < 0.001$ for gender comparison.

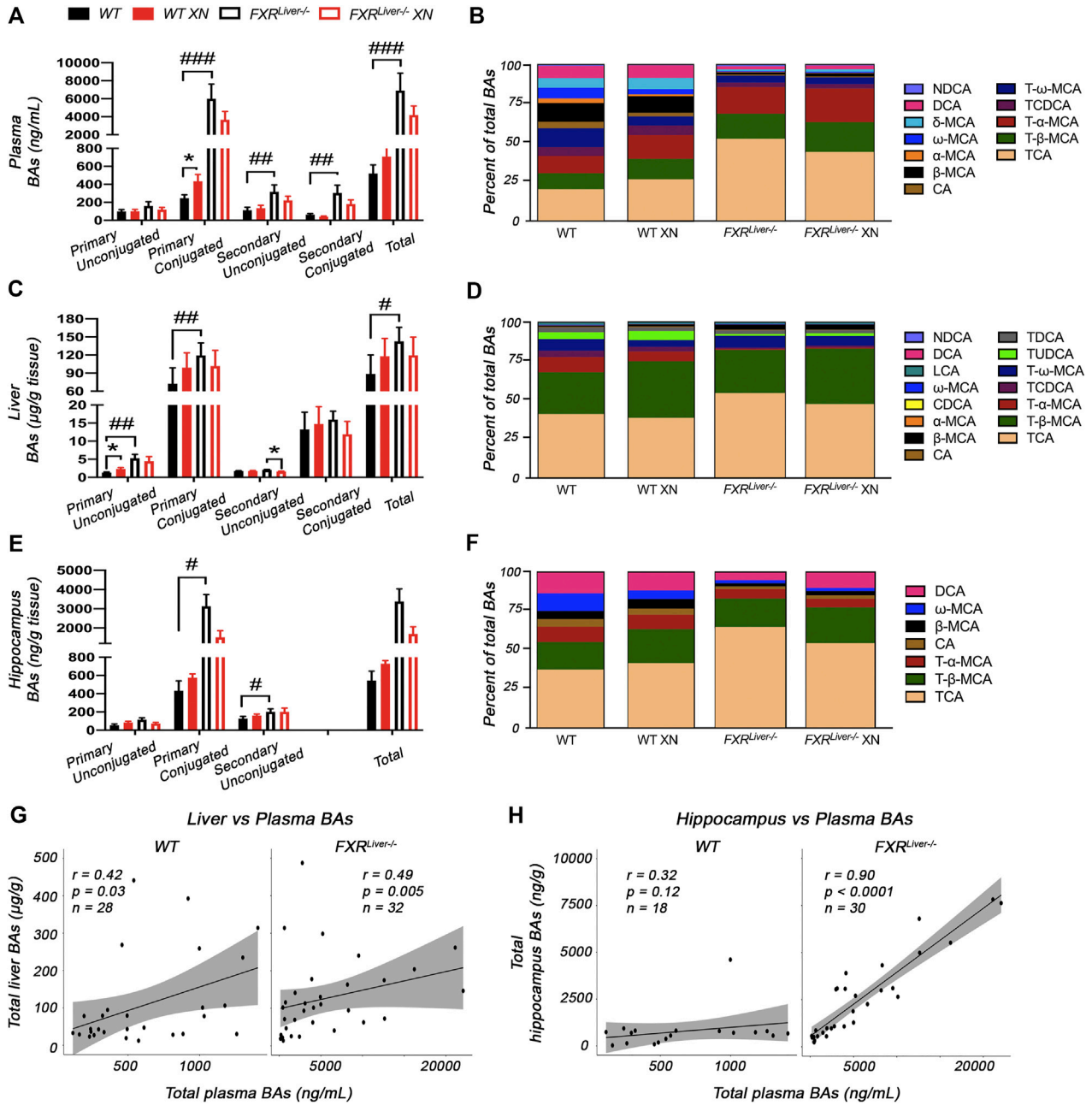

FIGURE 3 | XN modulates BA composition **(A)** Total BAs and **(B)** composition of the BA pool in the plasma of HFD-fed WT and FXR$^{Liver-/-}$ mice **(C)** Total BAs and **(D)** composition of BAs in the liver of HFD-fed WT and FXR$^{Liver-/-}$ mice **(E)** Total BAs and **(F)** composition of BAs in the hippocampus of HFD-fed WT and FXR$^{Liver-/-}$ mice **(G)** Correlations between liver and plasma BAs in WT and FXR$^{Liver-/-}$ mice **(H)** Correlations between hippocampus and plasma BAs in WT and FXR$^{Liver-/-}$ mice. Bar graphs values are mean ± SEM (n = 15–18 per group). *p < 0.05, **p < 0.01, ***p < 0.001 for effect of XN treatment; #p < 0.05, ##p < 0.01, ###p < 0.001 for genotype comparison. Abbreviations: chenodeoxycholic acid (CDCA), cholic acid (CA), deoxycholic acid (DCA), lithocholic acid (LCA), muricholic acid (MCA), nordeoxycholic acid (NDCA), taurochenodeoxycholic acid (TCDCA), taurocholic acid (TCA), tauromuricholic acid (T-MCA), taurodeoxycholic acid (TDCA), tauroursodeoxycholic acid (TUDCA).

In WT mice, presence of XN was associated with increased expression of genes involved in lipid and xenobiotic metabolism such as Mgat1, Cyp1a1, Ugt1a7c, and decreased expression of genes involved in energy metabolism (Clock, Rnf146) and inflammation (Cebpg, Saa1, Saa2) (**Figure 5C**; **Table 3**). CCAAT-enhancer binding proteins (C/EBP) interact with the proximal promoter of the Saa genes and regulate hepatic expression of SAA (Ray et al., 1995). The concurrent decrease in Cebpg and Saa expression suggests that XN-mediated repression of Saa1 and Saa2 is mediated by inhibition of Cebpg expression.

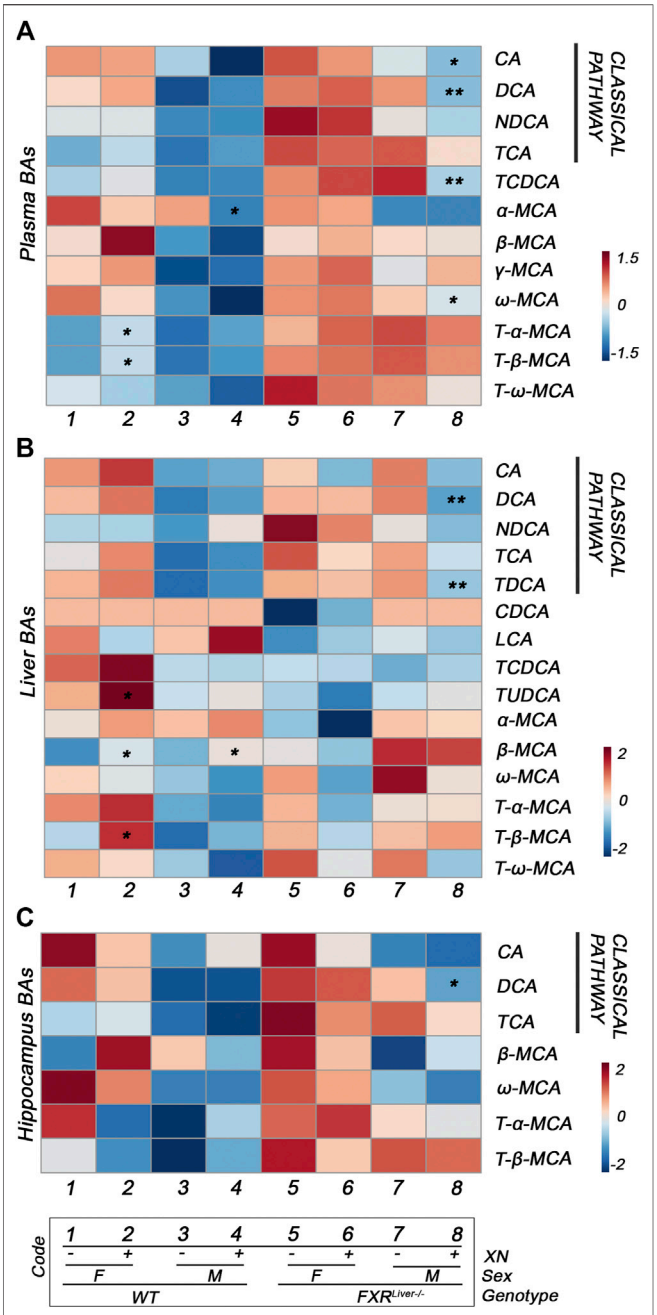

FIGURE 4 | XN differentially modulates classical and alternative pathways of synthesis in WT vs. $FXR^{Liver-/-}$ mice. Heatmaps of individual BA concentrations in the plasma **(A)**, liver **(B)** and hippocampus **(C)** of HFD-fed WT and $FXR^{Liver-/-}$ mice. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ for effect of XN treatment (n = 7–10 per group). Abbreviations: chenodeoxycholic acid (CDCA), cholic acid (CA), deoxycholic acid (DCA), lithocholic acid (LCA), muricholic acid (MCA), nordeoxycholic acid (NDCA), taurochenodeoxycholic acid (TCDCA), taurocholic acid (TCA), tauromuricholic acid (T-MCA), taurodeoxycholic acid (TDCA), tauroursodeoxycholic acid (TUDCA).

In $FXR^{Liver-/-}$ mice, XN induced Vsig4, which attenuates macrophage-mediated hepatic inflammation (Li et al., 2019), Acvr1, which is involved in activin signaling (Rao et al., 2017), and Timd4, that controls adaptive immunity by clearing antigen-specific T-cells (Albacker et al., 2010) (**Figure 5D**; **Table 3**). While XN induced increased expression of Cyp7a1, which is involved in the classical pathway of BA synthesis, there were no changes in the expression of genes involved in the alternative pathway of BA synthesis. RhoGDI2, encoded by Arhgdib, is involved in the molecular pathogenesis of liver fibrosis (Utsunomiya et al., 2007) and acts as a positive regulator of Rac1 (Kardol-Hoefnagel et al., 2020). The decreased expression of both genes indicated that XN inhibition of Rac1 transcription might be mediated by repression of Arhgdib in $FXR^{Liver-/-}$ mice (**Table 3**). These data show that XN promotes lipid and BA metabolism and decreases acute inflammation in WT mice, while XN attenuates inflammation by controlling immune response, inhibits cell proliferation and liver fibrosis in $FXR^{Liver-/-}$ mice.

XN Induces Gene Expression of FXR-Independent NRs

RNA sequencing revealed several genes involved in phase II reactions such as glucuronidation (UGTs), sulfation (SULTs) and glutathione conjugation (GSTs) were regulated by XN. Therefore, we performed a quantitative analysis of the hepatic expression of a panel of NRs known to regulate phase II BA metabolism including CAR, PXR and GR. BAs interact with CAR (Moore et al., 2002), PXR (Staudinger et al., 2001), as well as GR (Tanaka and Makino, 1992), which also regulates biosynthesis and transport of bile salts (Xiao et al., 2016). XN treatment resulted in higher CAR expression in all sex and genotype groups (**Figure 6A**), although XN effect was stronger in males. XN also induced gene expression of PXR in both WT and $FXR^{Liver-/-}$ mice, while an increase in GR transcript levels was observed in XN-treated $FXR^{Liver-/-}$ mice only (p = 0.009, **Figure 6A**).

Changes in these NR expression profiles were linked to hepatic BA concentrations. Correlation analyses revealed that relative expression levels of these receptors were negatively correlated with unconjugated BAs in the liver absence of FXR (**Figure 6B**). Additionally, relative expression of GR was negatively correlated with relative abundances of lipids regulated by XN, i.e. CE (r = -0.78, p = 0.01), cholesterol (r = -0.67, p = 0.04) and ceramide (r = -0.68, p = 0.03) in absence of FXR. Collectively, these data suggest that, in absence of hepatic FXR, induction of CAR, PXR and GR is involved in XN-mediated decrease of lipid and BA concentrations.

DISCUSSION

NRs regulate ligand-activated transcriptional activation of a myriad of genes for the elimination and detoxification of potentially toxic biliary constituents accumulating in cholestasis (Halilbasic et al., 2013). FXR controls the transcriptional activation of several genes involved in the regulation of glucose and lipid metabolism and maintenance of BA homeostasis, thereby protecting the host against liver damage associated with lipid and BA accumulation. BAs act as signaling molecules through BA receptors such as FXR, TGR5,

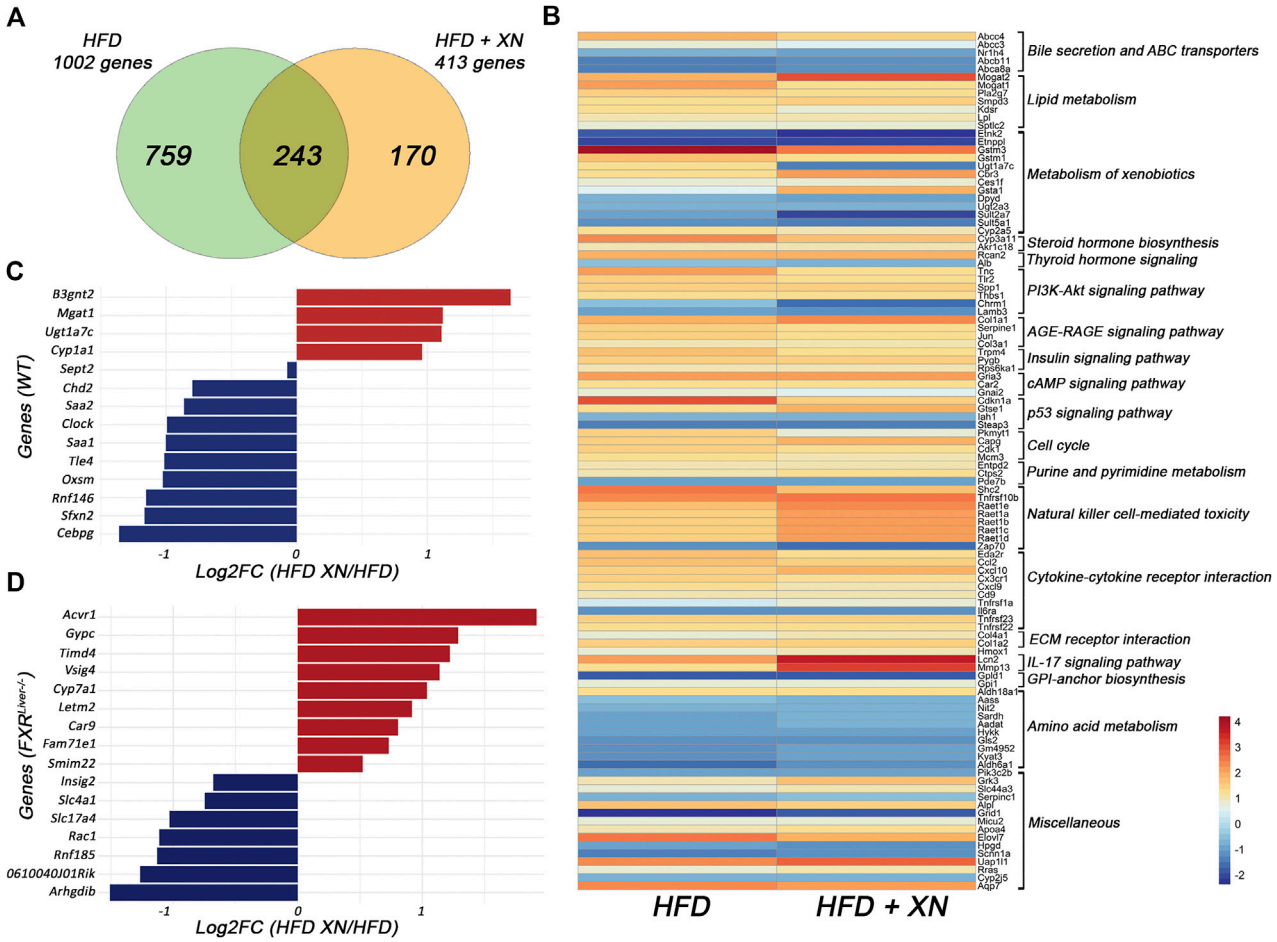

FIGURE 5 | XN alters hepatic gene profiles **(A)** Venn diagram comparing genes in HFD-fed mice vs. HFD-fed mice treated with XN **(B)** Relative expression of 106 shared genes classified according to KEGG pathway (Log2FC of $FXR^{Liver-/-}$ ± XN/WT) **(C)** Genes regulated by XN in WT mice **(D)** Genes regulated by XN in $FXR^{Liver-/-}$ mice. Values are mean of n = 4-5 mice (males and females) per group.

PXR and VDR to regulate TG, cholesterol, glucose, and energy homeostasis (Schaap et al., 2014; Fiorucci and Distrutti, 2015). BAs inhibit their own synthesis mainly via FXR-mediated negative feedback of CYP7A1, the rate-limiting enzyme in the catabolism of cholesterol into BAs (Goodwin et al., 2000). As a result, FXR knockout mice exhibit dyslipidemia (Sinal et al., 2000; Kok et al., 2003) and hepatic steatosis that progresses to NASH (Armstrong and Guo, 2017).

Liver histology revealed lipid vacuoles characteristic of fatty liver disease in $FXR^{Liver-/-}$ mice and elevated hepatic levels of triglycerides, free cholesterol, cholesterol esters and ceramides. The increased liver enzymes indicate liver injury associated with inflammation and NASH. HFD-induced dyslipidemia was aggravated by FXR deficiency with sex differences. Risk factors associated with HFD-induced obesity including fasting glucose and dyslipidemia were more pronounced in males than females. This is in accordance with previous report that male western diet-fed $FXR^{-/-}$ mice had higher hepatic and serum lipids than their female counterparts (Sheng et al., 2017). In fact, NRs play a crucial role in the calibration of sex-specific metabolic pathways and androsterone (Wang et al., 2006) as well as estrogen (Song et al., 2014) were reported to modulate FXR activity. This suggests that interactions between the receptor and gonadal hormones warrant further investigation. The severity of the $FXR^{Liver-/-}$ phenotype was further aggravated by the accumulation of BAs, which regulate several signaling pathways independent from FXR. In our study, elevated BA concentrations in the plasma of $FXR^{Liver-/-}$ mice were accompanied with higher concentrations of BAs in the hippocampus indicating passage of BAs through the BBB. The concentrations measured in HFD-fed $FXR^{Liver-/-}$ mice are comparable to BA concentrations in the blood and brain of $FXR^{-/-}$ mice with hepatic enceph-alopathy (Huang et al., 2015). In fact, in pathological conditions such as acute liver failure and cholestasis, elevated plasma BAs were reported to increase permeability of the BBB (Quinn et al., 2014; McMillin et al., 2016), warranting the investigation of therapeutic alternatives regulating BA concentrations.

XN anti-hyperlipidemic effect was more accentuated in male $FXR^{Liver-/-}$ mice that developed severe dyslipidemia. Our observation that XN and metabolites reached higher

TABLE 3 | Hepatic genes regulated by XN in WT and $FXR^{Liver-/-}$ mice and their roles in metabolic function.

Genes	Log2FC	Definition	Function	Ref
		WT mice (HFD + XN/HFD)		
B3gnt2	+ 1.64	Beta-1,3-N-acetylglucosaminyltransferase	lycosphingolipid biosynthesis	Togayachi et al. (2010)
Mgat1	+ 1.12	Monoacylglycerol acyltransferase	Triglyceride synthesis	Lee and Kim. (2017)
Ugt1a7c	+ 1.1	Uridine 5′-diphospho-glucuronosyltransferase 1A7c	Lipid and xenobiotic metabolism	Guillemette. (2003)
Cyp1a1	+ 0.96	Cytochrome P450 1A1	Lipid and xenobiotic metabolism	Stejskalova and Pavek. (2011)
Sept2	–0.07	Septin 2	Apoptosis and cell proliferation	Cao et al. (2015)
Chd2	–0.79	Chromodomain-helicase-DNA-binding protein 2	Epigenetic signature during liver development	Lu et al. (2012)
Clock	–0.99	Circadian locomotor output cycles kaput protein	Energy metabolism and obesity	Turek et al. (2005). Vieira et al. (2014)
Saa1	–1.00	Serum amyloid a protein 1	Inflammation and systemic complications of obesity	Poitou et al. (2005). Jumeau et al. (2019)
Saa2	–0.86	Serum amyloid a protein 2		
Tle4	–1.01	Transducin-like enhancer protein 4	Transcriptional corepressor associated with type 2 diabetes	Ali, (2013)
Oxsm	–1.02	3-Oxoacyl-ACP synthase II	Fatty acid metabolism	Gao et al. (2019)
Rnf146	–1.15	E3 ubiquitin-protein ligase Rnf146	Energy metabolism	Matsumoto et al. (2017)
Sfxn2	–1.16	Sideroflexin2	Mitochondrial biogenesis	Mon et al. (2019)
Cebpg	–1.36	CCAAT/enhancer binding protein gamma	Transcriptional regulation of adipogenesis and inflammation	Ray et al. (1995); Tanaka et al. (1997)
		$FXR^{Liver-/-}$ mice (HFD + XN/HFD)		
Acvr1	+ 1.92	Activin a receptor type 1	TGF-β signaling pathway	Rao et al. (2017)
Gypc	+ 1.29	Glycophorin C	Membrane properties of erythrocytes	Yiangou et al. (2016)
Timd4	+ 1.22	T-cell immunoglobulin and mucin domain containing 4	Adaptative immunity	Albacker et al. (2010). Dai et al. (2020)
Vsig4	+ 1.14	V-set and immunoglobulin domain-containing 4	Macrophage-mediated hepatic inflammation	Li et al. (2019)
Cyp7a1	+ 1.04	Cholesterol 7 alpha-monooxygenase	Cholesterol and bile acid metabolism	Chiang, (2009)
Letm2	+ 0.92	Leucine zipper and EF-hand containing transmembrane 2	Mitochondrial ion uptake	Waldeck-Weiermair et al. (2011)
Car9	+ 0.8	Carbonic anhydrase 9	Hypoxia-inducible	Olive et al. (2001)
Fam71e1	+ 0.73	Family with sequence similarity 71 member E1		
Smim22	+ 0.52	Small integral membrane protein 22	Cell proliferation	Polycarpou-Schwarz et al. (2018); Li et al. (2019)
Insig2	–0.68	Insulin induced gene 2	Lipid and glucose metabolism	Dong and Tang, (2010)
Slc4a1	–0.74	Solute carrier family 4A1	Efflux transport	Hediger et al. (2004)
Slc17a4	–1.03	Solute carrier family 17A4		
Rac1	–1.11	Ras-related C3 botulinum toxin substrate 1	Cell proliferation	Choi et al. (2006)
Rnf185	–1.13	E3 ubiquitin-protein ligase Rnf185	Autophagy	Tang et al. (2011)
0610040J01Rik	–1.26	RIKEN cDNA 0610040J01 gene		
Arhgdib	–1.51	Anti-rho guanosine diphosphate dissociation inhibitor beta	Liver fibrosis	Utsunomiya et al. (2007)

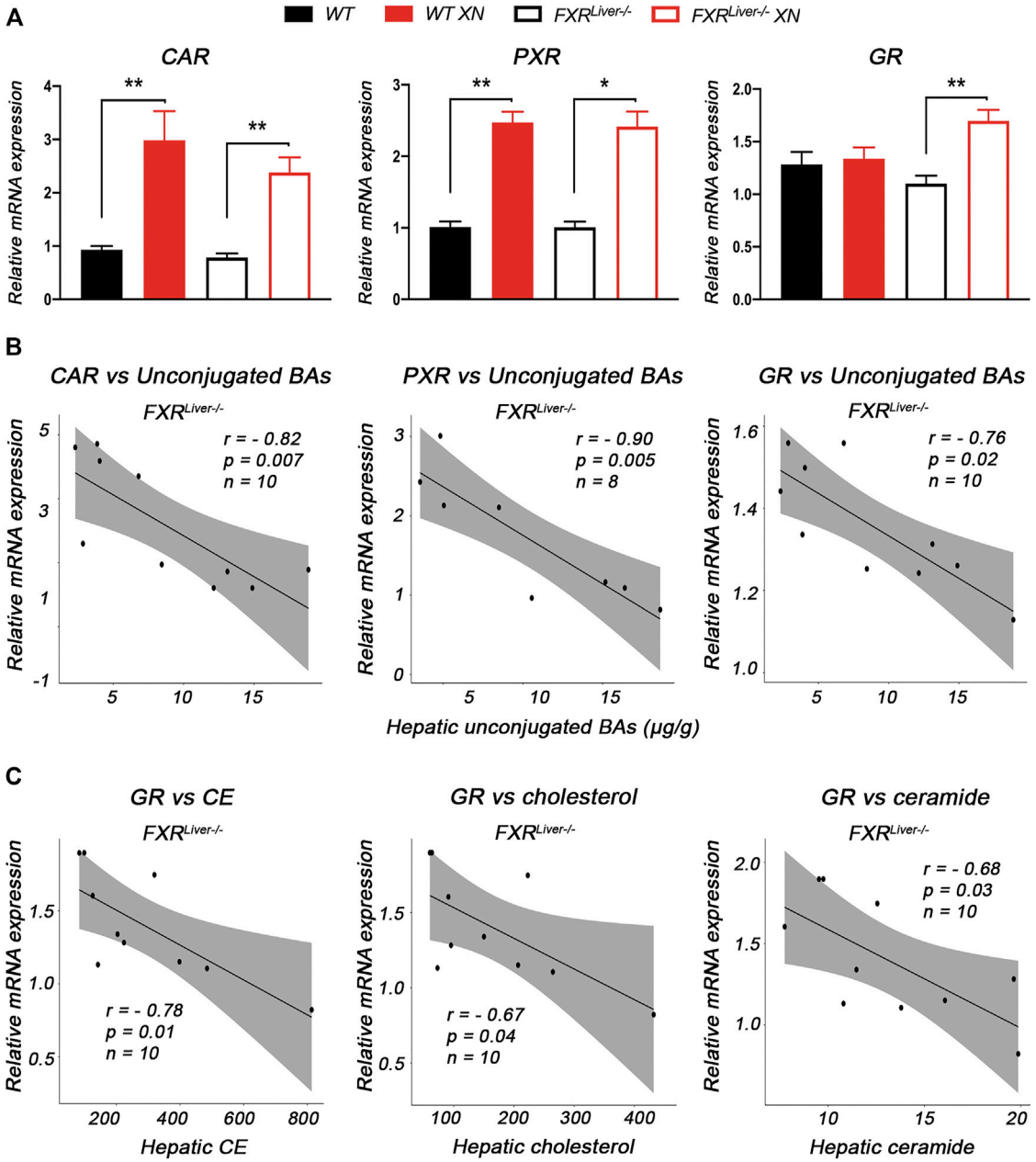

FIGURE 6 | XN induces hepatic expression of FXR-independent NRs **(A)** Quantitative relative expression of CAR, PXR and GR in the liver of HFD-fed WT and $FXR^{Liver-/-}$ mice **(B)** Correlations between relative mRNA expression of CAR, PXR and GR vs. hepatic unconjugated BAs concentrations in $FXR^{Liver-/-}$ mice **(C)** Correlations between relative mRNA expression of GR vs. relative abundances of hepatic CE, cholesterol and ceramide in $FXR^{Liver-/-}$ mice. Bar graphs values are mean ± SEM, n = 4-5 mice (males and females) per group. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ for effect of XN treatment.

concentrations in female mice suggests that pharmacodynamic effect of XN *in vivo* depends more on the severity of the phenotype and less on XN concentrations in biological tissues. Ceramide abundances were heavily influenced by their precursor, SM. The lipotoxicity of ceramides is well-documented (Summers, 2006; Chaurasia and Summers, 2015), but the role of SM in NAFLD and NASH is controversial. Several studies report lower SM levels in NASH patients (Kartsoli et al., 2020). In fact, SM are important components of biological cell membranes (Slotte and Ramstedt, 2007) with no demonstrated intrinsic lipotoxicity. The increase in hepatic SM in XN-treated WT males was accompanied by an increase in hepatic ceramides that did not correlate with metabolic improvements. Therefore, we used the ratio of ceramides over total sphingolipids to estimate ceramide relative abundances more precisely. This ratio was decreased in XN-treated WT and $FXR^{Liver-/-}$ mice.

Taken together, XN protected $FXR^{Liver-/-}$ mice from liver damage, as evaluated by liver transaminase activity, liver

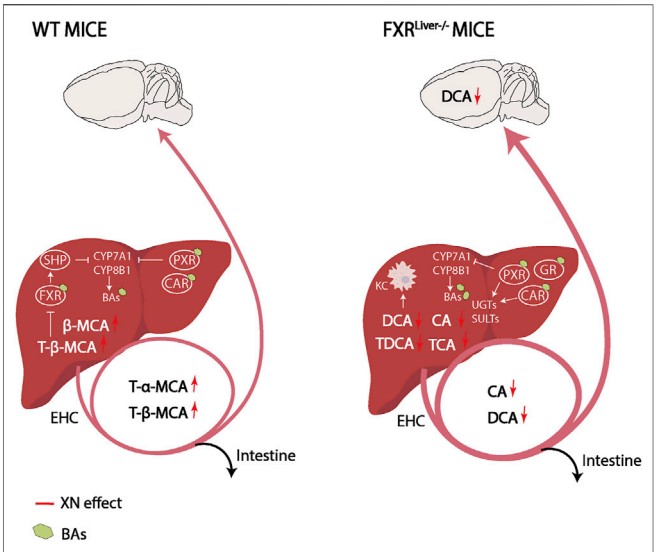

FIGURE 7 | Working hypothesis on XN mechanism of control of BA synthesis and detoxification in WT and $FXR^{Liver-/-}$ mice. XN promotes BA synthesis in WT mice and increases FXR antagonists in the liver and plasma. This results in downstream expression of enzymes involved in BA synthesis, CYP7A1 and CYP8B1. Negative feedback on BA synthesis is exerted by BAs agonists of FXR, PXR and CAR. Compared to WT mice, $FXR^{Liver-/-}$ mice had increased BA synthesis, bigger BA pool sizes and increased passage of BAs through the BBB. In $FXR^{Liver-/-}$ mice, BAs such as CA and DCA impair hepatic Kupffer cells (KCs) activity, which XN might attenuate by decreasing CA and DCA concentrations. XN-mediated activation of PXR and CAR, and GR slows down *de novo* BA synthesis by inhibition of CYP7A1 and induces metabolizing enzymes to stimulate BA excretion. Abbreviations: EHC (Enterohepatic circulation), KC (Kupffer cell).

histopathology and hepatic expression levels of anti-inflammatory genes. These observations indicate that XN effect is not exclusively mediated by hepatic FXR. We hypothesized that the observed XN effect might also derive from its potential as SBARM and from XN-dependent modulation of BA composition because BAs regulate energy expenditure in mice (Chiang, 2002). In WT mice, XN treatment led to an FXR-dependent increase in the most hydrophilic BAs among which are T-α-MCA and T-β-MCA. FXR antagonists, including T-α-MCA and T-β-MCA, have been reported to improve HFD-induced metabolic dysfunction by inducing thermogenesis and repressing intestinal FXR-FGF15 signaling (Sayin et al., 2013; Jiang et al., 2015). Intestinal FXR-FGF15 signaling regulates CYP7A1 gene expression (Kim et al., 2007). In turn, CYP7A1 regulates T-α-MCA and T-β-MCA synthesis, while CYP8B1 is required for TCA synthesis (Jiang et al., 2015; Chiang, 2009; Qi et al., 2015). In accordance with our results, previous studies have demonstrated that CYP7A1 mRNA levels are induced after XN treatment (Nozawa, 2005; Paraiso et al., 2020). This resulted in the increase of T-α-MCA and T-β-MCA, which repress intestinal FXR-FGF15 signaling to increase hepatic BA synthesis and prevent HFD-induced insulin resistance and obesity (Li et al., 2013; Sayin et al., 2013; Jiang et al., 2015; Gonzalez et al., 2016). Contrary to WT mice, XN decreased hydrophobic BAs such as DCA, CA and their taurine conjugates in $FXR^{Liver-/-}$ mice. Moreover, XN promoted genes that attenuate macrophage-mediated inflammation suggesting a shift toward detoxification in absence of hepatic FXR (**Figure 7**). Increased concentrations of hydrophobic BAs impairs phagocytosis activity of tissue-resident macrophages called Kupffer cells (KCs), induce neutrophil-mediated inflammation and alter hepatic T-cell immunity (Zhu et al., 2016). The selective depletion of liver-resident KCs restores hepatic insulin sensitivity and improves whole-body and hepatic fat accumulation (Neyrinck et al., 2009; Huang et al., 2010).

The major mechanisms underlying XN-mediated attenuation of liver damage in $FXR^{Liver-/-}$ mice are the reduction of BA concentrations and the mitigation of hepatic inflammation due the activation of NRs CAR/PXR/GR. FXR, PXR and CAR have complementary roles in the protection against BA toxicity (Guo et al., 2003). RNA sequencing revealed that several genes involved in metabolism of xenobiotics such as CYPs, UGTs, SULTs and GSTs were regulated by XN in $FXR^{Liver-/-}$ mice. Conjugation of hydrophilic groups by UGTs, SULTs, and GSTs increases the water solubility of BAs and xenobiotics to facilitate their renal elimination (Garcia et al., 2018). Hepatic xenobiotic-sensing receptors CAR and PXR mediate phase I and II BA metabolism by regulating CYP450s, UGTs, SULTs and GSTs that catalyze synthesis, oxidation, sulfonation and glucuronidation of BAs (Keppler, 2011; Garcia et al., 2018; Lv and Huang, 2020). Phase III clearance of BA is also regulated by FXR, PXR and CAR. Once BAs are transformed into more hydrophilic metabolites in the liver, they are pumped into the bile via efflux transporters BSEP and MRP2 as a route for fecal elimination (Wagner et al., 2005; Garcia et al., 2018). Moreover, GR enhances CAR/PXR-mediated transcriptional regulation of target genes such as UGT1A1 (Sugatani et al., 2005). CAR, PXR and GR are activated by unconjugated BAs such as LCA, CA, DCA and UDCA (Tanaka and Makino, 1992; Staudinger et al., 2001; Moore et al., 2002; Carazo et al., 2017). The strong correlations between these NRs and hepatic unconjugated BA concentrations suggest that BA metabolism in XN-treated $FXR^{Liver-/-}$ mice was primarily regulated by the activation of CAR, PXR and GR. Since the expression levels of these receptors were exclusively correlated with unconjugated BAs, it is conceivable that endogenous BAs might also play a role in the activation of these NRs *in vivo*. Consistent with our observations in $FXR^{Liver-/-}$ mice, taurine conjugated BA species such as TCA are increased in the liver of NASH patients (Lake et al., 2013). The decreased concentrations of hepatic TCA by XN also indicated a normalization of BA metabolism independent from FXR. Due to the affinity of TCA to FXR, it has been hypothesized that TCA is elevated as a compensatory effect to activate the receptor and normalize metabolism (Sheng et al., 2017). GR activation by XN in $FXR^{Liver-/-}$ mice is involved in the lipid-lowering effect of the flavonoid. Partial agonism of GR reverses NAFLD by preventing hepatic TG and cholesterol accumulation (Koorneef et al., 2018). Moreover, the anti-inflammatory effects of GR activation have a positive impact on hepatic lipid accumulation (Rando and Wahli, 2011; Scheschowitsch et al., 2017).

Our results support the hypothesis of a compensatory interaction between FXR, CAR and PXR. XN-mediated induction of CAR and PXR was not FXR-dependent. In the absence of FXR, the complementary regulation by CAR, PXR and GR might be involved in XN-mediated decrease of BAs

concentrations. Chronic BA overload and prolonged activation of detoxification pathways may lead to the desensitization of BA-sensing receptors, which would contribute to the chronicity of BA-mediated damage. By improving the efficiency of phase II metabolism and phase III hepatic clearance of BAs, XN alleviated the sustained activation of detoxification pathways, improved BA signaling, lipid metabolism and relieved inflammation. Future studies assessing the pharmacodynamic activity of XN metabolites and quantifying glucuronidated and sulfated BA metabolites are necessary to evaluate XN effect more accurately.

Functional annotation clustering revealed that hepatic genes involved in inflammation and neoplastic processes were altered in $FXR^{Liver-/-}$ mice. Hydrophobic BAs are well-known for their cancer-promoting effects and promote carcinogenesis in several tumor models, including hepatocellular carcinoma, colon cancer, and breast cancer (Nagengast et al., 1995; Debruyne et al., 2001; Kim et al., 2006). Our research sheds new light on the chemopreventive potential of XN and highlights the potential of XN as adjuvant therapy in cancers associated with accumulation of BAs such as bile duct cancer and hepatocellular carcinoma.

CONCLUSION

BA synthesis and transport are tightly regulated by BA and xenobiotic-sensing NRs, which regulate genes in synthesis, metabolism and clearance of BAs and play a critical role in BA detoxification. Our current study extends previous research and shows the novel findings that 1) XN ameliorates HFD-induced inflammation and tissue damage in $FXR^{Liver-/-}$ mice; 2) XN improves HFD-induced dysfunctional lipid and BA metabolism via FXR-dependent and independent signaling, including the induction of CAR, PXR and GR. To the best of our knowledge, the potential of XN as adjuvant therapy in the management of cholestatic diseases has not been reported and merits further investigation.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA687670.

ETHICS STATEMENT

The animal study was reviewed and approved by the Institutional Animal Care Use Committee, Oregon State University, Corvallis, Oregon (USA), IACUC-2019-0001. The animal studies were carried out in accordance with the approved protocol.

AUTHOR CONTRIBUTIONS

IP, CK, JS designed the experiments. IP, CK, AM performed the experiments. IP, TT, GB. PK. JR. CK. JS. analyzed the data and performed statistical analyses. All the authors provided scientific support, wrote and reviewed the manuscript.

FUNDING

This work was funded by the Linus Pauling Institute, the National Institutes of Health (NIH grants S10RR022589, S10RR027878, and R01AT009168-04S1), the OSU College of Pharmacy and the OSU Foundation Buhler-Wang Research Fund.

ACKNOWLEDGMENTS

The authors thank Grace L. Guo for providing the $FXR^{Liver-/-}$ mice and Cristobal Miranda for providing technical guidance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.643857/full#supplementary-material.

REFERENCES

Albacker, L. A., Karisola, P., Chang, Y. J., Umetsu, S. E., Zhou, M., Akbari, O., et al. (2010). TIM-4, a receptor for phosphatidylserine, controls adaptive immunity by regulating the removal of antigen-specific T cells. *J. Immunol.* 185, 6839–6849. doi:10.4049/jimmunol.1001360

Ali, O. (2013). Genetics of type 2 diabetes. *Wjd* 4 (4), 114–123. doi:10.4239/wjd.v4.i4.114

Armstrong, L. E., and Guo, G. L. (2017). Role of FXR in liver inflammation during nonalcoholic steatohepatitis. *Curr. Pharmacol. Rep.* 3 (2), 92–100. doi:10.1007/s40495-017-0085-2

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodological)* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Cajka, T., Smilowitz, J. T., and Fiehn, O. (2017). Validating quantitative untargeted lipidomics across nine liquid chromatography-high-resolution mass spectrometry platforms. *Anal. Chem.* 89 (22), 12360–12368. doi:10.1021/acs.analchem.7b03404

Cao, L. Q., Shao, Z. L., Liang, H. H., Zhang, D. W., Yang, X. W., Jiang, X. F., et al. (2015). Activation of peroxisome proliferator-activated receptor-γ (PPARγ) inhibits hepatoma cell growth via downregulation of SEPT2 expression. *Cancer Lett.* 359 (1), 127–135. doi:10.1016/j.canlet.2015.01.004

Carazo, A., Hyrsova, L., Dusek, J., Chodounska, H., Horvatova, A., Berka, K., et al. (2017). Acetylated deoxycholic (DCA) and cholic (CA) acids are potent ligands of pregnane X (PXR) receptor. *Toxicol. Lett.* 265, 86–96. doi:10.1016/j.toxlet.2016.11.013

Chang, C. N., Singh, A. J., Gross, M. K., and Kioussi, C. (2019). Requirement of Pitx2 for skeletal muscle homeostasis. *Dev. Biol.* 445 (1), 90–102. doi:10.1016/j.ydbio.2018.11.001

Chaurasia, B., and Summers, S. A. (2015). Ceramides - lipotoxic inducers of metabolic disorders. *Trends Endocrinol. Metab.* 26 (10), 538–550. doi:10.1016/j.tem.2015.07.006

Chawla, A., Repa, J. J., Evans, R. M., and Mangelsdorf, D. J. (2001). Nuclear receptors and lipid physiology: opening the X-files. *Science* 294 (5548), 1866. doi:10.1126/science.294.5548.1866

Chiang, J. Y. L. (2013). Bile acid metabolism and signaling. *Compr. Physiol.* 3 (3), 1191–1212. doi:10.1002/cphy.c120023

Chiang, J. Y. L. (2017). Bile acid metabolism and signaling in liver disease and therapy. *Liver Res.* 1 (1), 3–9. doi:10.1016/j.livres.2017.05.001

Chiang, J. Y. L. (2002). Bile acid regulation of gene expression: roles of nuclear hormone receptors. *Endocr. Rev.* 23 (4), 443–463. doi:10.1210/er.2000-0035

Chiang, J. Y. L. (2009). Bile acids: regulation of synthesis. *J. lipid Res.* 50 (10), 1955–1966. doi:10.1194/jlr.r900010-jlr200

Choi, J., Leonard, S. W., Kasper, K., McDougall, M., Stevens, J. F., Tanguay, R. L., et al. (2015). Novel function of vitamin E in regulation of zebrafish (*Danio rerio*) brain lysophospholipids discovered using lipidomics. *J. Lipid Res.* 56 (6), 1182–1190. doi:10.1194/jlr.m058941

Choi, S. S., Sicklick, J. K., Ma, Q., Yang, L., Huang, J., Qi, Y., et al. (2006). Sustained activation of Rac1 in hepatic stellate cells promotes liver injury and fibrosis in mice. *Hepatology* 44 (5), 1267–1277. doi:10.1002/hep.21375

Dai, W., Zhang, B., Jiang, X. M., Su, H., Li, J., Zhao, Y., et al. (2020). Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* 368 (6497), 1331. doi:10.1126/science.abb4489

Debruyne, P. R., Bruyneel, E. A., Li, X., Zimber, A., Gespach, C., and Mareel, M. M. (2001). The role of bile acids in carcinogenesis. *Mutat. Research/Fundamental Mol. Mech. Mutagenesis* 480-481, 359–369. doi:10.1016/s0027-5107(01)00195-6

Dong, X.-Y., and Tang, S.-Q. (2010). Insulin-induced gene: a new regulator in lipid metabolism. *Peptides* 31 (11), 2145–2150. doi:10.1016/j.peptides.2010.07.020

Fiorucci, S., and Distrutti, E. (2015). Bile acid-activated receptors, intestinal microbiota, and the treatment of metabolic disorders. *Trends Mol. Med.* 21 (11), 702–714. doi:10.1016/j.molmed.2015.09.001

Forman, B. M., Goode, E., Chen, J., Oro, A. E., Bradley, D. J., Perlmann, T., et al. (1995). Identification of a nuclear receptor that is activated by farnesol metabolites. *Cell* 81 (5), 687–693. doi:10.1016/0092-8674(95)90530-8

Gao, J., and Xie, W. (2012). Targeting xenobiotic receptors PXR and CAR for metabolic diseases. *Trends Pharmacol. Sci.* 33 (10), 552–558. doi:10.1016/j.tips.2012.07.003

Gao, T., Qian, S., Shen, S., Zhang, X., Liu, J., Jia, W., et al. (2019). Reduction of mitochondrial 3-oxoacyl-ACP synthase (OXSM) by hyperglycemia is associated with deficiency of α-lipoic acid synthetic pathway in kidney of diabetic mice. *Biochem. Biophysical Res. Commun.* 512 (1), 106–111. doi:10.1016/j.bbrc.2019.02.155

Garcia, M., Thirouard, L., Sedès, L., Monrose, M., Holota, H., Caira, F., et al. (2018). Nuclear receptor metabolism of bile acids and xenobiotics: a coordinated detoxification System with impact on health and diseases. *Ijms* 19 (11), 3630. doi:10.3390/ijms19113630

Gonzalez, F. J., Jiang, C., and Patterson, A. D. (2016). An intestinal microbiota-farnesoid X receptor *Axis* modulates metabolic disease. *Gastroenterology* 151 (5), 845–859. doi:10.1053/j.gastro.2016.08.057

Goodwin, B., Jones, S. A., Price, R. R., Watson, M. A., McKee, D. D., Moore, L. B., et al. (2000). A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. *Mol. Cel* 6 (3), 517–526. doi:10.1016/s1097-2765(00)00051-4

Guillemette, C. (2003). Pharmacogenomics of human UDP-glucuronosyltransferase enzymes. *Pharmacogenomics J.* 3 (3), 136–158. doi:10.1038/sj.tpj.6500171

Guo, G. L., Lambert, G., Negishi, M., Ward, J. M., Brewer, H. B., Kliewer, S. A., et al. (2003). Complementary roles of farnesoid X receptor, pregnane X receptor, and constitutive androstane receptor in protection against bile acid toxicity. *J. Biol. Chem.* 278 (46), 45062–45071. doi:10.1074/jbc.m307145200

Halilbasic, E., Baghdasaryan, A., and Trauner, M. (2013). Nuclear receptors as drug targets in cholestatic liver diseases. *Clin. Liver Dis.* 17 (2), 161–189. doi:10.1016/j.cld.2012.12.001

Hediger, M. A., Romero, M. F., Peng, J.-B., Rolfs, A., Takanaga, H., and Bruford, E. A. (2004). The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. *Pflugers Archiv Eur. J. Physiol.* 447 (5), 465–468. doi:10.1007/s00424-003-1192-y

Huang, C., Wang, J., Hu, W., Wang, C., Lu, X., Tong, L., et al. (2016). Identification of functional farnesoid X receptors in brain neurons. *FEBS Lett.* 590 (18), 3233–3242. doi:10.1002/1873-3468.12373

Huang, F., Wang, T., Lan, Y., Yang, L., Pan, W., Zhu, Y., et al. (2015). Deletion of mouse FXR gene disturbs multiple neurotransmitter systems and alters neurobehavior. *Front. Behav. Neurosci.* 9, 70. doi:10.3389/fnbeh.2015.00070

Huang, W., Metlakunta, A., Dedousis, N., Zhang, P., Sipula, I., Dube, J. J., et al. (2010). Depletion of liver Kupffer cells prevents the development of diet-induced hepatic steatosis and insulin resistance. *diabetes* 59 (2), 347–357. doi:10.2337/db09-0016

Jiang, C., Xie, C., Lv, Y., Li, J., Krausz, K. W., Shi, J., et al. (2015). Intestine-selective farnesoid X receptor inhibition improves obesity-related metabolic dysfunction. *Nat. Commun.* 6, 10166. doi:10.1038/ncomms10166

Jumeau, C., Awad, F., Assrawi, E., Cobret, L., Duquesnoy, P., Giurgea, I., et al. (2019). Expression of SAA1, SAA2 and SAA4 genes in human primary monocytes and monocyte-derived macrophages. *PLOS ONE* 14 (5), e0217005. doi:10.1371/journal.pone.0217005

Kardol-Hoefnagel, T., van Logtestijn, S. A. L. M., and Otten, H. G. (2020). A review on the function and regulation of ARHGDIB/RhoGDI2 expression including the hypothetical role of ARHGDIB/RhoGDI2 autoantibodies in kidney transplantation. *Transplant. Direct* 6 (5), e548. doi:10.1097/txd.0000000000000993

Kartsoli, S., Kostara, C. E., Tsimihodimos, V., Bairaktari, E. T., and Christodoulou, D. K. (2020). Lipidomics in non-alcoholic fatty liver disease. *Wjh* 12 (8), 436–450. doi:10.4254/wjh.v12.i8.436

Keppler, D. (2011). "Multidrug resistance proteins (MRPs, ABCCs): importance for pathophysiology and drug therapy," in *Drug transporters*. Editors M. F. Fromm and R. B. Kim (Berlin, Heidelberg: Springer), 299–323. doi:10.1007/978-3-642-14541-4_8

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37 (8), 907–915. doi:10.1038/s41587-019-0201-4

Kim, I., Ahn, S. H., Inagaki, T., Choi, M., Ito, S., Guo, G. L., et al. (2007). Differential regulation of bile acid homeostasis by the farnesoid X receptor in liver and intestine. *J. Lipid Res.* 48 (12), 2664–2672. doi:10.1194/jlr.m700330-jlr200

Kim, I., Morimura, K., Shah, Y., Yang, Q., Ward, J. M., and Gonzalez, F. J. (2006). Spontaneous hepatocarcinogenesis in farnesoid X receptor-null mice. *Carcinogenesis* 28 (5), 940–946. doi:10.1093/carcin/bgl249

Kirkwood, J. S., Legette, L. L., Miranda, C. L., Jiang, Y., and Stevens, J. F. (2013). A metabolomics-driven elucidation of the anti-obesity mechanisms of xanthohumol. *J. Biol. Chem.* 288 (26), 19000–19013. doi:10.1074/jbc.m112.445452

Kok, T., Hulzebos, C. V., Wolters, H., Havinga, R., Agellon, L. B., Stellaard, F., et al. (2003). Enterohepatic circulation of bile salts in farnesoid X receptor-deficient mice. *J. Biol. Chem.* 278 (43), 41930–41937. doi:10.1074/jbc.m306309200

Kong, B., Zhu, Y., Li, G., Williams, J. A., Buckley, K., Tawfik, O., et al. (2016). Mice with hepatocyte-specific FXR deficiency are resistant to spontaneous but susceptible to cholic acid-induced hepatocarcinogenesis. *Am. J. Physiology-Gastrointestinal Liver Physiol.* 310 (5), G295–G302. doi:10.1152/ajpgi.00134.2015

Koorneef, L. L., van den Heuvel, J. K., Kroon, J., Boon, M. R., Hoen, P. A. C., Hettne, K. M., et al. (2018). Selective glucocorticoid receptor modulation prevents and reverses nonalcoholic fatty liver disease in male mice. *Endocrinology* 159 (12), 3925–3936. doi:10.1210/en.2018-00671

Lake, A. D., Novak, P., Shipkova, P., Aranibar, N., Robertson, D., Reily, M. D., et al. (2013). Decreased hepatotoxic bile acid composition and altered synthesis in progressive human nonalcoholic fatty liver disease. *Toxicol. Appl. Pharmacol.* 268 (2), 132–140. doi:10.1016/j.taap.2013.01.022

Larter, C. Z., Chitturi, S., Heydet, D., and Farrell, G. C. (2010). A fresh look at NASH pathogenesis. Part 1: the metabolic movers. *J. Gastroenterol. Hepatol.* 25 (4), 672–690. doi:10.1111/j.1440-1746.2010.06253.x

Lee, F. Y., Lee, H., Hubbert, M. L., Edwards, P. A., and Zhang, Y. (2006). FXR, a multipurpose nuclear receptor. *Trends Biochem. Sci.* 31 (10), 572–580. doi:10.1016/j.tibs.2006.08.002

Lee, Y. J., and Kim, J. W. (2017). Monoacylglycerol O-acyltransferase 1 (MGAT1) localizes to the ER and lipid droplets promoting triacylglycerol synthesis. *BMB Rep.* 50 (7), 367–372. doi:10.5483/bmbrep.2017.50.7.036

Li, F., Ji, J. P., Xu, Y., and Liu, R. L. (2019). Identification a novel set of 6 differential expressed genes in prostate cancer that can potentially predict biochemical

recurrence after curative surgery. *Clin. Transl Oncol.* 21 (8), 1067–1075. doi:10.1007/s12094-018-02029-z

Li, F., Jiang, C., Krausz, K. W., Li, Y., Albert, I., Hao, H., et al. (2013). Microbiome remodelling leads to inhibition of intestinal farnesoid X receptor signalling and decreased obesity. *Nat. Commun.* 4, 2384. doi:10.1038/ncomms3384

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi:10.1093/bioinformatics/btp352

Li, Y., Sun, J. P., Wang, J., Lu, W. H., Xie, L. Y., Lv, J., et al. (2019). Expression of Vsig4 attenuates macrophage-mediated hepatic inflammation and fibrosis in high fat diet (HFD)-induced mice. *Biochem. Biophysical Res. Commun.* 516 (3), 858–865. doi:10.1016/j.bbrc.2019.06.045

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

Lu, H., Cui, J., Gunewardena, S., Yoo, B., Zhong, X.-b., and Klaassen, C. (2012). Hepatic ontogeny and tissue distribution of mRNAs of epigenetic modifiers in mice using RNA-sequencing. *Epigenetics* 7 (8), 914–929. doi:10.4161/epi.21113

Lv, C., and Huang, L. (2020). Xenobiotic receptors in mediating the effect of sepsis on drug metabolism. *Acta Pharmaceutica Sinica. B* 10 (1), 33–41. doi:10.1016/j.apsb.2019.12.003

Ma, K., Saha, P. K., Chan, L., and Moore, D. D. (2006). Farnesoid X receptor is essential for normal glucose homeostasis. *J. Clin. Invest.* 116 (4), 1102–1109. doi:10.1172/jci25604

Makishima, M., Okamoto, A. Y., Repa, J. J., Tu, H., Learned, R. M., Luk, A., et al. (1999). Identification of a nuclear receptor for bile acids. *Science* 284 (5418), 1362–1365. doi:10.1126/science.284.5418.1362

Matsumoto, Y., La Rose, J., Lim, M., Adissu, H. A., Law, N., Mao, X., et al. (2017). Ubiquitin ligase RNF146 coordinates bone dynamics and energy metabolism. *J. Clin. Invest.* 127 (7), 2612–2625. doi:10.1172/jci92233

McMillin, M., and DeMorrow, S. (2016). Effects of bile acids on neurological function and disease. *FASEB j.* 30 (11), 3658–3668. doi:10.1096/fj.201600275r

McMillin, M., Frampton, G., Quinn, M., Ashfaq, S., de los Santos, M., Grant, S., et al. (2016). Bile acid signaling is involved in the neurological decline in a murine model of acute liver failure. *Am. J. Pathol.* 186 (2), 312–323. doi:10.1016/j.ajpath.2015.10.005

Miranda, C. L., Elias, V. D., Hay, J. J., Choi, J., Reed, R. L., and Stevens, J. F. (2016). Xanthohumol improves dysfunctional glucose and lipid metabolism in diet-induced obese C57BL/6J mice. *Arch. Biochem. Biophys.* 599, 22–30. doi:10.1016/j.abb.2016.03.008

Miranda, C. L., Johnson, L. A., de Montgolfier, O., Elias, V. D., Ullrich, L. S., Hay, J. J., et al. (2018). Non-estrogenic xanthohumol derivatives mitigate insulin resistance and cognitive impairment in high-fat diet-induced obese mice. *Sci. Rep.* 8 (1), 613. doi:10.1038/s41598-017-18992-6

Mon, E. E., Wei, F. Y., Ahmad, R. N. R., Yamamoto, T., Moroishi, T., and Tomizawa, K. (2019). Regulation of mitochondrial iron homeostasis by sideroflexin 2. *J. Physiol. Sci.* 69 (2), 359–373. doi:10.1007/s12576-018-0652-2

Moore, L. B., Maglich, J. M., McKee, D. D., Wisely, B., Willson, T. M., Kliewer, S. A., et al. (2002). Pregnane X receptor (PXR), constitutive androstane receptor (CAR), and benzoate X receptor (BXR) define three pharmacologically distinct classes of nuclear receptors. *Mol. Endocrinol.* 16 (5), 977–986. doi:10.1210/mend.16.5.0828

Nagengast, F. M., Grubben, M. J. A. L., and van Munster, I. P. (1995). Role of bile acids in colorectal carcinogenesis. *Eur. J. Cancer* 31 (7), 1067–1070. doi:10.1016/0959-8049(95)00216-6

Neyrinck, A. M., Cani, P. D., Dewulf, E. M., De Backer, F., Bindels, L. B., and Delzenne, N. M. (2009). Critical role of Kupffer cells in the management of diet-induced diabetes and obesity. *Biochem. biophysical Res. Commun.* 385 (3), 351–356. doi:10.1016/j.bbrc.2009.05.070

Nozawa, H. (2005). Xanthohumol, the chalcone from beer hops (*Humulus lupulus L.*), is the ligand for farnesoid X receptor and ameliorates lipid and glucose metabolism in KK-A mice. *Biochem. Biophysical Res. Commun.* 336 (3), 754–761. doi:10.1016/j.bbrc.2005.08.159

Olive, P. L., Aquino-Parsons, C., MacPhail, S. H., Liao, S. Y., Raleigh, J. A., Lerman, M. I., et al. (2001). Carbonic anhydrase 9 as an endogenous marker for hypoxic cells in cervical cancer. *Cancer Res.* 61 (24), 8924–8929.

Paraiso, I. L., Plagmann, L. S., Yang, L., Zielke, R., Gombart, A. F., Maier, C. S., et al. (2019). Reductive metabolism of xanthohumol and 8-prenylnaringenin by the intestinal bacterium eubacterium ramulus. *Mol. Nutr. Food Res.* 63 (2), e1800923. doi:10.1002/mnfr.201970006

Paraiso, I. L., Revel, J. S., Choi, J., Miranda, C. L., Lak, P., Kioussi, C., et al. (2020). Targeting the liver-brain *Axis* with hop-derived flavonoids improves lipid metabolism and cognitive performance in mice. *Mol. Nutr. Food Res.* 64, 2000341. doi:10.1002/mnfr.202000341

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33 (3), 290–295. doi:10.1038/nbt.3122

Poitou, C., Viguerie, N., Cancello, R., De Matteis, R., Cinti, S., Stich, V., et al. (2005). Serum amyloid A: production by human white adipocyte and regulation by obesity and nutrition. *Diabetologia* 48 (3), 519–528. doi:10.1007/s00125-004-1654-6

Polycarpou-Schwarz, M., Groß, M., Mestdagh, P., Schott, J., Grund, S. E., Hildenbrand, C., et al. (2018). The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. *Oncogene* 37 (34), 4750–4768. doi:10.1038/s41388-018-0281-5

Porez, G., Prawitt, J., Gross, B., and Staels, B. (2012). Bile acid receptors as targets for the treatment of dyslipidemia and cardiovascular disease. *J. lipid Res.* 53 (9), 1723–1737. doi:10.1194/jlr.r024794

Qi, Y., Jiang, C., Cheng, J., Krausz, K. W., Li, T., Ferrell, J. M., et al. (2015). Bile acid signaling in lipid metabolism: metabolomic and lipidomic analysis of lipid and bile acid markers linked to anti-obesity and anti-diabetes in mice. *Biochim. Biophys. Acta (Bba) - Mol. Cel Biol. Lipids* 1851 (1), 19–29. doi:10.1016/j.bbalip.2014.04.008

Quinn, M., McMillin, M., Galindo, C., Frampton, G., Pae, H. Y., and DeMorrow, S. (2014). Bile acids permeabilize the blood brain barrier after bile duct ligation in rats via Rac1-dependent mechanisms. *Dig. Liver Dis.* 46 (6), 527–534. doi:10.1016/j.dld.2014.01.159

Rando, G., and Wahli, W. (2011). Sex differences in nuclear receptor-regulated liver metabolic pathways. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1812 (8), 964–973. doi:10.1016/j.bbadis.2010.12.023

Rao, S., Zaidi, S., Banerjee, J., Jogunoori, W., Sebastian, R., Mishra, B., et al. (2017). Transforming growth factor-β in liver cancer stem cells and regeneration. *Hepatol. Commun.* 1 (6), 477–493. doi:10.1002/hep4.1062

Ray, A., Hannink, M., and Ray, B. K. (1995). Concerted participation of NF-κB and C/EBP heteromer in lipopolysaccharide induction of serum amyloid A gene expression in liver. *J. Biol. Chem.* 270 (13), 7365–7374. doi:10.1074/jbc.270.13.7365

Sayin, S. I., Wahlström, A., Felin, J., Jäntti, S., Marschall, H. U., Bamberg, K., et al. (2013). Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cel Metab.* 17 (2), 225–235. doi:10.1016/j.cmet.2013.01.003

Schaap, F. G., Trauner, M., and Jansen, P. L. M. (2014). Bile acid receptors as targets for drug development. *Nat. Rev. Gastroenterol. Hepatol.* 11 (1), 55–67. doi:10.1038/nrgastro.2013.151

Scheschowitsch, K., Leite, J. A., and Assreuy, J. (2017). New insights in glucocorticoid receptor signaling-more than just a ligand-binding receptor. *Front. Endocrinol.* 8, 16. doi:10.3389/fendo.2017.00016

Schmitt, J., Kong, B., Stieger, B., Tschopp, O., Schultze, S. M., Rau, M., et al. (2015). Protective effects of farnesoid X receptor (FXR) on hepatic lipid accumulation are mediated by hepatic FXR and independent of intestinal FGF15 signal. *Liver Int.* 35 (4), 1133–1144. doi:10.1111/liv.12456

Schubring, S. R., Fleischer, W., Lin, J. S., Haas, H. L., and Sergeeva, O. A. (2012). The bile steroid chenodeoxycholate is a potent antagonist at NMDA and GABAA receptors. *Neurosci. Lett.* 506 (2), 322–326. doi:10.1016/j.neulet.2011.11.036

Sheng, L., Jena, P. K., Liu, H. X., Kalanetra, K. M., Gonzalez, F. J., French, S. W., et al. (2017). Gender differences in bile acids and microbiota in relationship with gender dissimilarity in steatosis induced by diet and FXR inactivation. *Scientific Rep.* 7 (1), 1748. doi:10.1038/s41598-017-01576-9

Sinal, C. J., Tohkin, M., Miyata, M., Ward, J. M., Lambert, G., and Gonzalez, F. J. (2000). Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid

and lipid homeostasis. *Cell* 102 (6), 731–744. doi:10.1016/s0092-8674(00)00062-3

Singh, A. J., Chang, C. N., Ma, H. Y., Ramsey, S. A., Filtz, T. M., and Kioussi, C. (2018). FACS-Seq analysis of Pax3-derived cells identifies non-myogenic lineages in the embryonic forelimb. *Scientific Rep.* 8 (1), 7670. doi:10.1038/s41598-018-25998-1

Slotte, J. P., and Ramstedt, B. (2007). The functional role of sphingomyelin in cell membranes. *Eur. J. Lipid Sci. Technol.* 109 (10), 977–981. doi:10.1002/ejlt.200700024

Song, X., Vasilenko, A., Chen, Y., Valanejad, L., Verma, R., Yan, B., et al. (2014). Transcriptional dynamics of bile salt export pump during pregnancy: mechanisms and implications in intrahepatic cholestasis of pregnancy. *Hepatology* 60 (6), 1993–2007. doi:10.1002/hep.27171

Staels, B., and Kuipers, F. (2007). Bile acid sequestrants and the treatment of type 2 diabetes mellitus. *Drugs* 67 (10), 1383–1392. doi:10.2165/00003495-200767100-00001

Staudinger, J. L., Goodwin, B., Jones, S. A., Hawkins-Brown, D., MacKenzie, K. I., LaTour, A., et al. (2001). The nuclear receptor PXR is a lithocholic acid sensor that protects against liver toxicity. *Proc. Natl. Acad. Sci.* 98 (6), 3369. doi:10.1073/pnas.051551698

Stejskalova, L., and Pavek, P. (2011). The function of cytochrome P450 1A1 enzyme (CYP1A1) and aryl hydrocarbon receptor (AhR) in the placenta. *Cpb* 12 (5), 715–730. doi:10.2174/138920111795470994

Sugatani, J., Nishitani, S., Yamakawa, K., Yoshinari, K., Sueyoshi, T., Negishi, M., et al. (2005). Transcriptional regulation of human UGT1A1 gene expression: activated glucocorticoid receptor enhances constitutive androstane receptor/pregnane X receptor-mediated UDP-glucuronosyltransferase 1A1 regulation with glucocorticoid receptor-interacting protein 1. *Mol. Pharmacol.* 67 (3), 845–855. doi:10.1124/mol.104.007161

Summers, S. (2006). Ceramides in insulin resistance and lipotoxicity. *Prog. Lipid Res.* 45 (1), 42–72. doi:10.1016/j.plipres.2005.11.002

Sun, L., Cai, J., and Gonzalez, F. J. (2021). The role of farnesoid X receptor in metabolic diseases, and gastrointestinal and liver cancer. *Nat. Rev. Gastroenterol. Hepatol.*. doi:10.1038/s41575-020-00404-2

Tanaka, H., and Makino, I. (1992). Ursodeoxycholic acid-dependent activation of the glucocorticoid receptor. *Biochem. Biophysical Res. Commun.* 188 (2), 942–948. doi:10.1016/0006-291x(92)91146-h

Tanaka, T., Yoshida, N., Kishimoto, T., and Akira, S. (1997). Defective adipocyte differentiation in mice lacking the C/EBPbeta and/or C/EBPdelta gene. *EMBO J.* 16 (24), 7432–7443. doi:10.1093/emboj/16.24.7432

Tang, F., Wang, B., Li, N., Wu, Y., Jia, J., Suo, T., et al. (2011). RNF185, a novel mitochondrial ubiquitin E3 ligase, regulates autophagy through interaction with BNIP1.6 (9), e24367. doi:10.1371/journal.pone.0024367

Togayachi, A., Kozono, Y., Kuno, A., Ohkura, T., Sato, T., Hirabayashi, J., et al. (2010). β3GnT2 (B3GNT2), a major polylactosamine synthase: analysis of B3gnt2-deficient mice. *Methods Enzymol.* 479, 185–204. doi:10.1016/s0076-6879(10)79011-x

Trauner, M., Claudel, T., Fickert, P., Moustafa, T., and Wagner, M. (2010). Bile acids as regulators of hepatic lipid and glucose metabolism. *Dig. Dis.* 28 (1), 220–224. doi:10.1159/000282091

Turek, F. W., Joshu, C., Kohsaka, A., Lin, E., Ivanova, G., McDearmon, E., et al. (2005). Obesity and metabolic syndrome in circadian clock mutant mice. *Science* 308 (5724), 1043. doi:10.1126/science.1108750

Uppal, H., Toma, D., Saini, S. P. S., Ren, S., Jones, T. J., and Xie, W. (2005). Combined loss of orphan receptors PXR and CAR heightens sensitivity to toxic bile acids in mice. *Hepatology* 41 (1), 168–176. doi:10.1002/hep.20512

Utsunomiya, T., Okamoto, M., Wakiyama, S., Hashimoto, M., Fukuzawa, K., Ezaki, T., et al. (2007). A specific gene-expression signature quantifies the degree of hepatic fibrosis in patients with chronic liver disease. *Wjg* 13 (3), 383–390. doi:10.3748/wjg.v13.i3.383

Vieira, E., Ruano, Eg., Figueroa, A. L., Aranda, G., Momblan, D., Carmona, F., et al. (2014). Altered clock gene expression in obese visceral adipose tissue is associated with metabolic syndrome. *PloS one* 9 (11), e111678. doi:10.1371/journal.pone.0111678

Wagner, M., Halilbasic, E., Marschall, H.-U., Zollner, G., Fickert, P., Langner, C., et al. (2005). CAR and PXR agonists stimulate hepatic bile acid and bilirubin detoxification and elimination pathways in mice. *Hepatology* 42 (2), 420–430. doi:10.1002/hep.20784

Waldeck-Weiermair, M., Jean-Quartier, C., Rost, R., Khan, M. J., Vishnu, N., Bondarenko, A. I., et al. (2011). Leucine zipper EF hand-containing transmembrane protein 1 (Letm1) and uncoupling proteins 2 and 3 (UCP2/3) contribute to two distinct mitochondrial Ca2+ uptake pathways. *J. Biol. Chem.* 286 (32), 28444–28455. doi:10.1074/jbc.m111.244517

Wang, S., Lai, K., Moy, F. J., Bhat, A., Hartman, H. B., and Evans, M. J. (2006). The nuclear hormone receptor farnesoid X receptor (FXR) is activated by androsterone. *Endocrinology* 147 (9), 4025–4033. doi:10.1210/en.2005-1485

Xiao, Y., Yan, W., Zhou, K., Cao, Y., and Cai, W. (2016). Glucocorticoid treatment alters systemic bile acid homeostasis by regulating the biosynthesis and transport of bile salts. *Dig. Liver Dis.* 48 (7), 771–779. doi:10.1016/j.dld.2016.03.022

Yang, J. Y., Della-Fera, M. A., Rayalam, S., and Baile, C. A. (2007). Effect of xanthohumol and isoxanthohumol on 3T3-L1 cell apoptosis and adipogenesis. *Apoptosis* 12 (11), 1953–1963. doi:10.1007/s10495-007-0130-4

Yang, L., Broderick, D., Campbell, Y., Gombart, A. F., Stevens, J. F., Jiang, Y., et al. (2016). Conformational modulation of the farnesoid X receptor by prenylflavonoids: insights from hydrogen deuterium exchange mass spectrometry (HDX-MS), fluorescence titration and molecular docking studies. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1864 (12), 1667–1677. doi:10.1016/j.bbapap.2016.08.019

Yiangou, L., Montandon, R., Modrzynska, K., Rosen, B., Bushell, W., Hale, C., et al. (2016). A stem cell strategy identifies glycophorin C as a major erythrocyte receptor for the rodent malaria parasite plasmodium berghei 11 (6), e0158238. doi:10.1371/journal.pone.0158238

Zhao, A., Yu, J., Lew, J. L., Huang, L., Wright, S. D., and Cui, J. (2004). Polyunsaturated fatty acids are FXR ligands and differentially regulate expression of FXR targets. *DNA Cel Biol.* 23 (8), 519–526. doi:10.1089/1044549041562267

Zhu, C., Fuchs, C. D., Halilbasic, E., and Trauner, M. (2016). Bile acids in regulation of inflammation and immunity: friend or foe?. *Clin. Exp. Rheumatol.* 34 (4 Suppl. 98), 25–31.

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.

frontiers in Plant Science

REVIEW
published: 07 June 2021
doi: 10.3389/fpls.2021.670103

Metabolic Engineering of Isoflavones: An Updated Overview

Soo In Sohn[1], Subramani Pandian[1], Young Ju Oh[2], Hyeon Jung Kang[1], Woo Suk Cho[1] and Youn Sung Cho[1]*

[1] Biosafety Division, Department of Agricultural Biotechnology, National Institute of Agricultural Sciences, Jeonju, South Korea, [2] Institute for Future Environmental Ecology Co., Ltd., Jeonju, South Korea

Isoflavones are ecophysiologically active secondary metabolites derived from the phenylpropanoid pathway. They were mostly found in leguminous plants, especially in the pea family. Isoflavones play a key role in plant–environment interactions and act as phytoalexins also having an array of health benefits to the humans. According to epidemiological studies, a high intake of isoflavones-rich diets linked to a lower risk of hormone-related cancers, osteoporosis, menopausal symptoms, and cardiovascular diseases. These characteristics lead to the significant advancement in the studies on genetic and metabolic engineering of isoflavones in plants. As a result, a number of structural and regulatory genes involved in isoflavone biosynthesis in plants have been identified and characterized. Subsequently, they were engineered in various crop plants for the increased production of isoflavones. Furthermore, with the advent of high-throughput technologies, the regulation of isoflavone biosynthesis gains attention to increase or decrease the level of isoflavones in the crop plants. In the review, we begin with the role of isoflavones in plants, environment, and its benefits in human health. Besides, the main theme is to discuss the updated research progress in metabolic engineering of isoflavones in other plants species and regulation of production of isoflavones in soybeans.

Keywords: genistein, isoflavones, metabolic engineering, MYB transcription factors, phenylpropanoid pathway, soybean

OPEN ACCESS

Edited by:
Cristina Garcia-Viguera, Consejo Superior de Investigaciones Científicas (CSIC), Spain

Reviewed by:
Xuebin Zhang, Henan University, China
Nik Kovinich, York University, Canada

****Correspondence:***
Soo-In Sohn
sisohn@korea.kr

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

Received: *20 February 2021*
Accepted: *21 April 2021*
Published: *07 June 2021*

Citation:
Sohn SI, Pandian S, Oh YJ, Kang HJ, Cho WS and Cho YS (2021) Metabolic Engineering of Isoflavones: An Updated Overview. Front. Plant Sci. 12:670103. doi: 10.3389/fpls.2021.670103

INTRODUCTION

Isoflavones are a class of flavonoids mostly available in leguminous plants where they play pivotal roles in plant–microbe interactions such as rhizobia–legume symbiosis and defense responses (Sugiyama, 2019). Isoflavones are involved in nodulation process in the leguminous plants by inducing the nodulation genes (Subramanian et al., 2006). They also act as a phytoalexins in plants, i.e., compounds produced by the plants during stress or pathogen attacks (Rípodas et al., 2013). Soybeans produce the maximum amount of isoflavones of all the leguminous crops, and they are the only significant dietary source of these groups of compounds (Kraszewska et al., 2007). Isoflavones have a similar size and chemical structure to the human estrogens that binds to both estrogen α and β receptors. Therefore, they are commonly referred to as "phytoestrogens" (Messina and Wood, 2008). Isoflavones are present in soybean as glycosylated form; however, their biological activity is from their aglycones. When soy foods are consumed, the soy isoflavones are converted to their aglycones by β-glucosidase from enteric bacteria (Tsuchihashi et al., 2008).

In recent years, scientists have been increasingly interested in isoflavones because of their potential health benefits. This can also be seen in the increased number of isoflavone containing nutritional health products in the market. Isoflavones have also been linked to cancer prevention, reduced alcohol intake, prevention of osteoporosis, and cardiovascular diseases (Dixon and Steele, 1999; Pandey et al., 2014). However, this does not mean that consuming isoflavone-rich foods is the ultimate solution to preventing diseases. In certain cases, isoflavones may be needed to consume in impossible quantities to achieve desired health benefits, although there may be some negative effects on human health. Therefore, it is imperative to increase the level of isoflavones in natural environment through metabolic engineering. Hence, understanding the molecular mechanism of isoflavone biosynthesis in various crops is important. This could pave the way to improved production of isoflavone and subsequently helps in the functional food production.

STRUCTURE AND NATURAL ROLE OF ISOFLAVONES IN PLANTS AND ENVIRONMENT

The general structure of isoflavone (**Figure 1**) is made up of a 3-phenylchromen-4-one backbone, with the rings denoted by the letters A, C, and B, beginning from the left (**Figure 1**). The position of phenyl ring in the structure of an isoflavone varies from that of a flavone, which is in position 2 in flavone but in position 3 in isoflavones (**Figure 1**). Although these compounds have structural similarities, they differ in their chemical behavior, and hence, the synthetic approaches for flavones cannot be used with isoflavones. Isoflavones belong to the large isoflavonoid family, which includes the following groups: isoflavones, isoflavans, isoflavanone, isoflavonols, isoflav-3-enes, α-methyldeoxybenzoins, rotenoids, pterocarpans, coumestans, 2-arylbenzofurans, 3-arylcoumarins, and coumaronochromones (Reynaud et al., 2005).

Isoflavones are known to have many effects on plant–microbe interactions, including control of nodulation, having an antifungal activity, and being precursors to phytoalexin (Yu et al., 2000). Phytoalexins act against both prokaryotic and eukaryotic microorganisms with their large spectrum of defense mechanisms (Paxton, 1981). In soybean, both the simple and complex derivatives of isoflavones act as phytoalexins. The main isoflavone phytoestrogens are genistein, daidzein, and glycitein (Křížová et al., 2019) (**Figure 1**). A rapid increase in the isoflavonoid levels in soybeans has been reported upon treatment with various pathogens (Graham, 1991, 1995; Lozovaya et al., 2004; Jeandet et al., 2013). Isoflavones also play a role in plant–environment interactions by mediating the symbiosis between plants and N_2 fixing bacteria. As plants could not use atmospheric N_2, the legumes develop symbiosis with the bacteria and used for its metabolism (Mulligan and Long, 1985). In respect to plant–microbe interaction, the rhizobia attracted by the root exudates move toward the legume roots through positive chemotaxis (Gaworzewska and Carlile, 1982; Caetano-Anolles et al., 1992; Compton et al., 2020). The *Rhizobium* genes are classified into two classes. Genes that determine the synthesis of exopolysaccharides (*exo* genes), lipopolysaccharides (*lps* genes), capsular polysaccharides of K antigens, and β-1,2-glucans (*ndv* genes) belong to the one class of genes involved in the synthesis of bacterial cell surface (Iyer and Rajkumar, 2017). The second class of genes comprises nodulation (*nod*) genes. Isoflavonoids from the plants act as key factor to induce the activation of rhizobial nodulation genes (Philips and Tsai, 1992; Liu and Murray, 2016; Ahmad et al., 2020) and require the participation of the transcriptional-activator protein NodD (**Figure 2**). In the first step, flavonoids excreted by the plant form a complex with the NodD protein, promoting the transcription of bacterial nod genes (Fisher and Long, 1992; Oldroyd et al., 2011; Del Cerro et al., 2019). In the second step, the bacterium produces lipooligosaccharide signals (Nod factors) (Spaink, 1992; Del Cerro et al., 2019) that cause various root responses through structural nod genes (Spaink et al., 1991). The role of isoflavonoids in root nodule formation is extensively studied in soybean by overexpression and RNAi-mediated gene silencing of *IFS* (*isoflavone synthase*) genes (Subramanian et al., 2005). This is not only specific to N_2 fixing bacteria, but it also plays an important role in symbiosis with mutualistic fungi (**Figure 2**). Isoflavones may promote spore germination, hyphal formation and growth, root colonization, and arbuscule formation within the root during the establishment of fungal symbiosis (Abdel-Lateif et al., 2012). The specific characteristics of utilization of atmospheric N_2 for their metabolism signify the legumes as important plant species for the development of soil quality and as an alternative for chemical N_2 fertilizers.

ROLE OF ISOFLAVONES IN HUMAN HEALTH

In the last few decades, the consumption of isoflavone-rich functional foods is highly recommended owing to the potential health protection against some aging-associated diseases such as cardiovascular disease, osteoporosis, hormone-related cancer, and cognitive impairment (Gilbert and Liu, 2013; Vitale et al., 2013; Chi et al., 2016) (**Figure 3**). Even though isoflavones present in a variety of plant-derived products such as cereals, potatoes, vegetables, and fruits, the richest sources in the human diet are soy-derived foods (Kraszewska et al., 2007). For instance, the US Food and Drug Administration announced in 1999 that consuming soy protein (25 g/day) (i.e., soy isoflavone) on a daily basis may decrease the risk of coronary heart disease by reducing the blood cholesterol content. Various clinical studies have revealed that isoflavone favorably lowers the risk of cardiovascular disease because of its estrogenic property (Yan et al., 2017; Nachvak et al., 2019). A systemic review that evaluated the impact of isoflavone diet on cardiovascular disease in 1,307 menopausal (n = 139) and postmenopausal (n = 1,268) women concluded that supplementation of soy isoflavone through diet reduces the cardiovascular risk by lowering the cholesterol and triglyceride plasma concentrations and also oxidative stress (Perna et al., 2016). The clinical study by Tikkanen and Adlercreutz (2000) suggested that

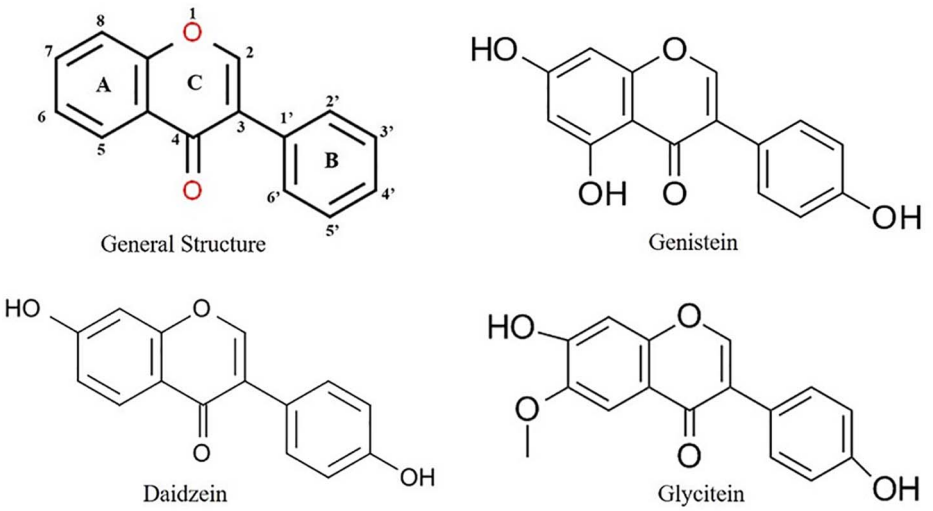

FIGURE 1 | Structure of important isoflavones.

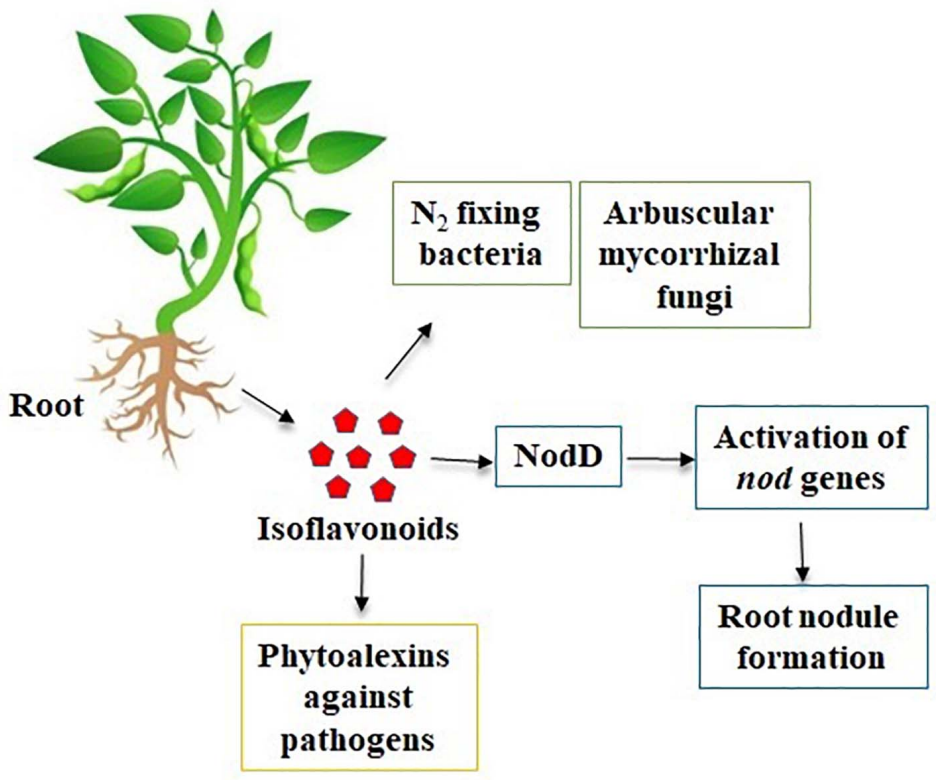

FIGURE 2 | Natural role of isoflavones in plants and environmental interactions.

dietary supplementation of isoflavone reduces low-density lipoprotein cholesterol concentrations and increased high-density lipoprotein concentration. However, the recent studies demonstrated that other constituents of soy, including proteins, fiber, and phospholipids, may play an important role in balancing the cholesterol profile than isoflavones. Some evidence hints that equol, a substance converted from soy isoflavones by the action of the intestinal microflora, seems to be playing a huge role in the reduction of cholesterol content (Nestel et al., 2004; Wu et al., 2007).

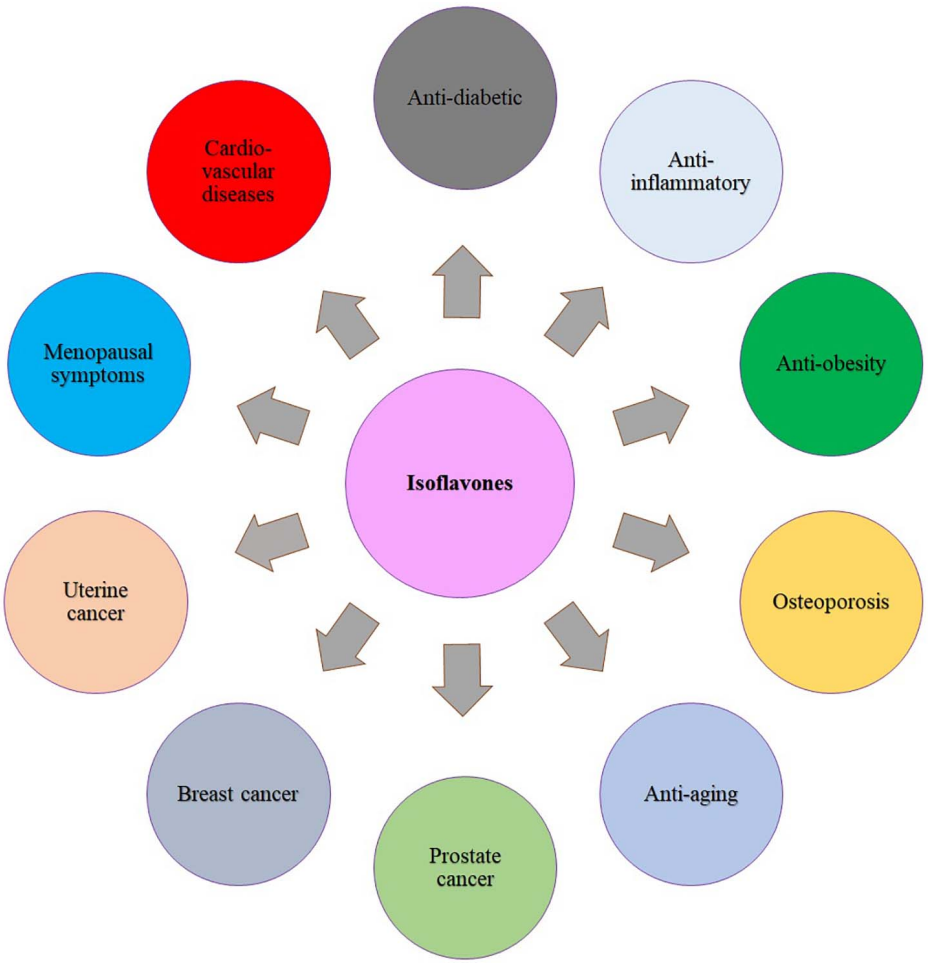

FIGURE 3 | Benefits of isoflavones in human health.

As the ovarian growth hormone (estrogen) has been linked to breast and uterine cancers, isoflavones with estrogenic effects may be used as an effective treatment to thwart breast and uterine cancer (Kumar et al., 2002). However, because of the estrogen-like properties of soy isoflavones, it may stimulate the development of estrogen-sensitive breast tumors in a few cases (Messina and Loprinzi, 2001); hence, understanding the action behind the intake of soy isoflavones and breast cancer reduction is still controversial. In contrast to these conclusions, epidemiologic results suggest that higher soy intake is linked to an approximately one-third lower incidence of breast cancer in Asian women (Wu et al., 2008). Other studies found that Japanese breast cancer patients have better survival rates than Western women, even after the diagnosis (Wu et al., 1996, 2002; Yamamoto et al., 2003). The anti–prostate cancer efficacy of isoflavones was elaborately reviewed by the Mahmoud et al. (2014). Both epidemiological and clinical studies revealed that dietary supplements containing isoflavones could be an effective alternative treatment for various hormonal disorders (Chen et al., 2019; Daily et al., 2019). One of such age-related hormonal diseases is osteoporosis, a bone loss disease that mostly occurs in women who are at menopause time. At the stage of menopause, the low secretion of ovarian hormone, i.e., estrogen, causes an imbalance between resorption and formation of new bone, and subsequent bone loss occurs (Hooper et al., 2010; Taku et al., 2011). The genistein is a well-known isoflavone phytoestrogen that plays an important role in prevention of osteoporosis by acting as an estrogen receptor modulator (**Figure 3**). Data from epidemiological studies revealed that the risk of osteoporosis has reduced in women who consume foods with high isoflavone content (Abdi et al., 2016).

A meta-analysis showed that supplementation of isoflavones greatly increases bone mineral density and decreases the marker of bone resorption, i.e., urinary deoxypyridinoline (Wei et al., 2012; Atcharaporn et al., 2014). They also found that there are numerous factors that could significantly influence the function of isoflavone on bone resorption and formation, such as menopausal status, dose of isoflavone, and intervention duration. Moreover, the menopausal period is associated with cardiovascular diseases due to low production of estrogen. As genistein has antiaging efficacy in various estrogen-dependent aging conditions, it can be used for cosmetic preparations

to improve skin tone and reduce wrinkles and skin dryness (Wang et al., 2010; Geeta et al., 2019). The estrogenic effects of isoflavones, such as genistein and daidzein, were also used to improve the quality of sleep in Japanese adults (Cui et al., 2015). Furthermore, a previous study indicated that administration of soy isoflavones (daidzein, genistein, and glycetin) (20 mg/day) synergistically improved irritable bowel disease by combination with vitamin D in female patients (Jalili et al., 2016). Owing to its importance in natural processes and in human health, a plethora of researchers started to concentrate on the improved productions of isoflavones through metabolic engineering.

ISOFLAVONE BIOSYNTHESIS AND THE ROLE OF BIOSYNTHETIC GENES

Isoflavonoids are limited primarily to the Leguminosae and a few other species. Isoflavones are synthesized via the phenylpropanoid pathway from which plants produce most of the secondary metabolites including lignin, flavone, flavonol, anthocyanin, and tannin, etc. Isoflavones are produced by using intermediate substrates of phenylpropanoid pathway, naringenin, and liquiritigenin, respectively. Naringenin is common in most plants, and other compounds of phenylpropanoid pathway, such as flavones, flavonol, and anthocyanin, are also derived from it (**Figure 4**). Another intermediate substrate of phenylpropanoid pathway, liquiritigenin, is produced by chalcone isomerase (CHI) and chalcone reductase (CHR). CHI is present in most plants, whereas CHR is specific for legumes. Isoflavone reaction comprised two steps. Cytochrome P450–mediated hydroxylation associated with 2,3-aryl migration of the B ring in 2S flavanones forms a 2-hydroxyisoflavanone using IFS (**Figure 4**). Then, this is dehydrated into the isoflavones (genistein and daidzein) through the natural reaction catalyzed by specific dehydratase enzyme (Kochs and Grisebach, 1986; Hakamatsuka et al., 1990, 1998).

Generally, genes on the phenylpropanoid pathway are known to be triggered by environmental stresses (nutrient deficiency, excessive heat, and pathogen attack, etc.) through developmental and tissue-specific regulation (Dixon and Paiva, 1995). Many of the upstream phenylpropanoid pathway enzymes, namely, phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), *p*-coumaroyl-CoA (4CL), CHS, and CHI, have been well characterized. PAL catalyzes the deamination of phenylalanine to produce *trans*-cinnamic acid, which is then transformed to *p*-coumaric acid through an oxidation reaction catalyzed by C4H. After 4CL activates the thio esterification, *p*-coumaroyl-CoA directed into the branched pathway resulted in the production of lignins and flavonoids. C4H belongs to the CYP73 family of the large group of cytochrome P450 monooxygenases (Teutsch et al., 1993). Cytochrome P450 monooxygenases are playing a role in the biosynthesis of a wide range of metabolites in plants (e.g., fatty acids, phenylpropanoids, alkaloids, and terpenoids) and detoxification of herbicides and pesticides (Chapple, 1998). The accumulation of chalcone in plant tissue is rare. CHI catalyzes the stereospecific isomerization of chalcones into corresponding (2S)-flavanones, naringenin, and liquiritigenin (**Figure 4**). Even in the absence of CHI, chalcones can be spontaneously isomerized to form (2S)-flavanones, at a slower rate (Jez et al., 2000). CHR, which is not present in nonlegume species, belongs to the aldo-keto-reductase superfamily and is involved in the synthesis of 6′-deoxychalcone, which is the precursor for daidzein. Among the genes, flavanone-3-hydroxylase (F3H) uses naringenin as a substrate; therefore, it competes with IFS for the formation of isoflavones. The phenylpropanoid pathway genes and their roles are elaborated below.

Phenylalanine Ammonia-Lyase Genes

Phenylalanine ammonia-lyase is a key enzyme in the phenylpropanoid pathway that catalyzes the first step of the pathway (Kervinen et al., 1997; Wang et al., 2007) found in plants (Ritter and Schulz, 2004; Francini et al., 2008), fungi (Hyun et al., 2011), and bacteria (Xiang and Moore, 2005). There is a positive correlation relationship between PAL activity and anthocyanin content in various fruits such as grapes (Hrazdina et al., 1984), strawberries (Given et al., 1988), and apples (Steyn et al., 2004), whereas negative correlation was detected in litchi fruit (Wang et al., 2004). *PAL* genes have been transformed and characterized in different plant species, including tobacco (Nagai et al., 1994), ginkgo (Cheng et al., 2005), Mongolian milkvetch (Liu et al., 2006), banana (Wang et al., 2007), Salvia (Hu et al., 2009), coffee (Lepelley et al., 2012), and so on.

Synthesis of phenylpropanoids is regulated by multiple steps, including the entry of sugars into the shikimic acid pathway, Phe into the general phenylpropanoid pathway, and the activated coenzyme A (CoA) esters into various subbranches of the phenylpropanoid pathway. PAL catalyzes the nonoxidative deamination of L-Phe to produce cinnamic acid, a reaction that is thought to be a central control point for which there is carbon flux into this pathway. *PAL* tends to be a gene family that exists universally in higher plants, and observation of *PAL* isoforms is common. The importance of this diversity is uncertain, but evidence of metabolic channeling within phenylpropanoid metabolism indicates that partitioning of photosynthesis into specific branches of phenylpropanoid metabolism could be involved by labile multienzyme complexes containing specific isoforms of PAL (Hrazdina and Wagner, 1985; Rasmussen and Dixon, 1999). PAL is localized in crucial metabolic position, linking the primary and secondary metabolism. PAL as a rate-limiting enzyme is thought to regulate overall flux into phenylpropanoid metabolism (Bate et al., 1994). *In vitro* PAL activity shows that feedback inhibitory regulation of its own product, *trans*-cinnamate and *trans*-cinnamic acid, was proposed to alter the transcription of *PAL* genes *in vivo* (Jorrin and Dixon, 1990; Mavandad et al., 1990; Appert et al., 1994).

Chalcone Synthase Genes

The first flavonoid biosynthetic gene *CHS* was isolated from parsley (Kreuzaler et al., 1983; Schulze-Lefert et al., 1989). CHS forms a tetraketide intermediate that is cyclized into 4,2′,4′,6′-tetrahydroxychalcone (chalcone) using *p*-coumaroyl-CoA three malonyl-CoA extender molecules. *CHS* belongs to the multigene family, for example, 12 in petunia (Koes et al., 1989) and

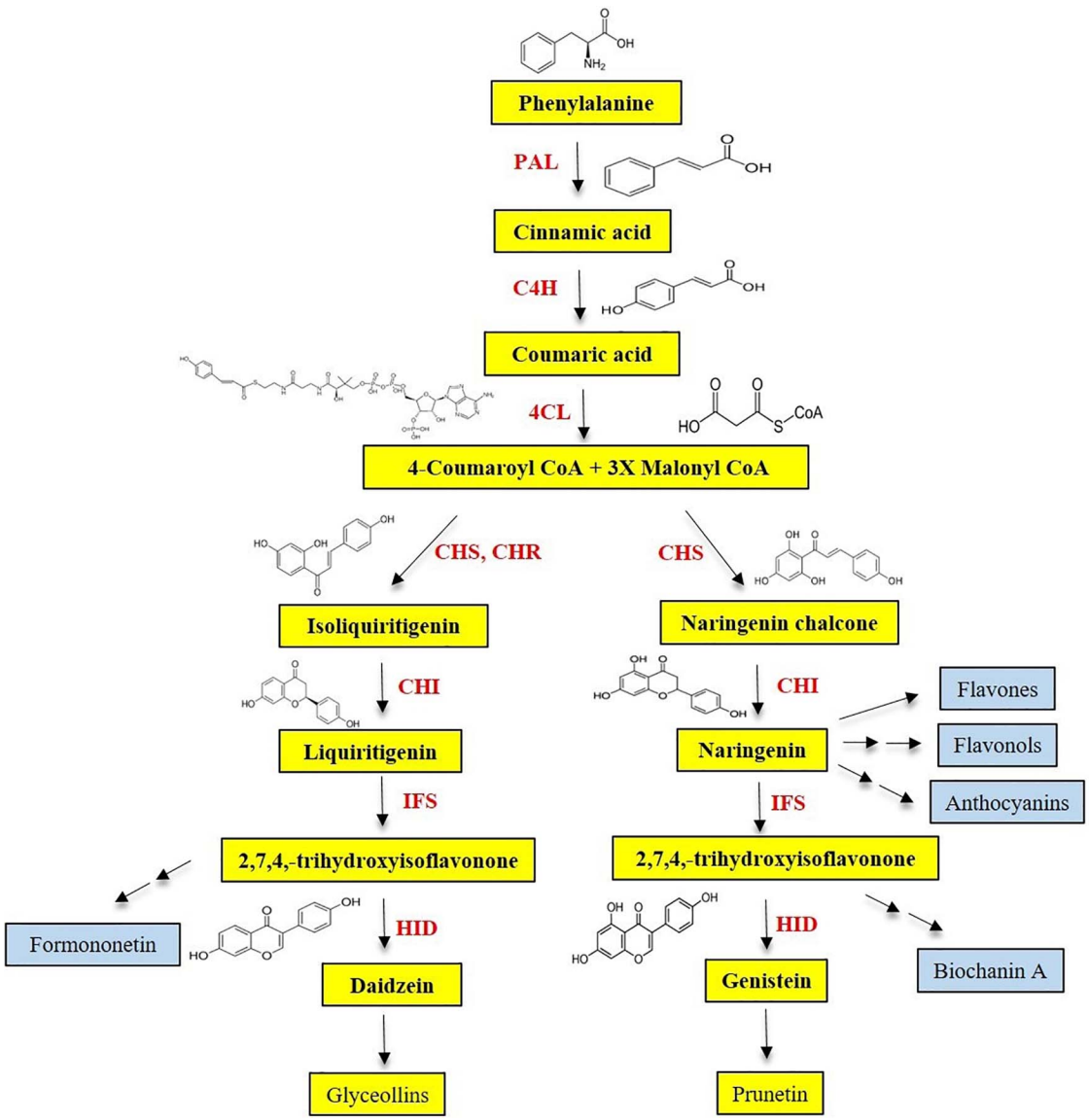

FIGURE 4 | Partial phenylpropanoid pathway for the isoflavone biosynthesis.

8 in soybean (Matsumura et al., 2005). Few of the species has single-copy genes also such as *Arabidopsis* (Feinbaum and Ausubel, 1988) and *Petroselinum crispum* (Weisshaar et al., 1991). Alfalfa *CHS2* with three-dimensional structure reveals the active site that catalyzes chalcone formation. The catalytic center of CHS is formed by four residues (Cys164, His303, Asn336, and Phe215), which are exclusively conserved in other CHS-like enzymes, such as 2-pyrone synthase, bibenzyl synthase, stilbene synthase, acridone synthase, and the *rppA* CHS-like protein. Structural and functional studies revealed that Cys164 acts as the nucleophilic thiolate in the loading reaction and as the covalent thioester-anchor for the acyl-enzyme chain during the elongation reactions. Furthermore, in the elongation reaction, His303 and Asn336 catalyze the decarboxylation of malonyl-CoA and stabilize the intermediate state in the condensation phases of polyketide formation. Substrates and reaction intermediates at the active site are oriented by Phe215. CHS supplies the important substrate chalcone for the isoflavone biosynthesis, and its expression level plays an important role in isoflavonoid accumulation in plants (Tuteja et al., 2004). Downregulation of *CHS* genes plants may lead to the complete absence of flavones and isoflavone biosynthesis. However, these may lead to the several functions of the plants such as male sterility (Napoli et al., 1999).

Chalcone Reductase Genes

Other than CHS, leguminous plants also have CHR, which catalyze the intermediate of the multistep CHS reaction, producing chalcone and 4,2′,4;-trihydroxychalcone (deoxychalcone) as a result of their combined catalytic activity.

In response to herbivore or pathogen attacks, CHR sits at a key branch of the biosynthetic pathway and synthesizes a set of deoxychalcone-derived phytoalexins such as isoflavonoids, coumestans, pterocarpans, and isoflavans. CHR occurrence was first discovered in *Glycyrrhiza echinata* crude extracts (licorice) (Ayabe et al., 1988). Subsequently, CHR was purified, and the CHR cDNA (*GmCHR1*) from *Glycine max* (soybean) was cloned (Welle and Grisebach, 1988). Other homologs of *GmCHR* (*GmCHR2*, *GmCHR3*, *GmCHR4*, and *GmCHR2A*) have been identified and tested for their significance (Subramanian et al., 2006; Sepiol et al., 2017). Graham et al. (2007) have identified four *GmCHR* genes through mining of soybean ESTs. Until now, CHR-like enzymes have been discovered in a wide range of leguminous and some nonleguminous plant species, such as *Medicago sativa* (alfalfa) (Ballance and Dixon, 1995), *Glycyrrhiza glabra* (licorice), and *Fragaria ananassa* (strawberry) (Manning, 1998), *Pueraria lobata* (Kudzu vine), *G. echinata* (licorice) (Akashi et al., 1999), *Lotus japonicas* (Shimada et al., 2006), *Pueraria montana* (He et al., 2011), and *G. glabra* (Hayashi et al., 2013). As a result of sequence analysis, CHR was found unrelated to known ketoreductase domains of fatty acid synthases (Price et al., 2001; Cohen-Gonsaud et al., 2002) and β-polyketide synthases (Staunton and Weissman, 2001). However, aldo-ketoreductase superfamily enzymes from primary metabolism, such as aldose reductase and 3α-hydroxy steroid dehydrogenase, were found to have sequence similarities with CHR (Mameda et al., 2018).

Chalcone Isomerase Genes

Most of the plants do not accumulate the chalcones. During the early stages of flavonoids biosynthesis, CHI catalyzes the intramolecular cyclization of chalcone and 6′-deoxychalcone both generated by the upstream enzyme chalcone synthase, into (2S)-naringenin and (2S)-5-deoxyflavanone, respectively (Holton and Cornish, 1995; Jez and Noel, 2000). CHI ensures the formation of biologically active (2S)-flavanones as chalcones naturally cyclize in solution to yield an enantiomeric mixture of flavanones. For instance, the metabolic precursor of anthocyanin pigments such as (2S)-naringenin and mutations in the CHI genes have been linked to variations in floral pigmentation (Burdick, 1958). Recently, the introduction of the petunia CHI gene was recently introduced into the tomato plants, resulting in fruits with higher flavanol content (Kang et al., 2014).

In the legume *L. japonicas*, two types of CHIs coexist with distinctive phylogenic lineages (Shimada et al., 2003). The type I CHIs are commonly present in the plant kingdom, which converts naringenin chalcone to naringenin. On the other hand, the type II CHI tends to be legume-specific and have additional catalytic activity, allowing them to convert 4,2′,4-tryhydroxychalcone (isoliquiritigenin) into (2S)-7,4′-dihydroxyflavanone (liquiritigenin). A type II CHI isolated from alfalfa (*M. sativa*) has been extensively studied structurally and mechanistically (Jez et al., 2000, 2002; Hur et al., 2004; Sun et al., 2019). According to structure–function analyses, the formation of a hydrogen bond network between the active site of CHI and its substrates appears to be important for the enzyme's catalytic activity (Jez et al., 2002; Hur et al., 2004). Ralston et al. (2005) reported that there are four types of CHI based on the phylogenetic relationships, but types III and IV do not have chalcone cyclization activity like types I and II. Until now, CHI genes have been transformed and identified in several plants including *Pterolophia hybrida* (Van et al., 1988), *M. sativa* (Jez et al., 2000), *L. japonicas* (Shimada et al., 2003), *Oryza sativa* (Druka et al., 2003), *G. max* (Ralston et al., 2005), *Ginkgo biloba* (Cheng et al., 2011), *Ipomoea batatas* (Guo et al., 2015), tomato (Muir et al., 2001; Lim and Li, 2016), and *Chamaemelum nobile* (Wang et al., 2018), etc.

Isoflavone Synthases Gene

Isoflavone synthases belong to the cytochrome P450 family, and they are extremely labile in the cells. In the phenylpropanoid pathway, IFS plays an important role that redirects the intermediates of flavonoid pathway to the isoflavonoid pathway. It catalyzes the committed step of isoflavonoid biosynthesis by producing the 2-hydroxyisoflavone from the flavone intermediates such as naringenin and liquiritigenin (Liu et al., 2002). The resulted 2-hydroxyisoflavone is dehydrated by the isoflavone dehydratase (HID) to produce basic isoflavone compounds such as genistein and daidzein. The IFS gene has been cloned and characterized in various plants, including *Arabidopsis*, tobacco, rice, and tomato, etc. (**Table 1**). Cloning of IFS encoding genes to various plant species paved way for the genetic engineering for the synthesis of isoflavone compounds in the plants which naturally do not produce isoflavones (**Table 1**).

Initially, it is thought that IFS is specific to legumes and encoding genes only in leguminous plants, but researchers found that *IFS* genes in other plants such as *Beta vulgaris* (Jung et al., 2000), *Psoralea corilyfolia* (Misra et al., 2010) suggested that *IFS* can be found in other crops and medicinal plants. There are two isoflavone synthases, *IFS1* and *IFS2*, isolated from soybean which is a having high level of genetic similarity, but they have different efficiency in converting the flavones in to the isoflavones (Jung et al., 2000). Expression analysis revealed that both the *IFS1* and *IFS2* had different level of expression in different environmental conditions with *IFS1* mainly found in root and seed coat, whereas *IFS2* can be found in embryos and seed pods (Dhaubhadel et al., 2003; Subramanian et al., 2004). *In vitro* system using yeast, when the naringenin and liquiritigen were used as substrates, *IFS1* had a twofold higher activity compared to *IFS2*. During the soybean embryo development, the expression of *IFS2* gene was significantly increased for 70 days after the pollination (Dhaubhadel et al., 2007).

Isoflavone Dehydratase Genes

Isoflavone dehydratase belongs to the carboxylesterase gene family, which is involved in the final step of isoflavonoid biosynthesis that produces the genistein and daidzein from the isoflavone skeleton. HID was first identified and purified from the *P. lobata* by rapid enzyme assay method (Hakamatsuka et al., 1998), but the amino acid sequence is not available. Compared to other genes involved in the phenylpropanoid pathway, such as *CHS*, *CHI*, and *IFS*, this gene (*HID*) is less characterized. The problem in characterizing the HID is the instability of its substrate 2-hydroxyisoflavanones. Later *HID* genes were

TABLE 1 | Metabolic engineering of isoflavone biosynthetic genes in various plants.

S. no.	Plants	Genes	Promoter	Purpose	References
1	*Arabidopsis thaliana*	*GmIFS*	*CaMV 35S*	Accumulation of genistein	Jung et al., 2000
2	Tobacco (*Nicotiana tabacum*) Maize BMS cell lines	*GmIFS*	*CaMV 35S*	Not identified	Yu et al., 2000
3	Maize BMS cell lines	*GmIFS and Maize CRC*	*CaMV 35S*	Accumulation of genistein	Yu et al., 2000
4	*Arabidopsis thaliana*	*GmIFS*, *MtCHI*	*CaMV 35S*	Accumulation of genistein	Liu et al., 2002
5	Alfalfa (*Medicago sativa*)	*MtIFS1*	*CaMV 35S*	Accumulation of genistein, glycosides of biochanin A, formononetin, and daidzein	Deavours and Dixon, 2005
6	Rice (*Oryza sativa* L.)	*GmIFS*	*CaMV 35S*	Accumulation of genistein	Sreevidya et al., 2006
7	Rice (*O. sativa* L.)	*maize C1-R-S*	*prolamin*	Dihydroquercetin, dihydroisorhametin, 3′-*O*-methyl quercetin	Shin et al., 2006
8	Tobacco (*N. tabacum*)	*IFS/CHI* fusion protein	*1. CaMV 35S*	Increased accumulation of genistein and genistein glycosides	Tian and Dixon, 2006
9	Tobacco (*N. tabacum*)	*GmIFS*	*CaMV 35S*	Accumulation of genistein	Liu et al., 2007
10	Lettuce (*Lactuca sativa*)	*GmIFS*	*CaMV 35S*	Accumulation of genistein	Liu et al., 2007
11	Petunia (*Petunia hybrida Vilm.*)	*GmIFS*	*CaMV 35S*	Accumulation of genistein	Liu et al., 2007
12	Tomato (*Solanum lycopersicum*)	*GmIFS2*	*CaMV 35S*	Accumulation of genistein	Shih et al., 2008
13	Tobacco	*PcIFS*	*CaMV 35S*	Accumulation of genistein and daidzein	Misra et al., 2010
14	Rapeseed (*Brassica napus*)	*GmIFS2*	*CaMV 35S*	Accumulation of Genistein	Li et al., 2011
15	Rice (*O. sativa* L.)	*PAL*, *CHS*, and *GmIFS1*	*Glu-B1* and *oleosin*	Accumulation of naringenin, kaempferol, genistein and apigenin	Ogo et al., 2013
16	Rice (*O. sativa* L.)	*Maize Lc*	*Glutelin Gt1*	Not identified	Song et al., 2013
17	Tobacco (*N. tabacum*)	*AtMYB12* and *GmIFS1*	*CaMV 35S*	Enhanced biosynthesis of isoflavones and flavonols	Pandey et al., 2014
18	Rice (*O. sativa* L.)	*SpdIFS1* and *SpdIFS2*	*Glb-1*	Accumulation of genistein	Sohn et al., 2014
19	Barrel medic (*Medicago truncatula*)	*GmIFS1*, *GmCHS7*, and *GmCHI1A*	*CaMV 35S*	Increased accumulation of genistein, daidzein, biochanin A, and formononetin	Gou et al., 2016
20	Onion (*Allium cepa* L.)	*GmIFS*	*CaMV 35S*	Increased accumulation of genistein	Malla et al., 2021

identified and characterized from two plants, including licorice and soybean (Akashi et al., 2005). Site-specific mutagenesis in the *GmHID* revealed that the oxyanion and catalytic triad are important for the dehydratase and esterase activity of these genes. Shimamura et al. (2007) have overexpressed both *IFS* and *HID* genes in lotus, to understand the functional role of the *HID* in the isoflavone biosynthesis. The *GmHID* introduced was produced and increased the amount of genistein and daidzein in the lotus as the *GmHID* has broad specificity to both the 4′-methoxy and 4′-hydroxy substrates. The HID enzyme activity is an important limiting factor in the isoflavone biosynthesis and its level in the legumes (Du et al., 2010).

REGULATION OF ISOFLAVONE PRODUCTION THROUGH GENETIC ENGINEERING IN SOYBEAN

As the consumption of isoflavones is associated with a variety of health benefits to the humans, several attempts were made to develop soybeans that accumulate much higher levels of isoflavones than in wild-type seed. Generally, the isoflavone contents in *G. max* (soybean) were improved through metabolic engineering of the complex phenylpropanoid biosynthetic pathway. In a first study, the phenylpropanoid pathway genes were activated by expressing the maize *CRC* fusion gene, resulting in a decrease in genistein and the increase in daidzein levels with a marginal increase in total isoflavone levels (Grotewold et al., 1998). Cosuppression of flavanone 3-hydroxylase to block the anthocyanin branch of the pathway, in conjunction with *CRC* expression, resulted in higher levels of isoflavones (Yu et al., 2003). The use of transcription factor–driven gene activation combined with the suppression of a competing pathway resulted in increased isoflavone accumulation in soybean seeds. These high isoflavone soybeans could be used to make soy foods with greater health benefits to consumers.

The regulation of isoflavone biosynthesis is carried out by multiple genes and complex metabolic pathways. In order to understand the specific functions/regulation of the genes involved in the phenylpropanoid pathway for isoflavone biosynthesis, the overexpression or gene-silencing methods of specific genes can be carried out. Hence, Subramanian et al. (2005) have studied the RNAi-mediated gene silencing of *GmIFS* genes in the soybean. The study resulted in the reduced level of isoflavone accumulation in the gene silenced plants compared to control (**Table 2**). In another study, overexpression of *CHS6*

TABLE 2 | Studies on the regulation of isoflavone biosynthesis in soybean.

S. no.	Genes	Regulation	Results	References
1	*GmIFS*	Positive regulation	RNAi-mediated gene silencing reduced the level of isoflavone content in the soybean	Subramanian et al., 2005
2	*CHS6* and *IFS2*	Negative regulation	Transformed plants found reduced accumulation of isoflavones and increased accumulation of phenolic compounds	Lozovaya et al., 2007
3	*GmMYB176*	Positive regulation	No significant improvement in the overexpression but in transient expression	Yi et al., 2010
4	*GmFNSII*	Negative regulation	RNAi-mediated silencing increased the genistein level significantly	Jiang et al., 2010
5	*GmMYB39*	Negative regulation	Overexpression reduced the transcript levels of *PAL*, *C4H*, *CHI*, and *CHS* genes	Liu et al., 2013
6	*F3H* and *GmFNSII*	Negative regulation	RNAi-mediated silencing increased the daidzein and genistein level significantly	Jiang et al., 2014
7	*GmMYB100*	Negative regulation	Overexpression reduced the transcript level, whereas RNAi-mediated gene silencing increased the transcript level of isoflavonoids genes	Yan et al., 2015
8	*GmMYB29*	Positive regulation	Overexpression increased the isoflavone level in 1.6- to 3.3-fold, whereas RNAi gene silencing decreased the isoflavone content in twofold	Chu et al., 2017
9	*GmIMaT1* and *GmIMaT3*	Positive regulation	Overexpression significantly increased the level of isoflavone aglycones, glucosides. and malonylates, whereas knockdown the genes reduced the contents	Ahmad et al., 2017
10	*GmMYBJ3*	Positive regulation	Activates CHS8 and CHI1A and increases the accumulation of isoflavones	Zhao et al., 2017
11	*Gma-miRNA12*, *Gma-miRNA24*, *Gma-miRNA26*, *Gma-miRNA28*, and *Gma-miRNA29*	Positive regulation	Corresponding genes for the *Gma-miRNA26*, *Gma-miRNA28* have increased the isoflavone content in soybean	Gupta et al., 2017
12	*GmCHI1A*	Positive regulation	Overexpression of *CHI1A* leads to the increased accumulation of daidzein in transgenic plants	Zhou et al., 2018
13	*GmMYB133*	Positive regulation	Increased the transcript level of *GmIFS2* and *GmCHS8* also increased the total isoflavone contents in the hairy roots	Bian et al., 2018
14	*GmMYB102*, *GmMYB280*, and *GmMYB502*	Positive regulation	Two- to fourfold increased accumulation of isoflavones was found	Sarkar et al., 2019
15	*GmMYB176*	Positive regulation	Activating *CHS8* gene and identified 25 metabolic genes and six metabolites upon overexpression	Anguraj Vadivel et al., 2019
16	*GmF3H1*, *GmF3H2*, *and GmFNSII-1*	Positive regulation	CRISPR/Cas9-mediated targeted gene editing leads to the increased accumulations of isoflavones	Zhang et al., 2020
17	*GmMYB29A2*	Positive regulation	Overexpression increased glyceollin accumulation and RNAi gene silencing decreased the accumulation. Accumulation of glyceollin leads to the resistance to *Phytophthora sojae*	Jahan et al., 2020
18	*GmCHI1A*	Positive regulation	T_2 transgenic plants accumulated the higher level of genistein and daidzein than the control plants	Nguyen et al., 2020
19	*GmMYB176*	Positive regulation	Improved the accumulation of isoflavone biosynthetic genes based on the diurnal regulation system	Matsuda et al., 2020
20	*GmMYB176* and *GmbZIP5*	Positive regulation	Enhance accumulation of multiple isoflavonoid phytoalexins, namely, glyceollin, isowighteone, and *O*-methyl hydroxyl isoflavone in soybean hairy roots	Vadivel et al., 2021

and *IFS2* genes resulted in the reduced level of isoflavones and increased level of phenolic compounds (Lozovaya et al., 2007). Jiang et al. (2010; 2014) explored the negative regulation of F3H and *GmFNSII* genes by the RNAi-mediated gene silencing studies. Previously, the gene silencing of *GmFNSII* resulted in the increased accumulation of genistein in soybean hairy roots (Jiang et al., 2010). Later, RNAi-mediated silencing of both genes (*F3H* and *GmFNSII*) was found to increase accumulation of isoflavones such as daidzein and genistein (Jiang et al., 2014) (**Table 2**).

Overexpression of *GmIMaT1*, *GmIMaT3* genes significantly increased the different forms of isoflavones such as aglycones, glucosides, and malonylates, but in knockdown of the genes, the isoflavone levels were reduced drastically (Ahmad et al., 2017) (**Table 2**). Zhang et al. (2020) have used a new methodology for the regulation of isoflavone biosynthetic genes by introducing CRISPR/Cas9-mediated targeting mutation of multiple genes (*GmF3H1*, *GmF3H2*, and *GmFNSII-1*) in soybean. The T3 generation of triple gene mutants produced increased levels (twice the amount) of genistein compared to controls. The

increased isoflavone content also leads to the enhanced resistance to soybean mosaic virus infections (Zhang et al., 2020). Recently, Nguyen et al. (2020) found that overexpression of *GmCHI1A* was found to increase higher levels of both genistein and daidzein in T2 generations, but previously, Zhou et al. (2018) reported that it can increase the daidzein content alone (**Table 2**).

Not only changing the specific genes in the phenylpropanoid pathway can notably increase/decrease isoflavone content but also the transcription factors. Therefore, the identification and application of transcription factors, particularly for the isoflavone pathway, may significantly resolve this problem (Yu and McGonigle, 2005; Chu et al., 2017). The transcription factors of the Myeloblastosis (MYB) family play crucial roles in the regulation of isoflavone biosynthesis. Yi et al. (2010) demonstrated through the functional genomic approach that *GmMYB176* (R1 MYB protein) regulates *CHS8* expression and affects the synthesis of isoflavonoids in soybean. In this study, cotransfection analysis with *Arabidopsis* leaf protoplast resulted that *GmMYB176* transactivate the *CHS8* promoter with maximum activity. As a result of transient expression in soybean embryo protoplast, after 48 h, the endogenous transcript levels were increased up to 149-fold. Subcellular localization assay indicates that *GmMYB176* is a nuclear protein. RNAi-mediated gene silencing of *GmMYB176* in hairy roots resulted in reduced levels of isoflavonoids, but overexpression of *GmMYB176* did not significantly increase the levels of *CHS8* transcript (Yi et al., 2010) (**Table 2**). Liu et al. (2013) reported that soybean MYB transcription factor *GmMYB39* was potentially regulating the isoflavone biosynthesis. *GmMYB39* that contained N-terminal R2R3 repeats corresponds to DNA-binding domains of plant MYB-type proteins, which were highly conserved among R2R3-MYB proteins. Quantitative reverse transcriptase–polymerase chain reaction results revealed that overexpression of *GmMYB39* was found to be varied in different parts of the plants. Interestingly, the higher level of transcripts was found in flowers, and lower level in the pods. *GmMYB39* overexpression in hairy roots resulted in drastic reduction of the transcript levels of *PAL*, *C4H*, *CHS*, *4CL*, and *CHR*. However, the transcript level of IFS was slightly increased, whereas there is no change in CHI expression between overexpressed hairy roots and control roots. Overall, this report suggests that *GmMYB39* is involved in the inhibition role in regulation of isoflavone biosynthesis in soybean (**Table 2**).

Yan et al. (2015) found that R2R3-MYB transcription factor *GmMYB100* is involved in soybean isoflavone biosynthesis. Generally, *GmMYB100* is expressed in flowers, leaves, and immature embryo, but its level will decrease after pod ripening. The subcellular localization study found the nuclear localization of *GmMYB100*. Initially, yeast functional assay revealed the transactivation ability of *GmMYB100*, but the bioinformatics analyses suggested its negative role in flavonoid biosynthesis. Finally, the overexpression of *GmMYB100* reduced transcript levels of transgenic hairy roots and *Arabidopsis* and reduced flavonoid and flavonol productions, respectively, whereas the RNAi-mediated silencing resulted in the higher level of transcripts of six flavonoid related genes and accumulated higher level of flavonoids in transgenic hairy roots (**Table 2**). Similarly, Genome-Wide Association Study for the identification of SNPs related to isoflavone concentration in soybean found that another R2R3-MYB transcription factor *GmMYB29* is significantly involved in regulation of isoflavone biosynthesis (Chu et al., 2017). The subcellular localization analysis found that *GmMYB29* was located in the nucleus. It is found to be activated *IFS2* and *CHS8* gene promoters by the transient gene expression assays. Furthermore, *GmMYB29* overexpression and RNAi-mediated silencing in soybean hairy roots resulted in 1.6- to 3.3-fold increase and isoflavone contents in twofold decrease (Chu et al., 2017). Zhao et al. (2017) explored the role of *GmMYBJ3* in regulation of the isoflavone biosynthetic pathway. This result suggests that *GmMYBJ3* can activate the CHS8 and CHI1A genes; therefore, increased accumulation of isoflavones has been witnessed (Zhao et al., 2017).

Gupta et al. (2017) have identified miRNAs in the regulation of isoflavone biosynthesis in two contrast genotypes of soybean. *In silico* analysis identified 31 new miRNAs along with the 245 putative target genes from the seed-specific ESTs. Based on the Kyoto Encyclopedia of Genes and Genomes pathway analyses, five genes (*Gma-miRNA12*, *Gma-miRNA24*, *Gma-miRNA26*, *Gma-miRNA28*, and *Gma-miRNA29*) were found to be involved in isoflavone biosynthesis, among which *Gma-miRNA26* and *Gma-miRNA28* and their corresponding genes (*Glyma.10G197900* and *Glyma.09G127200*) exhibited their direct relationship with the isoflavone content of the soybean (**Table 2**). Bian et al. (2018) identified that *GmMYB133* (CCA1-like MYB) positively regulates isoflavone biosynthesis in soybean. Overexpression of *GmMYB133* has led to the expression of two important isoflavonoid biosynthetic genes such as *GmIFS2* and *GmCHS8* and increased total isoflavonoid contents in the hairy roots. Furthermore, the protein–protein interaction results revealed that it can form heterodimers with another isoflavone regulator *GmMYB176* and homodimers with another *GmMYB133* (Bian et al., 2018).

Sarkar et al. (2019) found that *GmMYB102*, *GmMYB280*, and *GmMYB502* were the potential transcription factors that can activate the promoters of the CHS gene (*GmCHS8*) and the IFS genes (*GmIFS1* and *GmIFS2*) in the isoflavone biosynthetic pathway by hairy root transformation assay. They have assessed the functional regulatory role of these genes by hairy root transformation assay, resulting in increased accumulation of isoflavones (two- to fourfold) in the three MYB overexpressing lines compared to vector control (**Table 2**). Anguraj Vadivel et al. (2019) have identified that *GmMYB176* (R1 MYB transcription factor) activates the CHS8 gene and regulates the isoflavonoid biosynthesis in soybean. They have identified 25 metabolic genes and six metabolites by the targeting approach that is differentially regulated during overexpression and silencing of *GmMYB176* in soybean hairy roots (**Table 2**).

Jahan et al. (2020) have identified that *GmMYB29A2* transcription factor positively regulates glyceollin biosynthesis in soybean. Glyceollins are pathogen-inducible defensive metabolites (phytoalexins) that play important roles in pathogen defense (Jahan et al., 2019). Overexpression of *GmMYB29A2* increased the expression of *GmNAC42*-1, *GmMYB29A1*, and glyceollin biosynthesis genes and metabolites, whereas

RNAi-mediated gene silencing had opposite effects. Previously, Jahan et al. (2019) have studied the role of *GmNAC42*-1 in activation of glyceollin biosynthesis by expression analysis. In this study, they confirmed the positive regulation of *GmMYB29A2* that leads to the increased conversion of isoflavonoids into the glyceollin and thus develops resistance against *Phytophthora sojae* (Jahan et al., 2020). Matsuda et al. (2020) studied the diurnal metabolic regulation of isoflavones and soy saponins in soybean roots. The transcriptome and metabolite analysis of soybean plants at 6-h intervals for 48 h in a 12-h light–12-h dark condition. In the root tissues, isoflavone and soy saponin biosynthetic genes showed opposite patterns; the former are highly expressed in the day, whereas the latter are strongly triggered at night. *GmMYB176*, which encodes an isoflavone biosynthesis transcription factor, was upregulated from ZT0 (6:00 A.M.) to ZT6 (12:00 A.M.), accompanied by the stimulation of isoflavone biosynthetic genes at ZT6 (Matsuda et al., 2020). Recently, Vadivel et al. (2021) have found that the RNAi silencing of transcription factor *GmbZIP5* reduced the isoflavone accumulation in hairy roots. Furthermore, the co-overexpression of *GmMYB176* and *GmbZIP5* enhanced the accumulation of multiple isoflavonoids such as glyceollin, malonyl glycitin, isowighteone, and *O*-methyl hydroxyl isoflavone in soybean hairy roots. An ample of studies in the regulation of isoflavone biosynthetic genes provided the role of key players in soybean isoflavone biosynthesis, which could be useful for the development of soybean with desired level of isoflavones.

METABOLIC ENGINEERING OF ISOFLAVONE IN NONSOYBEAN CROPS

Epidemiologic studies show that a high intake of soybean-derived foods is linked to a low incidence of hormone-related cancers, menopausal symptoms, osteoporosis, menopausal symptoms, and cardiovascular diseases. Furthermore, metabolic engineering of isoflavonoids in common nonlegume vegetables, grains, and fruits to increase dietary intake of these compounds has piqued researchers' interest (Dixon and Sumner, 2003; McCue and Shetty, 2004; Deavours and Dixon, 2005).

Initially, in a monocot cell system, introduced expression of a transcription factor controlling anthocyanin pathway genes was successful in the production of genistein in the presence of the IFS gene (Yu et al., 2000). The genistein produced in tobacco, *Arabidopsis*, and maize cells is present in conjugated forms, indicating that endogenous enzymes were capable of recognizing genistein as a substrate. Introducing foreign IFS gene in *Arabidopsis* seedlings (Jung et al., 2000), tobacco petals, and maize cells in which the phenylpropanoid pathway was activated by C1 and R transcription factors resulted in successful accumulation of genistein (Yu et al., 2000), demonstrating that heterogeneous IFS can use flavanone intermediates as substrates. However, in these cases, the genistein levels were two- to threefold lesser than in soybean seeds (Dixon and Ferreira, 2002). The higher genistein content in transgenic tobacco petals compared to leaves was thought to be due to a more active phenylpropanoid pathway leading to anthocyanin biosynthesis in petals, which also increased the level of intermediates available to IFS (Yu et al., 2000). The low transgene expression is not the cause of low genistein production in the tobacco leaves. The IFS mRNA was actually detected at a higher level in the leaves than in the flowers of tobacco transformants, and IFS protein and enzyme activity were also confirmed in the leaves. Thus, the poor genistein synthesis is most likely due to a lack of naringenin substrate. As IFS can fight for naringenin in tobacco flowers and *Arabidopsis* leaves, channeling might be the cause of unavailability.

Conversely, Liu et al. (2002) produced transgenic *Arabidopsis* with significantly increased genistein accumulation (31–169 nmol/g Fresh weight (FW)), but soybean IFS was introduced into *Arabidopsis tt6/tt3* mutant in which the expression of F3H and another flavonol/anthocyanin enzyme dihydroflavonol reductase was decreased (Liu et al., 2002). As F3H and IFS both use naringenin as a substrate, the researchers hypothesized that the competition for intermediate availability between IFS and other enzymes was a limiting factor for a genistein biosynthesis in genetically modified plants. Furthermore, blocking a competing branch pathway may be a useful method to promote genistein biosynthesis (Liu et al., 2007). UV-B irradiation of *Arabidopsis* transformants expressing the soybean IFS gene resulted in a 2.5-fold increase in genistein accumulation. Increased levels of UV absorption and anthocyanins in IFS-transformed plants depict the higher activity of phenylpropanoid pathway (Yu et al., 2000; Liu et al., 2002). Deavours and Dixon (2005) have improved the genistein production for up to 50 nmol/g FW by constitutively expressing *MtIFS1* in the alfalfa plants. Genistein levels of this study are much higher than previous studies (Jung et al., 2000; Yu et al., 2000; Liu et al., 2002). Even though *MtIFS1* expresses all the tissues in plants, genistein accumulation was specific to the leaves. Apart from the genistein, the plants also accumulated other isoflavones such as daidzein and formononetin in response to the UV-B treatments.

Tian and Dixon (2006) have improved the isoflavone metabolism by introducing the CHI, CHS, and CHI/CHS fusion protein into the tobacco plants. High-performance liquid chromatography (HPLC) analysis results showed the accumulation of genistein in both IFS- and IFS/CHI-transformed plants, but not in vector control and CHI plants. In transgenic tobacco, lettuce, and petunia, both overexpression and antisense suppression were used to control the expression of multiple genes essential enzymes in the flavonoids/isoflavonoids pathway (Liu et al., 2007) (**Table 1**). The introduction of soybean IFS (*GmIFS*) into these plants, which lack this leguminous enzyme and therefore do not produce isoflavonoids naturally, resulted in genistein biosynthesis in tobacco petals, petunia leaves and petals, and leaves of lettuce. In tobacco, when antisense suppression of *F3H* and overexpression of *GmIFS* were done simultaneously, the yield of genistein increased prominently. In addition, overexpression of PAL also led to an enhanced genistein production in tobacco petals and lettuce leaves in the presence of IFS than in the plants that overexpressed only IFS (Liu et al., 2007). Similarly, Misra et al. (2010) transformed the IFS gene from the medicinal plant *P. corilyfolia* (*PcIFS*) into the tobacco.

Overexpression of *PcIFS* resulted in the higher accumulation of isoflavones in the tobacco petals (**Table 1**).

As rice is an important crop and staple food for half of the world population, recently it caught extensive attention as bioreactors for the production of therapeutic compounds and proteins. Although production of isoflavones in rice plants was performed by several researchers, Shin et al. (2006) for the first time in rice plants have transformed *maize C1* and *R-S* regulatory genes for the production of isoflavones. The expression of transgenes is restricted to endosperm under the control of a rice prolamin gene promoter. The transformed plants showed several phenotypic variations such as changes in pericarp color, chalky endosperm, and opaque kernel. HPLC analysis showed that several types of isoflavones accumulated in the transgenic kernels. Among which taxifolin, 3′-*O*-methyl taxifolin and 3′-*O*-methyl quercetin were identified through liquid chromatography (LC)–tandem mass spectrometry (MS/MS) analysis (**Table 1**). Furthermore, they have confirmed the accumulation of isoflavones in the layers of endosperm by the florescence labeling. At the same time, Sreevidya et al. (2006) have introduced the *GmIFS* gene into the rice plants for the production of isoflavones under the control of *CaMV35S* promoter. The genistein accumulated as a glycosylated form in the rice plants. Rhizobia study for the nod gene expression confirms the accumulation of isoflavones inducing the nod genes in varied degrees.

Introduction of five biosynthetic genes (*OsPAL*, *OsC4H*, *Os4CL*, *OsCHS*, and *OsCHI*) with the endosperm-specific *GluB*-1 promoter into one vector has been done through *Agrobacterium*-mediated transformation (Ogo et al., 2013). As multiple-expression cassette is quite large, some of the parts were deleted in a few plants; as a result, none of the transgenic plants did not express all the five genes. Generally, plants were found to have one to four expression cassettes. The resulted plant that is lacking either *OsPAL* or *OsCHS* gene does not accumulate naringenin, but has both genes accumulating naringenin in the seeds. Furthermore, these selected genes were transformed exclusively with embryo- and endosperm-specific 18 kDa *oleosin* and *GluB-1* promoters for the identification of isoflavone localization. LC-MS analysis revealed the accumulation of naringenin and some flavones. Also the expression of additional genes for flavone (*PoFNSI*/*GmFNSII*), isoflavone (*GmIFS*), and flavonol (*AtF3H*/*AtFLS*) has resulted in the accumulation of kaempferol, genistein derivatives, quercetin, chrysoeriol, and tricin, etc. (**Table 1**). Song et al. (2013) aimed to generate colored rice seeds and increased accumulation of isoflavone that expressed *maize-Lc* gene under the control of endosperm-specific rice *glutelin Gt-1* promoter. As a result, the transgenic rice was found to be with dark color and the total isoflavone content also increased. Expression of two soybean isoflavone synthases (*SpdIFS1* and *SpdIFS2*) under the control of endosperm-specific 26-kDa *globulin* promoter in rice varieties accumulated a higher level of genistein in seeds (Sohn et al., 2014). As the two rice varieties has different seed color (black and normal white), the accumulation of isoflavone was also found to be varied. The white variety (103 μg/g) of transgenic rice has accumulated more genistein than black variety (87 μg/g) (**Table 1**).

In an attempt of isoflavone biosynthesis in tomato, Shih et al. (2008) have transformed soybean *IFS* (*GmIFS*) gene under the control of the *CaMV35S* promoter. The transgenic plants accumulated genistein in a tissue-specific manner. The LC-MS analysis has found that substantial amount of genistein has been accumulated in the leaves, whereas the little amount has been accumulated in fruit peel. Although *Brassica napus*, a nonleguminous oil crop, produced phenylpropanoids and flavonoids but not isoflavones because of the absence of the *IFS* gene. Incorporation of exogenous *GmIFS2* showed the accumulation of genistein in the leaves of transgenic plants up to 0.72 mg/g dry weight (Li et al., 2011) (**Table 1**). With the aim of increased accumulation of genistein in transgenic tobacco, Pandey et al. (2014) has engineered the tobacco with the coexpression of *GmIFS* and transcription factor *AtMYB12*. The transgenic plants accumulated higher levels of flavonols and genistein conjugates compared to control plants. Gou et al. (2016) have transformed different combinations of isoflavone pathway genes (*CHS*, *CHI*, *IFS*, and *F3H*) in *Medicago truncatula* for the increased accumulation of isoflavones and proanthocyanidins. Downregulation of *MtF3H* in combination with overexpression of *GmIFS1*, *GmCHS7*, and *GmCHI1A* was found to be more effective in elevated accumulation of isoflavones, flavones, and proanthocyanidins. Recently, Malla et al. (2021) have transformed *GmIFS* into onion (*Allium cepa* L.) for the accumulation of genistein in transformed onion plants through both biolistic gene transfer and *Agrobacterium*-mediated gene transfer methods. The results showed that a higher level of genistein was accumulated in biolistic gene transfer (62.65 nM/g FW) than *Agrobacterium*-mediated gene transfer method. Thus, introducing the transcription factor for the positive or negative regulation of the isoflavone biosynthesis proved the efficiency of transcription factors, suggesting that it can be utilized for both leguminous and nonleguminous plants for the regulation of isoflavone biosynthesis.

CONCLUDING REMARKS AND FUTURE PROSPECTS

Isoflavones have important roles in plants, environment, humans, and other animals with their compatible chemical structures. Their general occurrence in the soybean and its applications in various diseases, as we highlighted, necessitated the detailed understanding of the metabolic engineering process and the regulation of isoflavone biosynthesis in the plants. We have provided an overview of the isoflavone biosynthesis through metabolic engineering in various crop plants. The recent progression in the identification and characterization of important enzymes in the isoflavone biosynthetic pathway has helped to improve the level of isoflavones in plants. However, as phenylpropanoid pathway is a complex pathway for isoflavone biosynthesis, simultaneous engineering of multiple genes and understanding the crosstalk between pathways are important. That helps us to understand the role of a specific branch of the pathway in the biosynthesis of isoflavones and other secondary metabolites. The isoflavone-level variation

is a complex mechanism regulated by various genetic and environmental factors. Transcriptional regulation of isoflavone biosynthesis occurs by modifying the transcription factors such as MYBs either to increase or decrease the level of isoflavones in soybean (Yi et al., 2010; Yan et al., 2015; Bian et al., 2018; Jahan et al., 2020). In addition to that, posttranslational regulation of isoflavone biosynthesis through ubiquitination and SUMOylation processes has become necessary tools for clear understanding of how isoflavones can be produced. Likewise, synthetic metabolic engineering technologies, such as a multigene expression vector system (CRISPR/cas9)–based system, should be used efficiently for expression and regulation of isoflavone biosynthetic genes in a precise manner (Zhang et al., 2020). With the deep understanding of the isoflavone biosynthetic pathway, novel technologies in metabolic engineering and synthetic metabolic engineering will decipher light on the regulation of complex metabolic networks and can produce biofortified crops with required isoflavone levels to meet better human health.

AUTHOR CONTRIBUTIONS

SIS and SP conceived the review and wrote the manuscript and made a critical revision of the review. YJO, HJK, WSC, and YSC performed the literature search. SP and YJO prepared figures and tables. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the New Breeding Technologies Development Program (Project No. PJ01492902), Rural Development Administration, Republic of Korea.

REFERENCES

Abdel-Lateif, K., Bogusz, D., and Hocher, V. (2012). The role of flavonoids in the establishment of plant roots endosymbioses with arbuscular mycorrhiza fungi, rhizobia and *Frankia bacteria*. *Plant Signal. Behav.* 7, 636–641. doi: 10.4161/psb.20039

Abdi, F., Alimoradi, Z., Haqi, P., and Mahdizad, F. (2016). Effects of phytoestrogens on bone mineral density during the menopause transition: a systematic review of randomized, controlled trials. *Climacteric* 19, 535–545. doi: 10.1080/13697137.2016.1238451

Ahmad, M. Z., Li, P., Wang, J., Rehman, N. U., and Zhao, J. (2017). Isoflavone malonyltransferases GmiMaT1 and GmiMaT3 differently modify isoflavone glucosides in soybean (*Glycine max*) under various stresses. *Front. Plant Sci.* 8:735. doi: 10.3389/fpls.2017.00735

Ahmad, M. Z., Zhang, Y., Zeng, X., Li, P., Wang, X., Benedito, V. A., et al. (2020). Isoflavone malonyl-CoA acyltransferase GmMaT2 is involved in nodulation of soybean (*Glycine max*) by modifying synthesis and secretion of isoflavones. *J. Exp. Bot.* 101, 106–109. doi: 10.1093/jxb/eraa511

Akashi, T., Aoki, T., and Ayabe, S. (1999). Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. *Plant Physiol.* 121, 821–828. doi: 10.1104/pp.121.3.821

Akashi, T., Aoki, T., and Ayabe, S. (2005). Molecular and biochemical characterization of 2-hydroxyisoflavanone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis. *Plant Physiol.* 13, 882–891. doi: 10.1104/pp.104.056747

Anguraj Vadivel, A. K., Renaud, J., Kagale, S., and Dhaubhadel, S. (2019). GmMYB176 regulates multiple steps in isoflavonoid biosynthesis in soybean. *Front. Plant Sci.* 10:562. doi: 10.3389/fpls.2019.00562

Appert, C., Logemann, E., Hahlbrock, K., Schmid, J., and Amrhein, N. (1994). Structural and catalytic properties of the four phenylalanine ammonia-lyase isozymes from parsley (*Petroselinum crispum* Nym.). *Eur. J. Biochem.* 225, 491–499. doi: 10.1111/j.1432-1033.1994.00491.x

Atcharaporn, K., Thawunporn, P., and Lamsal, B. (2014). "Isoflavones–extraction and bioavailability," in *Functional Foods and Dietry Supplements*, eds A. K. Anal, A. Noomhorm, and I. Ahmad (Hoboken, NJ: Wiley), 255–293. doi: 10.10002/9781118227800.ch10

Ayabe, S.-I., Udagawa, A., and Furuya, T. (1988). NAD(P)H-dependent 6′-deoxychalcone synthase activity in *Glycyrrhiza echinata* cells induced by yeast extract. *Arch. Biochem. Biophys.* 261, 458–462. doi: 10.1016/0003-9861(88)90362-1

Ballance, G. M., and Dixon, R. A. (1995). Medicago sativa cDNAs encoding chalcone reductase. *Plant Physiol.* 107:1027. doi: 10.1104/pp.107.3.1027

Bate, N. J., Orrt, J., Nit, W., Meromit, A., Nadler-hassar, T., Doerner, P. W., et al. (1994). Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determining step in natural product synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 91, 7608–7612. doi: 10.1073/pnas.91.16.7608

Bian, S., Li, R., Xia, S., Liu, Y., Jin, D., Xie, X., et al. (2018). Soybean CCA1-like MYB transcription factor GmMYB133 modulates isoflavonoid biosynthesis. *Biochem. Biophys. Res. Commun.* 507, 324–329. doi: 10.1016/j.bbrc.2018.11.033

Burdick, A. B. (1958). New mutants. *Tomato Genet. Coop. Rep.* 8, 9–11.

Caetano-Anolles, G., Wrobel-Boerner, E., and Bauer, W. D. (1992). Growth and movement of spot inoculated *Rhizobium meliloti* on the root surface of alfalfa. *Plant Physiol.* 98, 1181–1189. doi: 10.1104/pp.98.3.1181

Chapple, C. (1998). Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases. *Ann. Rev. Plant. Physiol. Plant. Mol. Biol.* 49, 311–343. doi: 10.1146/annurev.arplant.49.1.311

Chen, L. R., Ko, N. Y., and Chen, K. H. (2019). Isoflavone supplements for menopausal women: a systematic review. *Nutrients* 11:2649. doi: 10.3390/nu11112649

Cheng, H., Li, L., Cheng, S., Cao, F., Wang, Y., and Yuan, H. (2011). Molecular cloning and function assay of a chalcone isomerase gene (GbCHI) from *Ginkgo biloba*. *Plant Cell Rep.* 30, 49–62. doi: 10.1007/s00299-010-0943-4

Cheng, S. Y., Wang, Y., Liu, W. H., Du, H. W., and Chen, K. S. (2005). Effects of plant growth regulators on phenylalanine ammonia-lyase (PAL) activities in leaves of *Ginkgo biloba*. *J. Plant Res. Environ.* 14, 20–22.

Chi, X. X., Zhang, T., Zhang, D. J., Yu, W., Wang, Q. Y., and Zhen, J. L. (2016). Effects of isoflavones on lipid and apolipoprotein levels in patients with type 2 diabetes in Heilongjiang Province in China. *J. Clinical Biochem. Nutr.* 59, 134–138. doi: 10.3164/jcbn.15-147

Chu, S., Wang, J., Zhu, Y., Liu, S., Zhou, X., Zhang, H., et al. (2017). An R2R3-type MYB transcription factor, GmMYB29, regulates isoflavone biosynthesis in soybean. *PLoS Genet.* 13:e1006770. doi: 10.1371/journal.pgen.1006770

Cohen-Gonsaud, M., Ducasse, S., Hoh, F., Zerbib, D., Labesse, G., and Quemard, A. (2002). Crystal structure of MabA from *Mycobacterium tuberculosis*, a reductase involved in long-chain fatty acid biosynthesis. *J. Mol. Biol.* 320, 249–261. doi: 10.1016/S0022-2836(02)00463-1

Compton, K. K., Hildreth, S. B., Helm, R. F., and Scharf, B. E. (2020). An updated perspective on *Sinorhizobium meliloti* chemotaxis to alfalfa flavonoids. *Front. Microbiol.* 11:581482. doi: 10.3389/fmicb.2020.581482

Cui, Y., Niu, K., Huang, C., Momma, H., Guan, L., Kobayashi, Y., et al. (2015). Relationship between daily isoflavone intake and sleep in Japanese adults: a cross-sectional study. *Nutr. J.* 14, 1–7. doi: 10.1186/s12937-015-0117-x

Daily, J. W., Ko, B. S., Ryuk, J., Liu, M., Zhang, W., and Park, S. (2019). Equol decreases hot flashes in postmenopausal women: a systematic review and meta-analysis of randomized clinical trials. *J. Med. Food.* 22, 127–139. doi: 10.1089/jmf.2018.4265

Deavours, B. E., and Dixon, R. A. (2005). Metabolic engineering of isoflavonoid biosynthesis in alfalfa. *Plant Physiol.* 138, 2245–2259. doi: 10.1104/pp.105.062539

Del Cerro, P., Megías, M., López-Baena, F. J., Gil-Serrano, A., Pérez-Montaño, F., and Ollero, F. J. (2019). Osmotic stress activates nif and fix genes and induces the *Rhizobium tropici* CIAT 899 Nod factor production via NodD2 by up-regulation of the nodA2 operon and the nodA3 gene. *PLoS One* 14:e0213298. doi: 10.1371/journal.pone.0213298

Dhaubhadel, S., Gijzen, M., Moy, P., and Farhangkhoee, M. (2007). Transcriptome analysis reveals a critical role of CHS7 and CHS8 genes for isoflavonoid synthesis in soybean seeds. *Plant Physiol.* 143, 326–338. doi: 10.1104/pp.106.086306

Dhaubhadel, S., McGarvey, B. D., Williams, R., and Gijzen, M. (2003). Isoflavonoid biosynthesis and accumulation in developing soybean seeds. *Plant Mol. Biol.* 53, 733–743. doi: 10.1023/B:PLAN.0000023666.30358.ae

Dixon, R. A., and Ferreira, D. (2002). Molecules of interest, genistein. *Phytochemistry* 60, 205–211. doi: 10.1016/S0031-9422(02)00116-4

Dixon, R. A., and Paiva, N. L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell.* 93, 385–392. doi: 10.1105/tpc.7.7.1085

Dixon, R. A., and Steele, C. L. (1999). Flavonoids and isoflavonoids a gold mine for metabolic engineering trends. *Plant Sci.* 4, 394–400. doi: 10.1016/s1360-1385(99)01471-5

Dixon, R. A., and Sumner, L. W. (2003). Legume natural products: understanding and manipulating complex pathways for human and animal health. *Plant Physiol.* 131, 878–885. doi: 10.1104/pp.102.017319

Druka, A., Kudrna, D., Rostoks, N., Brueggeman, R., Von, W. D., and Kleinhofs, A. (2003). Chalcone isomerase from rice (*Oryza sativa*) and barley (*Hordeum vulgare*): physical, genetic and mutation mapping. *Gene* 302, 171–178. doi: 10.1016/S0378-1119(02)01105-1108

Du, H., Huang, Y., and Tang, Y. (2010). Genetic and metabolic engineering of isoflavonoid biosynthesis. *Appl. Microbiol. Biotechnol.* 86, 1293–1312. doi: 10.1007/s00253-010-2512-8

Feinbaum, R. L., and Ausubel, F. M. (1988). Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. *Mol. Cell. Biol.* 8, 1985–1992. doi: 10.1128/mcb.8.5.1985

Fisher, R. F., and Long, S. R. (1992). Rhizobium–plant signal exchange. *Nature* 357, 655–660. doi: 10.1038/357655a0

Francini, A., Nali, C., Pellegrini, E., and Lorenzini, G. (2008). Characterization and isolation of some genes of the shikimate pathway in sensitive and resistant *Centaurea jacea* plants after ozone exposure. *Environ. Pollut.* 151, 272–279. doi: 10.1016/j.envpol.2007.07.007

Gaworzewska, E. T., and Carlile, M. J. (1982). Postive chemotaxis of *Rhizobium leguminosarum* and other bacteria towards root exudates from legumes and other plants. *J. Gen. Microbiol.* 128, 789–798. doi: 10.1099/00221287-128-6-1179

Geeta, G., Widodo, W. S., Widowati, W., Ginting, C. N., Lister, I. N. E., Armansyah, A., et al. (2019). Comparison of antioxidant and anti-collagenase activity of genistein and epicatechin. *Pharm. Sci. Res.* 6, 111–117. doi: 10.7454/psr.v6i2.4510

Gilbert, E. R., and Liu, D. (2013). Anti-diabetic functions of soy isoflavone genistein: mechanisms underlying its effects on pancreatic β-cell function. *Food. Funct.*4, 200–212. doi: 10.1039/c2fo30199g

Given, N. K., Venis, M. A., and Gierson, D. (1988). Hormonal regulation of ripening in the strawberry, a non-climacteric fruit. *Planta* 174, 402–406. doi: 10.1007/BF00959527

Graham, T. L. (1991). Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates. *Plant Physiol.* 95, 594–603. doi: 10.1104/pp.95.2.594

Graham, T. L. (1995). "Cellular biochemistry of phenylpropanoid responses of soybean to infection by *Phytophthora sojae*," in *Handbook of Phytoalexin Metabolism and Action*, eds M. Daniel and R. P. Purkayanstha (New York, NY: Marcel Dekker), 85–116. doi: 10.1201/9780203752647-5

Graham, T. L., Graham, M. Y., Subramanian, S., and Yu, O. (2007). RNAi silencing of genes for elicitation or biosynthesis of 5-deoxyisoflavonoids suppresses race-specific resistance and hypersensitive cell death in *Phytophthora sojae* infected tissues. *Plant Physiol.* 144, 728–740. doi: 10.1104/pp.107.097865

Grotewold, E., Chamberlin, M., Snook, M., Siame, B., Butler, L., Swenson, J., et al. (1998). Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. *Plant Cell.* 10, 721–740. doi: 10.1105/tpc.10.5.721

Guo, J., Zhou, W., Lu, Z., Li, H., Li, H., and Gao, F. (2015). Isolation and functional analysis of chalcone isomerase gene from purple-fleshed sweet potato. *Plant Mol. Biol. Rep.* 33, 1451–1463. doi: 10.1007/s11105-014-0842-x

Gou, L., Zhang, R., Ma, L., Zhu, F., Dong, J., and Wang, T. (2016). Multigene synergism increases the isoflavone and proanthocyanidin contents of *Medicago truncatula*. *Plant Biotechnol. J.* 14, 915–925. doi: 10.1111/pbi.12445

Gupta, O. P., Nigam, D., Dahuja, A., Kumar, S., Vinutha, T., Sachdev, A., et al. (2017). Regulation of isoflavone biosynthesis by miRNAs in two contrasting soybean genotypes at different seed developmental stages. *Front. Plant Sci.* 8:567. doi: 10.3389/fpls.2017.00567

Hakamatsuka, T., Mori, K., Ishida, S., Ebizuka, Y., and Sankawa, U. (1998). Purification of 2-hydroxyisoflavanone dehydratase from the cell cultures of *Pueraria lobata*. *Phytochemistry* 49, 497–505. doi: 10.1016/S0031-9422(98)00266-0

Hakamatsuka, T., Noguchi, H., Ebizuka, Y., and Sankawa, U. (1990). Isoflavone synthase from cell suspension cultures of *Pueraria lobata*. *Chem. Pharm. Bull.* 38, 1942–1945. doi: 10.1248/cpb.38.1942

Hayashi, H., Fujii, I., Iinuma, M., Shibano, M., Ozaki, K., and Watanabe, H. (2013). Characterization of a glycyrrhizin-deficient strain of *Glycyrrhiza uralensis*. *Biol. Pharm. Bull.* 36, 1448–1453. doi: 10.1248/bpb.b13-00164

He, X., Blount, J., Ge, S., Tang, Y., and Dixon, R. (2011). A genomic approach to isoflavone biosynthesis in kudzu (*Pueraria lobata*). *Planta* 233, 843–855. doi: 10.1007/s00425-010-1344-1

Holton, T. A., and Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell.* 7:1071. doi: 10.1105/tpc.7.7.1071

Hooper, L., Madhavan, G., Tice, J. A., Leinster, S. J., and Cassidy, A. (2010). Effects of isoflavones on breast density in pre-and post-menopausal women: a systematic review and meta-analysis of randomized controlled trials. *Hum. Reprod. Update* 16, 745–760. doi: 10.1093/humupd/dmq011

Hrazdina, G., Parsons, G. F., and Mattick, L. R. (1984). Physiological and biochemical events during development and maturation of grape berries. *Am. J. Enol. Vitic.* 35, 220–227.

Hrazdina, G., and Wagner, G. J. (1985). Metabolic pathways as enzyme complexes: evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. *Arch. Biochem. Biophys.* 237, 88–100. doi: 10.1016/0003-9861(85)90257-7

Hu, Y. S., Zhang, L., Di, P., and Chen, W. S. (2009). Cloning and induction of phenylalanine ammonia-lyase gene from *Salvia miltiorrhiza* and its effect on hydrophilic phenolic acids levels. *Chin. J. Nat. Med.* 7, 449–457. doi: 10.1016/S1875-5364(09)60069-8

Hur, S., Newby, Z. E., and Bruice, T. C. (2004). Transition state stabilization by general acid catalysis, water expulsion, and enzyme reorganization in Medicago savita chalcone isomerase. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2730–2735. doi: 10.1073/pnas.0308264100

Hyun, M. W., Yun, Y. H., Kim, J. Y., and Kim, S. H. (2011). Fungal and plant phenylalanine ammonia-lyase. *Mycobiology.* 39, 257–265. doi: 10.5941/MYCO.2011.39.4.257

Iyer, B., and Rajkumar, S. (2017). Host specificity and plant growth promotion by bacterial endophytes. *Curr. Res. Microbiol. Biotechnol.* 5, 1018–1030.

Jahan, M. A., Harris, B., Lowery, M., Coburn, K., Infante, A. M., Percifield, R. J., et al. (2019). The NAC family transcription factor GmNAC42–1 regulates biosynthesis of the anticancer and neuroprotective glyceollins in soybean. *BMC Genomics* 20:149. doi: 10.1186/s12864-019-5524-5

Jahan, M. A., Harris, B., Lowery, M., Infante, A. M., Percifield, R. J., and Kovinich, N. (2020). Glyceollin transcription factor GmMYB29A2 regulates soybean resistance to *Phytophthora sojae*. *Plant Physiol.* 183, 530–546. doi: 10.1104/pp.19.01293

Jalili, M., Hekmatdoost, A., Vahedi, H., Poustchi, H., Khademi, B., Saadi, M., et al. (2016). Co-Administration of soy isoflavones and Vitamin D in management of irritable bowel disease. *PLoS One* 11:e0158545. doi: 10.1371/journal.pone.0158545

Jeandet, P., Clement, C., Courot, E., and Cordelier, S. (2013). Modulation of phytoalexin biosynthesis in engineered plants for disease resistance. *Intl. J. Mol. Sci.* 14, 14136–14170. doi: 10.3390/ijms140714136

Jez, J. M., Bowman, M. E., Dixon, R. A., and Noel, J. P. (2000). Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nat. Struct. Biol.* 7, 786–791. doi: 10.1038/79025

Jez, J. M., Bowman, M. E., and Noel, J. P. (2002). Role of hydrogen bonds in the reaction mechanism of chalcone isomerase. *Biochemistry* 41, 5168–5176. doi: 10.1021/bi0255266

Jez, J. M., and Noel, J. P. (2000). Mechanism of chalcone synthase. *J. Biol. Chem.* 275, 39640–39646. doi: 10.1074/jbc.M008569200

Jiang, Y., Hu, Y., Wang, B., and Wu, T. (2014). Bivalent RNA interference to increase isoflavone biosynthesis in soybean (*Glycine max*). *Braz. Arch. Biol. Technol.* 57, 163–170. doi: 10.1590/S1516-89132013005000018

Jiang, Y. N., Wang, B., Li, H., Yao, L. M., and Wu, T. L. (2010). Flavonoid production is effectively regulated by RNAi interference of two flavone synthase genes from *Glycine max*. *J. Plant Biol.* 53, 425–432. doi: 10.1007/s12374-010-9132-9

Jorrin, J., and Dixon, R. A. (1990). Stress responses in alfalfa (*Medicago sativa* L.). Purification, characterization, and induction of phenylalanine ammonia-lyase isoforms from elicitor-treated cell suspension cultures. *Plant Physiol.* 92, 447–455. doi: 10.1104/pp.92.2.447

Jung, W., Yu, O., Lau, S. -M. C., O'Keefe, D. P., Odell, J., Fader, G., et al. (2000). Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nat. Biotech* 18, 208–212. doi: 10.1038/72671

Kang, J. H., McRoberts, J., Shi, F., Moreno, J. E., Jones, A. D., and Howe, G. A. (2014). The flavonoid biosynthetic enzyme chalcone isomerase modulates terpenoid production in glandular trichomes of tomato. *Plant Physiol.* 164, 1161–1174. doi: 10.1104/pp.113.233395

Kervinen, T., Peltonen, S., Utriainen, M., Kangasjärvi, J., Teeri, T. H., and Karjalainen, R. (1997). Cloning and characterization of cDNA clones encoding phenylalanine ammonia-lyase in barley. *Plant Sci.* 123, 143–150. doi: 10.1016/S0168-9452(96)04570-0

Kochs, G., and Grisebach, H. (1986). Enzymic synthesis of isoflavones. *Eur. J. Biochem.* 155, 311–318. doi: 10.1111/j.1432-1033.1986.tb09492.x

Koes, R. E., Spelt, C. E., and Mol, J. N. (1989). The chalcone synthase multigene family of Petunia hybrida (V30): differential, light-regulated expression during flower development and UV light induction. *Plant Mol. Biol.* 12, 213–225. doi: 10.1007/BF00020506

Kraszewska, O., Nynca, A., Kaminska, B., and Ciereszko, R. (2007). Fitoestrogeny. 1. Wystepowanie, metabolizm i znaczenie biologiczne u samic. *Postêpy Biol. Komórki* 1, 189–205.

Kreuzaler, F., Ragg, H., Fautz, E., Kuhn, D. N., and Hahlbrock, K. (1983). UV-induction of chalcone synthase mRNA in cell suspension cultures of *Petroselinum hortense*. *Proc. Natl. Acad. Sci. U.S.A.* 80, 2591–2593. doi: 10.1073/pnas.80.9.2591

Křížová, L., Dadáková, K., Kašparovská, J., and Kašparovský, T. (2019). Isoflavones. *Molecules* 24:1076. doi: 10.3390/molecules24061076

Kumar, N. B., Cantor, A., Allen, K., Riccardi, D., and Cox, C. E. (2002). The specific role of isoflavones on estrogen metabolism in premenopausal women. *Cancer* 94, 1166–1174. doi: 10.1002/cncr.10320

Lepelley, M., Mahesh, V., McCarthy, J., Rigoreau, M., Crouzillat, D., Chabrillange, N., et al. (2012). Characterization, high-resolution mapping and differential expression of three homologous PAL genes in *Coffea canephora* Pierre (Rubiaceae). *Planta* 236, 313–326. doi: 10.1007/s00425-012-1613-2

Li, X., Qin, J. C., Wang, Q. Y., Wu, X., Lang, C. Y., Pan, H. Y., et al. (2011). Metabolic engineering of isoflavone genistein in *Brassica napus* with soybean isoflavone synthase. *Plant Cell Rep.* 30, 1435–1442. doi: 10.1007/s00299-011-1052-8

Lim, W., and Li, J. (2016). Co-expression of onion chalcone isomerase in Del/Ros1-expressing tomato enhances anthocyanin and flavonol production. *Plant Cell Tiss. Org. Cult.* 128, 113–124. doi: 10.1007/s11240-016-1090-6

Liu, C. J., Blount, J. W., Steele, C. L., and Dixon, R. A. (2002). Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*. *Proc. Nat. Acad. Sci. U.S.A* 99, 14578–14583. doi: 10.1073/pnas.212522099

Liu, C. W., and Murray, J. D. (2016). The role of flavonoids in nodulation host-range specificity: an update. *Plants* 5:33. doi: 10.3390/plants5030033

Liu, R., Hu, Y., Li, J., and Lin, Z. (2007). Production of soybean isoflavone genistein in non-legume plants via genetically modified secondary metabolism pathway. *Metab. Eng.* 9, 1–7. doi: 10.1016/j.ymben.2006.08.003

Liu, R., Xu, S., Li, J., Hu, Y., and Lin, Z. (2006). Expression profile of a PAL gene from *Astragalus membranaceus* var. Mongholicus and its crucial role in flux into flavonoid biosynthesis. *Plant Cell Rep.* 25, 705–710. doi: 10.1007/s00299-005-0072-7

Liu, X., Yuan, L., Xu, L., Xu, Z., Huang, Y., He, X., et al. (2013). Over-expression of GmMYB39 leads to an inhibition of the isoflavonoid biosynthesis in soybean (*Glycine max*. L). *Plant Biotechnol. Rep.* 7, 445–455. doi: 10.1007/s11816-013-0283-2

Lozovaya, V. V., Lygin, A. V., Zernova, O. V., Li, S., Hartman, G. L., and Widholm, J. M. (2004). Isoflavonoid accumulation in soybean hairy roots upon treatment with Fusarium solani. *Plant Physiol. Biochem.* 42, 671–679. doi: 10.1016/j.plaphy.2004.06.007

Lozovaya, V. V., Lygin, A. V., Zernova, O. V., Ulanov, A. V., Li, S., Hartman, G. L., et al. (2007). Modification of phenolic metabolism in soybean hairy roots through down regulation of chalcone synthase or isoflavone synthase. *Planta* 225, 665–679. doi: 10.1007/s00425-006-0368-z

Mahmoud, A. M., Yang, W., and Bosland, M. C. (2014). Soy isoflavones and prostate cancer: a review of molecular mechanisms. *J. Steroid. Biochem. Mol. Biol.* 140, 116–132. doi: 10.1016/j.jsbmb.2013.12.010

Malla, A., Shanmugaraj, B., Srinivasan, B., Sharma, A., and Ramalingam, S. (2021). Metabolic engineering of isoflavonoid biosynthesis by expressing *Glycine max* isoflavone synthase in *Allium cepa* L. for genistein production. *Plants* 10:52. doi: 10.3390/plants10010052

Mameda, R., Waki, T., Kawai, Y., Takahashi, S., and Nakayama, T. (2018). Involvement of chalcone reductase in the soybean isoflavone metabolon: identification of GmCHR5, which interacts with 2-hydroxyisoflavanone synthase. *Plant J.* 96, 56–74. doi: 10.1111/tpj.14014

Manning, K. (1998). Isolation of a set of ripening-related genes from strawberry: their identification and possible relationship to fruit quality traits. *Planta* 205, 622–631. doi: 10.1007/s004250050365

Matsuda, H., Nakayasu, M., Aoki, Y., Yamazaki, S., Nagano, A. J., Yazaki, K., et al. (2020). Diurnal metabolic regulation of isoflavones and soyasaponins in soybean roots. *Plant Direct* 4:e00286. doi: 10.1002/pld3.286

Matsumura, H., Watanabe, S., Harada, K., Senda, M., Akada, S., Kawasaki, S., et al. (2005). Molecular linkage mapping and phylogeny of the chalcone synthase multigene family in soybean. *Theor. Appl. Genet.* 110, 1203–1209. doi: 10.1007/s00122-005-1950-7

Mavandad, M., Edwards, R., Liang, X., Lamb, C. J., and Dixon, R. A. (1990). Effects of trans-cinnamic acid on expression of the bean phenylalanine ammonia-lyase gene family. *Plant Physiol.* 113, 755–763. doi: 10.1104/pp.94.2.671

McCue, P., and Shetty, K. (2004). Health benefits of soy isoflavonoids and strategies for enhancement: a review. *Crit. Rev. Food. Sci.* 44, 361–367. doi: 10.1080/10408690490509591

Messina, M. J., and Loprinzi, C. L. (2001). Soy for breast cancer survivors: a critical review of the literature. *J. Nutr.* 131, 3095S–3108S. doi: 10.1093/jn/131.11.3095S

Messina, M. J., and Wood, C. E. (2008). Soy isoflavones, estrogen therapy, and breast cancer risk; analysis and commentary. *Nutr. J.* 7:17. doi: 10.1186/1475-2891-7-17

Misra, P., Pandey, A., Tewari, S. K., Nath, P., and Trivedi, P. K. (2010). Characterization of isoflavone synthase gene from *Psoralea corylifolia*: a medicinal plant. *Plant Cell Rep.* 29, 747–755. doi: 10.1007/s00299-010-0861-5

Muir, S. R., Collins, G. J., Robinson, S., Hughes, S., Bovy, A., Ric, D. V. C. H., et al. (2001). Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nat. Biotechnol.* 19, 470–474. doi: 10.1038/88150

Mulligan, J. T., and Long, S. R. (1985). Induction of *Rhizobium meliloti* nodC expression by plant exudate requires nodD. *Proc. Nat. Acad.Sci. U.S.A.* 82, 6609–6613. doi: 10.1073/pnas.82.19.6609

Nachvak, S. M., Moradi, S., Anjom-Shoae, J., Rahmani, J., Nasiri, M., Maleki, V., et al. (2019). Soy, soy isoflavones, and protein intake in relation to mortality from all causes, cancers, and cardiovascular diseases: a systematic review and dose–response meta-analysis of prospective cohort studies. *J. Acad. Nutr. Diet.* 119, 1483–1500. 10.1016/j.jand.2019.04.011

Nagai, N., Kitauchi, F., Shimosaka, M., and Okazaki, M. (1994). Cloning and sequencing of a full-length cDNA coding for phenylalanine ammonia-lyase from tobacco cell culture. *Plant Physiol.* 104:1091. doi: 10.1104/pp.104.3.1091

Napoli, C. A., Fahy, D., Wang, H. Y., and Taylor, L. P. (1999). White anther: a petunia mutant that abolishes pollen flavonol accumulation, induces male sterility, and is complemented by a chalcone synthase transgene. *Plant Physiol.* 120, 615–622. doi: 10.1104/pp.120.2.615

Nestel, P., Cehun, M., Chronopoulos, A., DaSilva, L., Teede, H., and McGrath, B. (2004). A biochanin-enriched isoflavone from red clover lowers LDL cholesterol in men. *Eur. J Clinic. Nutr.* 58, 403–408. doi: 10.1038/sj.ejcn.1601796

Nguyen, H. Q., Le, T. H. T., Nguyen, T. N. L., Nguyen, T. G., Sy, D. T., Tu, Q. T., et al. (2020). Overexpressing GmCHI1A increases the isoflavone content of transgenic soybean (*Glycine max* (L.) Merr.) seeds. *In Vitro Cell Dev. Biol. Plant* 56, 842–850. doi: 10.1007/s11627-020-10076-x

Ogo, Y., Ozawa, K., Ishimaru, T., Murayama, T., and Takaiwa, F. (2013). Transgenic rice seed synthesizing diverse flavonoids at high levels: a new platform for flavonoid production with associated health benefits. *Plant Biotechnol. J.* 11, 734–746. doi: 10.1111/pbi.12064

Oldroyd, G. E., Murray, J. D., Poole, P. S., and Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* 45, 119–144. doi: 10.1146/annurev-genet-110410-132549

Pandey, A., Misra, P., Khan, M. P., Swarnkar, G., Tewari, M. C., Bhambhani, S., et al. (2014). Co-expression of *Arabidopsis* transcription factor, AtMYB12, and soybean isoflavone synthase, GmIFS1, genes in tobacco leads to enhanced biosynthesis of isoflavones and flavonols resulting in osteoprotective activity. *Plant Biotechnol. J.* 12, 69–80. doi: 10.1111/pbi.12118

Paxton, J. D. (1981). Phytoalexins-a working redefinition. *J. Phytopathol.* 101, 106–109. doi: 10.1111/j.1439-0434.1981.tb03327.x

Perna, S., Peroni, G., Miccono, A., Riva, A., Morazzoni, P., Allegrini, P., et al. (2016). Multidimensional Effects of soy isoflavone by food or supplements in menopause women: a systematic review and bibliometric analysis. *Nat. Prod. Commun.* 11, 1733–1740. 10.1177/1934578X1601101127

Philips, D. A., and Tsai, S. M. (1992). Flavonoids as plant signals to rhizosphere microbes. *Mycorrhiza* 1, 55–58. doi: 10.1007/BF00206136

Price, A. C., Zhang, Y. M., Rock, C. O., and White, S. W. (2001). Structure of β-ketoacyl-[acyl carrier protein] reductase from *Escherichia coli*: negative cooperativity and its structural basis. *Biochemistry* 40, 12772–12781. doi: 10.1021/bi010737g

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

Rasmussen, S., and Dixon, R. A. (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. *Plant Cell.* 11, 1537–1551. doi: 10.1105/tpc.11.8.1537

Reynaud, J., Guilet, D., Terreux, R., Lussignol, M., and Walchshofer, N. (2005). Isoflavonoids in non-leguminous families, an update. *Nat. Prod. Rep.* 22, 504–515. doi: 10.1039/b416248j

Rípodas, C., Dalla, V., Aguilar, O. M., Zanetti, M. E., and Blanco, F. A. (2013). Knock-down of a member of the isoflavone reductase gene family impairs plant growth and nodulation in *Phaseolus vulgaris*. *Plant Physiol. Biochem.* 68, 81–89. doi: 10.1016/j.plaphy.2013.04.003

Ritter, H., and Schulz, G. E. (2004). Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. *Plant Cell.* 16, 3426–3436. doi: 10.1105/tpc.104.025288

Sarkar, M. A. R., Watanabe, S., Suzuki, A., Hashimoto, F., and Anai, T. (2019). Identification of novel MYB transcription factors involved in the isoflavone biosynthetic pathway by using the combination screening system with agroinfiltration and hairy root transformation. *Plant Biotechnol.* 36, 241–251. doi: 10.5511/plantbiotechnology.19.1025a

Schulze-Lefert, P., Becker-Andre, M., Schulz, W., Hahlbrock, K., and Dangl, J. L. (1989). Functional architecture of the light-responsive chalcone synthase promoter from parsley. *Plant Cell.* 1, 707–714. doi: 10.1105/tpc.1.7.707

Sepiol, C. J., Yu, J., and Dhaubhadel, S. (2017). Genome-wide identification of chalcone reductase gene family in soybean: insight into root-specific GmCHRs and *Phytophthora sojae* resistance. *Front. Plant Sci.* 8:2073. doi: 10.3389/fpls.2017.02073

Shih, C. H., Chen, Y., Wang, M., Chu, I. K., and Lo, C. (2008). Accumulation of isoflavone genistin in transgenic tomato plants overexpressing a soybean isoflavone synthase gene. *J. Agricult. Food Chem.* 56, 5655–5661. doi: 10.1021/jf800423u

Shimada, N., Aoki, T., Sato, S., Nakamura, Y., Tabata, S., and Ayabe, S. (2003). A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy(iso)flavonoids in *Lotus japonicus*. *Plant Physiol.* 131, 941–951. doi: 10.1104/pp.004820

Shimada, N., Nakatsuka, T., Nishihara, M., Yamamura, S., Ayabe, S., and Aoki, T. (2006). Isolation and characterization of a cDNA encoding polyketide reductase in *Lotus japonicus*. *Plant Biotechnol.* 23, 509–513. doi: 10.5511/plantbiotechnology.23.509

Shimamura, M., Akashi, T., Sakura, N., Suzuki, H., Saito, K., Shibata, D., et al. (2007). 2-Hydroxyisoflavanone dehydratase is a critical determinant of isoflavone productivity in hair root cultures of *Lotus japonicus*. *Plant Cell Physiol.* 48, 1652–1657. doi: 10.1093/pcp/pcm125

Shin, Y. M., Park, H. J., Yim, S. D., Baek, N. I., Lee, C. H., An, G., et al. (2006). Transgenic rice lines expressing maize C1 and R-S regulatory genes produce various flavonoids in the endosperm. *Plant Biotechnol. J.* 4, 303–315. doi: 10.1111/j.1467-7652.2006.00182.x

Sohn, S. I., Kim, Y. H., Kim, S. L., Lee, J. Y., Oh, Y. J., Chung, J. H., et al. (2014). Genistein production in rice seed via transformation with soybean IFS genes. *Plant Sci.* 217, 27–35. doi: 10.1016/j.plantsci.2013.11.015

Song, Y. E., Wang, X., Shen, Z. W., Xu, Y., and Li, J. Y. (2013). Expressing the maize anthocyanin regulatory gene Lc increased flavonoid content in the seed of white pericarp rice and purple pericarp rice. *Russian J. Genet.* 49, 1127–1133. doi: 10.7868/s0016675813100123

Spaink, H. P. (1992). Rhizobial lipo-oligosaccharides: answers and questions. *Plant Mol. Biol.* 20, 977–986. doi: 10.1007/BF00027167

Spaink, H. P., Sheeley, D. M., van Brussel, A. A., Glushka, J., York, W. S., Tak, T., et al. (1991). A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of Rhizobium. *Nature* 354, 125–130. doi: 10.1038/354125a0

Sreevidya, V. S., Rao, S. C., Sullia, S. B., Ladha, J. K., and Reddy, P. M. (2006). Metabolic engineering of rice with soybean isoflavone synthase for promoting nodulation gene expression in rhizobia. *J. Exp. Bot.* 57, 1957–1969. doi: 10.1093/jxb/erj143

Staunton, J., and Weissman, K. J. (2001). Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* 18, 380–416. doi: 10.1039/a909079g

Steyn, W. J., Holcroft, D. M., Wand, S. J. E., and Jacobs, G. (2004). Anthocyanin degradation in detached pome fruit with reference to preharvest red color loss and pigmentation patterns of blushed and fully red pears. *J. Am. Soc. Hort. Sci.* 129, 13–19. doi: 10.21273/JASHS.129.1.0013

Subramanian, S., Graham, M. Y., Yu, O., and Graham, T. L. (2005). RNA interference of soybean isoflavone synthase genes leads to silencing in tissues distal to the transformation site and to enhanced susceptibility to *Phytophthora sojae*. *Plant Physiol.* 137, 1345–1353. doi: 10.1104/pp.104.057257

Subramanian, S., Hu, X., Lu, G., Odell, J. T., and Yu, O. (2004). The promoters of two isoflavone synthase genes respond differentially to nodulation and defense signals in transgenic soybean roots. *Plant Mol. Biol.* 54, 623–639. doi: 10.1023/B:PLAN.0000040814.28507.35

Subramanian, S., Stacey, G., and Yu, O. (2006). Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. *Plant J.* 48, 261–273. doi: 10.1111/j.1365-313X.2006.02874.x

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

Sun, W., Shen, H., Xu, H., Tang, X., Tang, M., Ju, Z., et al. (2019). Chalcone isomerase a key enzyme for anthocyanin biosynthesis in *Ophiorrhiza Japonica*. *Front. Plant Sci.* 10:865. doi: 10.3389/fpls.2019.00865

Taku, K., Melby, M. K., Nishi, N., Omori, T., and Kurzer, M. S. (2011). Soy isoflavones for osteoporosis: an evidence-based approach. *Maturitas* 70, 333–338. doi: 10.1016/j.maturitas.2011.09.001

Teutsch, H. G., Hasenfratz, M. P., Lesot, A., Garnier, J. M., Jeltsch, J. M., Durst, F., et al. (1993). Isolation and sequence of a cDNA encoding the *Jerusalem artichoke* cinnamate-4-hydroylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. *Proc. Natl. Acad. Sci. U.S.A.* 90, 4102–4107. doi: 10.1073/pnas.90.9.4102

Tian, L., and Dixon, R. A. (2006). Engineering isoflavone metabolism with an artificial bifunctional enzyme. *Planta* 224, 496–507. doi: 10.1007/s00425-006-0233-0

Tikkanen, M. J., and Adlercreutz, H. (2000). Dietary soy-derived isoflavone phytoestrogens: could they have a role in coronary heart disease prevention? *Biochem. Pharmacol.* 60, 1–5. doi: 10.1016/s0006-2952(99)00409-8

Tsuchihashi, R., Sakamoto, S., Kodera, M., Nohara, T., and Kinjo, J. (2008). Microbial metabolism of soy isoflavones by human intestinal bacterial strains. *J. Nat. Med.* 62, 456–460. doi: 10.1007/s11418-008-0271-y

Tuteja, J. H., Clough, S. J., Chan, W. C., and Vodkin, L. O. (2004). Tissue-specific gene silencing mediated by a naturally occurring chalcone synthase gene cluster in *Glycine max*. *Plant Cell* 16, 819–835. doi: 10.1105/tpc.021352

Vadivel, A. K. A., McDowell, T., Renaud, J. B., and Dhaubhadel, S. (2021). A combinatorial action of GmMYB176 and GmbZIP5 controls isoflavonoid biosynthesis in soybean (*Glycine max*). *Commun. Biol.* 4, 1–10. doi: 10.1038/s42003-021-01889-6

Van, T. A., Koes, R. E., Spelt, C. E., van der Krol, A. R., Stuitje, A. R., and Mol, J. N. (1988). Cloning of the two chalcone flavanone isomerase genes from Petunia hybrida: coordinate, light-regulated and differential expression of flavonoid genes. *EMBO J.* 7, 1257–1263. doi: 10.1002/j.1460-2075.1988.tb02939.x

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, H. C., Huang, X. M., Hu, G. B., and Huang, H. B. (2004). Studies on the relationship between anthocyanin biosynthesis and related enzymes in litchi pericarp. *Sci. Agric. Sin.* 37, 2028–2032.

Wang, L., Liu, X., Meng, X., Wu, G., and Xu, F. (2018). Cloning and expression analysis of a chalcone isomerase (CnCHI) gene from *Chamaemelum nobile*. *Biotechnology* 17, 19–25. doi: 10.3923/biotech.2018.19.25

Wang, Y., Chen, J. Y., Jiang, Y. M., and Lu, W. J. (2007). Cloning and expression analysis of phenylalanine ammonia-lyase in relation to chilling tolerance in harvested banana fruit. *Postharvest Biol. Technol.* 44, 34–41. doi: 10.1016/j.postharvbio.2006.11.003

Wang, Y. N., Wu, W., Chen, H. C., and Fang, H. (2010). Genistein protects against UVB-induced senescence-like characteristics in human dermal fibroblast by p66Shc down-regulation. *J. Dermatol. Sci.* 58, 19–27. doi: 10.1016/j.jdermsci.2010.02.002

Wei, P., Liu, M., Chen, Y., and Chen, D. C. (2012). Systematic review of soy isoflavone supplements on osteoporosis in women. *Asian Pac. J. Trop. Med.* 5, 243–248. doi: 10.1016/S1995-7645(12)60033-9

Weisshaar, B., Block, A., Armstrong, G. A., Herrmann, A., Schulze-Lefert, P., and Hahlbrock, K. (1991). Regulatory elements required for light-mediated expression of the *Petroselinum crispum* chalcone synthase gene. *Symp. Soc. Exp. Biol.* 45, 191–210.

Welle, R., and Grisebach, H. (1988). Isolation of a novel NADPH-dependent reductase which coacts with chalcone synthase in the biosynthesis of 6′-deoxychalcone. *FEBS Lett.* 236, 221–225. doi: 10.1016/0014-5793(88)80318-1

Wu, A. H., Koh, W. P., Wang, R., Lee, H. P., and Yu, M. C. (2008). Soy intake and breast cancer risk in Singapore Chinese Health Study. *Br. J. Cancer* 99, 196–200. doi: 10.1038/sj.bjc.6604448

Wu, A. H., Wan, P., Hankin, J., Tseng, C. C., Yu, M. C., and Pike, M. C. (2002). Adolescent and adult soy intake and risk of breast cancer in Asian-Americans. *Carcinogenesis* 23, 1491–1496. doi: 10.1093/carcin/23.9.1491

Wu, A. H., Ziegler, R. G., Horn-Ross, P. L., Nomura, A. M., West, D. W., Kolonel, L. N., et al. (1996). Tofu and risk of breast cancer in Asian-Americans. *Cancer Epidemiol. Prev. Biomarkers.* 5, 901–906.

Wu, L. M., Qiao, H. X., Li, Y. K., and Li, L. D. (2007). Cardioprotective effects of the combined use of puerarin and danshensu on acute ischemic myocardial injury in rats. *Phytother. Res.* 21, 751–756. doi: 10.1002/ptr.2157

Xiang, L., and Moore, B. S. (2005). Biochemical characterization of a prokaryotic phenylalanine ammonia lyase. *J. Bacteriol.* 187, 4286–4289. doi: 10.1128/JB.187.12.4286-4289.2005

Yamamoto, S., Sobue, T., Kobayashi, M., Sasaki, S., and Tsugane, S. (2003). Soy, isoflavones, and breast cancer risk in Japan. *J. Natl Cancer Inst.* 95, 906–913. doi: 10.1093/jnci/95.12.906

Yan, J., Wang, B., Zhong, Y., Yao, L., Cheng, L., and Wu, T. (2015). The soybean R2R3 MYB transcription factor GmMYB100 negatively regulates plant flavonoid biosynthesis. *Plant Mol. Biol.* 89, 35–48. doi: 10.1007/s11103-015-0349-3

Yan, Z., Zhang, X., Li, C., Jiao, S., and Dong, W. (2017). Association between consumption of soy and risk of cardiovascular disease: a meta-analysis of observational studies. *Eur. J. Prev. Cardiol.* 24, 735–747. doi: 10.1177/2047487316686441

Yi, J., Derynck, M. R., Li, X., Telmer, P., Marsolais, F., and Dhaubhadel, S. (2010). A single-repeat MYB transcription factor, GmMYB176, regulates CHS8 gene expression and affects isoflavonoid biosynthesis in soybean. *Plant J.* 62, 1019–1034.

Yu, O., Jung, W. S., Shi, J., Croes, R. A., Fader, G. M., McGonigle, B., et al. (2000). Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant Physiol.* 124, 781–793. 10.1104/pp.124.2.781

Yu, O., and McGonigle, B. (2005). Metabolic engineering of isoflavone biosynthesis. *Adv. Agron.* 86, 147–190. doi: 10.1016/S0065-2113(05)86003-1

Yu, O., Shi, J., Hession, A. O., Maxwell, C. A., McGonigle, B., and Odell, C. A. (2003). Metabolic engineering to increase isoflavone biosynthesis in soybean seed. *Phytochemistry* 63, 753–763. doi: 10.1016/s0031-9422(03)00345-5

Zhang, P., Du, H., Wang, J., Pu, Y., Yang, C., Yan, R., et al. (2020). Multiplex CRISPR/Cas9-mediated metabolic engineering increases soya bean isoflavone content and resistance to soya bean mosaic virus. *Plant Biotechnol. J.* 18, 1384–1395. doi: 10.1111/pbi.13302

Zhao, M., Wang, T., Wu, P., Guo, W., Su, L., Wang, Y., et al. (2017). Isolation and characterization of GmMYBJ3, an R2R3-MYB transcription factor that affects isoflavonoids biosynthesis in soybean. *PLoS One.* 12:e0179990. doi: 10.1371/journal.pone.0179990

Zhou, Y., Huang, J. L., Zhang, X. L., Zhu, L. M., Wang, X. F., Guo, N., et al. (2018). Overexpression of chalcone isomerase (CHI) increases resistance against *Phytophthora sojae* in soybean. *J. Plant Biol.* 61, 309–319. doi: 10.1007/s12374-018-0017-7

Conflict of Interest: YJO was employed by the company Institute for Future Environmental Ecology Co., Ltd., Jeonju, South Korea.

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 June 2021
doi: 10.3389/fpls.2021.669315

Alternative Splicing of the Basic Helix–Loop–Helix Transcription Factor Gene *CmbHLH2* Affects Anthocyanin Biosynthesis in Ray Florets of Chrysanthemum (*Chrysanthemum morifolium*)

Sun-Hyung Lim[1], Da-Hye Kim[1,2], Jae-A. Jung[3] and Jong-Yeol Lee[2]**

[1]Division of Horticultural Biotechnology, School of Biotechnology, Hankyong National University, Anseong, South Korea, [2]National Academy of Agricultural Science, Rural Development Administration, Jeonju, South Korea, [3]Floriculture Research Division, National Institute of Horticultural & Herbal Science, Rural Development Administration, Wanju, South Korea

OPEN ACCESS

Edited by:
M. Carmen González-Mas, University of Valencia, Spain

Reviewed by:
Nick Albert, The New Zealand Institute for Plant and Food Research Ltd, New Zealand
Antje Feller, University of Tübingen, Germany

***Correspondence:**
Sun-Hyung Lim
limsh2@hknu.ac.kr
Jong-Yeol Lee
jy0820@korea.kr

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

Received: *18 February 2021*
Accepted: *17 May 2021*
Published: *10 June 2021*

Citation:
Lim S-H, Kim D-H, Jung J-A and Lee J-Y (2021) Alternative Splicing of the Basic Helix–Loop–Helix Transcription Factor Gene CmbHLH2 Affects Anthocyanin Biosynthesis in Ray Florets of Chrysanthemum (Chrysanthemum morifolium). Front. Plant Sci. 12:669315. doi: 10.3389/fpls.2021.669315

Chrysanthemum is an important ornamental crop worldwide. Some white-flowered chrysanthemum cultivars produce red ray florets under natural cultivation conditions, but little is known about how this occurs. We compared the expression of anthocyanin biosynthetic and transcription factor genes between white ray florets and those that turned red based on cultivation conditions to comprehend the underlying mechanism. Significant differences in the expression of *CmbHLH2* were detected between the florets of different colors. *CmbHLH2* generated two alternatively spliced transcripts, designated *CmbHLH2*^{Full} and *CmbHLH2*^{Short}. Compared with *CmbHLH2*^{Full}, *CmbHLH2*^{Short} encoded a truncated protein with only a partial MYB-interaction region and no other domains normally present in the full-length protein. Unlike the full-length form, the splicing variant protein CmbHLH2^{Short} localized to the cytoplasm and the nucleus and could not interact with CmMYB6. Additionally, CmbHLH2^{Short} failed to activate anthocyanin biosynthetic genes and induce pigment accumulation in transiently transfected tobacco leaves, whereas CmbHLH2^{Full} promoted both processes when simultaneously expressed with CmMYB6. Co-expressing CmbHLH2^{Full} and CmMYB6 also enhanced the promoter activities of *CmCHS* and *CmDFR*. Notably, the Arabidopsis *tt8-1* mutant, which lacks red pigmentation in the leaves and seeds, could be complemented by the heterologous expression of CmbHLH2^{Full}, which restored red pigmentation and resulted in red pigmentation in high anthocyanin and proanthocyanidin contents in the leaves and seeds, respectively, whereas expression of CmbHLH2^{Short} did not. Together, these results indicate that CmbHLH2 and CmMYB6 interaction plays a key role in the anthocyanin pigmentation changes of ray florets in chrysanthemum. Our findings highlight alternative splicing as a potential approach to modulate anthocyanin biosynthesis in specific tissues.

Keywords: alternative splicing, anthocyanin, basic helix–loop–helix, chrysanthemum, MBW complex

INTRODUCTION

Chrysanthemum, one of the most popular ornamental plants worldwide, is an important floriculture crop. Flower color is a key trait for its commercial value, with a differential accumulation of secondary metabolites determining ray floret color. Whereas white ray florets in chrysanthemum accumulate flavone derivatives such as apigenins and luteolins but not anthocyanins, pink and red ray florets accumulate the cyanidin derivatives anthocyanins (Chen et al., 2012; Brugliera et al., 2013; Park et al., 2015).

Anthocyanins play an important role in attracting pollinators and seed dispersers due to their bright colors (Harborne and Williams, 2000; Hoballah et al., 2007). In addition, anthocyanins enhance plant survival under several biotic and abiotic stresses (Chalker-Scott, 1999; Ahmed et al., 2014), which could help plants better adapt to climate change. Anthocyanin accumulation is usually induced under adverse environmental conditions such as high light, nutrient depletion, and low temperature, suggesting that anthocyanins protect cells by scavenging free radicals (Nakabayashi et al., 2014; Schulz et al., 2016). Under natural cultivation conditions, the white ray florets of chrysanthemum flowers often turn red due to anthocyanin accumulation, depending on the genetic background (Xiang et al., 2019).

Extensive studies exploring the molecular mechanism of anthocyanin biosynthesis have led to the isolation and functional verification of regulators that promote anthocyanin biosynthesis in several plant species (Grotewold, 2006; Lin-Wang et al., 2010; Lim and Ha, 2013). A ternary complex composed of MYB–basic helix–loop–helix (bHLH)–WD40 (MBW) family members functions as the major transcriptional regulator of flavonoid biosynthetic genes (Hichri et al., 2011; Xu et al., 2015). In particular, subgroup IIIf bHLH proteins involved in flavonoid biosynthesis share common features, including an MYB-interaction region (MIR) at the N-terminus, a WD40-interacting region in the acidic domain (WD/AD), a bHLH domain, and a C-terminal region, which mediates the formation of homodimers or heterodimers with other bHLH proteins (Heim et al., 2003; Feller et al., 2006). Furthermore, mutations in the coding sequences of bHLH proteins, including PhAN1, IpIVS, and BnTT8, lead to reduced expression of flavonoid biosynthetic genes, thereby resulting in little or no pigment accumulation (Spelt et al., 2002; Park et al., 2007; Chen et al., 2012). These findings suggest that bHLH proteins play crucial roles in anthocyanin biosynthesis by modulating interactions with protein partners such as MYB and/or WD40 proteins.

The MYB and bHLH transcription factors (TFs) CmMYB6 and CmbHLH2 promote anthocyanin biosynthesis in chrysanthemum flowers (Liu et al., 2015; Xiang et al., 2015). In addition, the expression of the recently identified R3 MYB repressor gene *CmMYB#7* is negatively correlated with anthocyanin biosynthesis (Xiang et al., 2019). Thus, CmMYB#7 might inhibit anthocyanin biosynthesis by interacting with CmbHLH2. However, it does not directly regulate the expression of anthocyanin biosynthetic genes, suggesting that competition between CmMYB#7 and CmMYB6 for binding to CmbHLH2 modulates anthocyanin biosynthesis. Therefore, the regulation of anthocyanin biosynthesis appears to be highly dependent on the presence of a partner capable of interacting with CmbHLH2.

Here, we isolated the *CmbHLH2* and *CmMYB6* TF genes and investigated their roles in anthocyanin biosynthesis in the white-flowered chrysanthemum cultivar "OhBlang" (OB), which produces red ray florets at low temperatures. Notably, we observed that the *CmbHLH2* transcript is alternatively spliced. We expressed these transcripts transiently in tobacco leaves and stably in Arabidopsis plants to examine the roles of alternative splicing forms of CmbHLH2 in anthocyanin biosynthesis. Functional characterization of CmbHLH2 variants ($CmbHLH2^{Full}$ and $CmbHLH2^{Short}$) revealed that $CmbHLH2^{Short}$ could not initiate anthocyanin biosynthesis because it cannot interact with CmMYB6. Our identification and analysis of splicing variants of *CmbHLH2* thus increase our understanding of the mechanism underlying anthocyanin pigmentation in the chrysanthemum cultivar OB.

MATERIALS AND METHODS

Plant Materials

Chrysanthemum (*Chrysanthemum morifolium* Ramat.) cultivar "OhBlang (OB)" was used in this study. This cultivar produced either white or red flowers cultivated in a greenhouse from October 2018 to March 2020 at the National Institute of Horticultural and Herbal Science (Wanju, Korea). Chrysanthemums cv. OB that grown in greenhouses at 17–25°C on the timing of flower bud emergence exhibited the fully white-flowered (WO). However, chrysanthemums cv. OB was cultivated below 15°C for 2–3 weeks on the timing of flower bud emergence. It turned pigmentation in the ray florets. We called it turning red OB, TRO. Both white OB (WO) and turning red OB (TRO) flowers were collected at the full-bloom stage. Ray florets were collected from WO and TRO flowers, frozen in liquid nitrogen, and stored at −80°C. Gene expression in the ray florets was analyzed by quantitative PCR (qPCR) and reverse-transcription PCR (RT-PCR).

Transformation experiments were conducted using the *Arabidopsis thaliana transparent testa 8-1* (*tt8-1*) mutant (SALK_030966), which was obtained from the Arabidopsis Biological Resource Center (ABRC). All plants were grown on Murashige and Skoog (MS) medium or soil under long-day conditions (LD, 16-h light/8-h dark, 100 μmol m^{-2} s^{-1}) at 22°C. In addition, transient expression experiments were conducted using tobacco (*Nicotiana tabacum*) plants grown in a greenhouse under natural light at 28°C.

RNA Extraction, cDNA Synthesis, and Genomic DNA Isolation

Total RNA was extracted from ray florets using Fruit-mate for RNA Purification solution (Takara, Otsu, Japan) and Plant RNA Purification Reagent (Invitrogen, Carlsbad, CA, United States) as described previously (Lim et al., 2020) and purified using a FavorPrep™ Plant Total RNA Mini Kit (Favorgen, Changzhi, Taiwan). Total RNA was prepared from tobacco leaves using TRIzol reagent (Invitrogen) and purified using a

FavorPrep™ Plant Total RNA Mini Kit (Favorgen), according to the manufacturer's instructions. DNA contamination was removed by DNase I treatment (Ambion, Thermo Fisher Scientific), and the first-strand cDNA was synthesized from 2 μg of total RNA using amfiRivert cDNA Synthesis Platinum Master Mix (GenDEPOT, Barker, TX, United States).

Genomic DNA was obtained from chrysanthemum leaves using a Plant Mini Kit (Qiagen, Valencia, CA, United States) according to the manufacturer's instructions.

Gene Cloning and Sequence Analysis

The full-length open reading frames (ORFs) of *CmbHLH2* and *CmMYB6* were amplified from cDNA and genomic DNA by PCR with PrimeSTAR® HS DNA Polymerase (Takara) using primer sets (*CmbHLH2* and *CmMYB* F/R) listed in **Supplementary Table S1**. In addition, all PCR fragments were subcloned into the pENTR/D-TOPO vector (Invitrogen) to validate their DNA sequences.

The nucleotide sequences deduced amino acid sequences, and ORFs of *CmbHLH2* and *CmMYB6* were analyzed online.[1] Structural analysis of the deduced protein was carried out using the ExPASy Molecular Biology Server.[2] Multiple sequence alignments were generated using the CLUSTALW program.[3] A phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei, 1987) with MEGA version 4 software (Kumar et al., 2001).

qPCR and RT-PCR Analysis

The expression levels of specific genes were quantified by qPCR using AccuPower 2x Greenstar qPCR Master Mix (Bioneer, Daejun, Korea) and the BioRad CFX96 Detection System (BioRad Laboratories, Hercules, CA, United States) according to the manufacturer's instructions. Gene expression was normalized to *Elongation factor 1α* (*CmEF1α*) and *Glyceraldehyde 3-phosphate dehydrogenase* (*NtGAPDH*) expression for chrysanthemum and tobacco, respectively, as internal references. Three biological replicates were performed for each sample. The primers used for qPCR analysis are listed in **Supplementary Table S1**.

To validate the alternative splicing of *CmbHLH2*, RT-PCR was performed with specific primers listed in **Supplementary Table S1**. The resulting DNA fragments were separated on a 1.5% agarose gel with ethidium bromide for visualization.

Measuring Total Anthocyanin Content

Total anthocyanin content was measured in the samples as described by Lim et al. (2020). Briefly, powdered ray florets from WO and TRO and tobacco leaf samples were incubated in 600 μl extraction buffer (methanol containing 1% HCl) for 6 h at 4°C with moderate shaking. After adding 200 μl water and 200 μl chloroform, the samples were centrifuged at 14,000 *g* for 5 min at 4°C to sediment the plant material. The absorbance of the supernatant was recorded at 530 nm (A_{530}) and 657 nm (A_{657}) using a microplate reader. Anthocyanin content was determined using the following equation: $A_{530} - 0.25 \times A_{657}$. The anthocyanin content in each sample was measured in three independent experiments.

Subcellular Localization Assay

The subcellular localization of *CmbHLH2*^{Full}, *CmbHLH2*^{Short}, and *CmMYB6* was analyzed in Arabidopsis protoplasts as described by Yoo et al. (2007). GFP fusion constructs were generated in the p326-sGFP plasmid, which contains the cauliflower mosaic virus 35S (CaMV 35s) promoter. For the C-terminal GFP fusion, the ORFs of *CmbHLH2*^{Full}, *CmbHLH2*^{Short}, and *CmMYB6* were individually amplified using gene-specific primer sets (p326-*CmbHLH2*^{Full}-F/R, p326-*CmbHLH2*^{Short}-F/R, and p326-*CmMYB6*-F/R), which introduced an *Xba*I site upstream of the ATG codon, using the InFusion Cloning System (Clontech). The resulting p326-*CmbHLH*^{Full}-sGFP, p326-*CmbHLH2*^{Short}-sGFP, and p326-*CmMYB*-sGFP plasmids were sequenced to confirm the absence of errors during PCR amplification. The plasmids were introduced into Arabidopsis protoplasts prepared from leaf tissues by polyethylene glycol-mediated transformation. Fusion construct expression was monitored 16–20 h after transformation, and images were captured by fluorescence confocal microscopy (Leica TCS SP8, Leica Microsystems, Germany).

Transactivation and Yeast Two-Hybrid Assays

To generate the CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6 BD constructs, sequences encoding complete and partial regions of CmbHLH2^{Full} and CmMYB6 were individually amplified using specific primer sets (**Supplementary Table S1**). The amplified fragments were cloned into pGBKT7, harboring the GAL4 DNA-binding domain (BD) using an In-Fusion Cloning System (Clontech). The individual BD constructs were transformed into yeast strain AH109 following the manufacturer's instructions (Takara). The transformed yeast cells were grown on SD medium lacking Trp and replicated on SD medium lacking Trp, His, and Ade containing X-α-gal for color development. After 2 days of incubation in the dark at 30°C, the plates were photographed.

To examine the interactions of CmMYB6 with CmbHLH2^{Full} and CmbHLH2^{Short}, three CmMYB6 BD constructs (CmMYB6N1, CmMYB6N2, and CmMYB6N3) were selected, which showed no autoactivation activity in yeast. To validate the protein interactions, complete and partial regions of the CmbHLH2^{Full} coding sequence were individually cloned into pGADT7 harboring the GAL4 AD. The AD and BD constructs were co-transformed into yeast strain MaV203 following the manufacturer's instructions (Takara). Yeast strains were selected on SD medium lacking Trp and Leu. They were replicated on SD medium lacking Trp, Leu, and His supplemented with 10 mM 3-AT, a competitive inhibitor of the *HIS3* gene product. After 2 days of incubation in the dark at 30°C, the plates were photographed.

Promoter Activation Assay

The promoter regions of *CmCHS* and *CmDFR* were individually amplified using specific primer sets listed in

[1]http://www.ncbi.nlm.nih.gov
[2]http://cn.expasy.org/tools/
[3]https://www.genome.jp/tools-bin/clustalw

Supplementary Table S1. The resultant PCR products were purified and cloned into the pENTR/D-TOPO vector (Invitrogen) for sequencing.

The reporter fusion construct was generated by inserting the *CmCHS* and *CmDFR* promoter into the pTr-GUS vector at the 5' end of the *GUS* gene (derived from pBI121) after removing the CaMV 35S promoter region. The constructs were transferred into plant expression vector pBAR as described by Lim et al. (2017) and used as reporter constructs. The ORFs of *CmbHLH2*^{Full}, *CmbHLH2*^{Short}, and *CmMYB6* were individually subcloned into the pENTR/D-TOPO vector (Invitrogen) and incorporated into Gateway destination vector pB2GW7 (VIB-Ghent University, Ghent, Belgium) using several Gateway cloning steps. The pB2GW7-*CmbHLH*^{Full} and pB2GW7-*CmbHLH*^{Short} constructs were used as effector constructs. All constructs were transformed into Agrobacterium strain GV3101.

Transient promoter activation assays were performed in *N. tabacum* as previously described (Lim et al., 2017). Briefly, Agrobacteria containing reporter and effector constructs were cultured in LB medium for 2 days at 28°C, pelleted by centrifugation at 3,500 *g* for 5 min at 4°C, resuspended in infiltration buffer (10 mM $MgCl_2$ and 100 μM acetosyringone) to an OD_{600} of 0.2 (approximately 10 ml of buffer), and incubated at room temperature without shaking for 2 h. Before infiltration into tobacco leaves, Agrobacteria harboring effector and reporter constructs were mixed at a ratio of 3:1. Then, tobacco leaves were infiltrated with Agrobacteria harboring the genes of interest, and the leaves were harvested to assay GUS activity at 3 days post-infiltration (dpi). Agrobacteria harboring only the GUS reporter construct were used as the control. At least three biological replicates were used for each experiment.

In Planta Assay of *CmbHLH2* Function

The pB2GW7-*CmbHLH*^{Full} and pB2GW7-*CmbHLH*^{Short} constructs were individually transferred in Agrobacterium strain GV3101 and transformed into the Arabidopsis *tt8-1* mutant (SALK_030966) using the floral dip method. Transformed Arabidopsis seeds were grown in soil under 16-h light/8-h dark conditions at 20°C. Transgenic Arabidopsis plants were selected by spraying the plants with 0.3% Basta solution. Homozygous T_2 lines were selected and used for further analysis.

RESULTS

Regulatory and Structural Genes in Anthocyanin Biosynthesis Pathways Are Differentially Expressed in Chrysanthemum Ray Florets of Different Colors

We measured the anthocyanin contents in TRO and WO ray florets (**Figures 1A,B**). The anthocyanin contents were consistent with the visible appearance of the ray florets. However, only TRO ray florets appeared red. These results indicate that differences in anthocyanin contents are responsible for the color differences in TRO and WO ray florets.

To investigate the mechanism underlying ray floret pigmentation, we examined the transcript levels of genes from TRO and WO at the full-bloom stage (**Figure 1C**): four regulatory genes, including the R2R3-MYB-type TF gene *CmMYB6*, the bHLH-type TF gene *CmbHLH2*, the WD40 class TF gene *CmTTG1*, and the R3-MYB-type TF gene *CmMYB#7*; and seven structural genes involved in anthocyanin biosynthesis, including early biosynthetic genes (EBGs) encoding chalcone synthase (*CmCHS*), chalcone isomerase (*CmCHI*), flavanone 3-hydroxylase (*CmF3H*), and flavonoid 3' hydroxylase (*CmF3'H*) and late biosynthetic genes (LBGs) encoding dihydroflavonol 4-reductase (*CmDFR*), anthocyanidin synthase (*CmANS*), and UDP glucose: flavonoid-3-O-glucosyltransferase (*CmUFGT*).

CmBHLH2 and *CmMYB6* transcript levels were higher in TRO vs. WO ray florets, whereas *CmTTG1* transcript levels were similar in TRO and WO. Notably, the transcript level of *CmMYB#7* was negatively correlated with the transcript levels of active regulators, including *CmBHLH2* and *CmMYB6*. In addition, the transcript levels of all anthocyanin biosynthetic genes were significantly associated with the transcript levels of *CmBHLH2* and *CmMYB6*. These results indicate that the high expression levels of *CmBHLH2* and *CmMYB6* were positively correlated with EBG and LBG transcript levels and anthocyanin contents in TRO and WO ray florets.

Characterization of Regulatory Genes Involved in Anthocyanin Biosynthesis

We generated cDNA sequences of the open reading frames (ORFs) of CmBHLH2 and CmMYB6 to validate the difference in anthocyanin biosynthesis between white and red ray florets primers designed based on previously reported sequences (Liu et al., 2015; Xiang et al., 2015). The *CmbHLH2* cDNA sequences from TRO ray florets shared 100% identity with the previously reported sequence of *CmbHLH2* (GenBank accession number KT724056), a 1,842-bp sequence with an ORF encoding a deduced protein of 613 amino acids, which we refer to hereafter as a CmbHLH2^{Full} (GenBank accession number MW532125). However, the transcript length of *CmbHLH2* from WO ray florets was 1,207 bp with an ORF encoding a deduced protein of 149 amino acids due to the premature stop codon; this predicted protein was designated as CmbHLH2^{Short} (GenBank accession number MW532126), which harbored a large deletion at the C-terminal region (**Supplementary Figure S1**).

Phylogenetic analysis grouped CmbHLH2^{Full} and CmbHLH2^{Short} into the TT8 clade, whose members function in PA and anthocyanin biosynthesis (**Figure 2A**). Sequence alignment showed that CmbHLH2^{Full} contained well-conserved domains like those of other flavonoid-related bHLH TFs, including an N-terminal MYB-interaction region (MIR) domain, a WD/activation domain (AD), a basic helix–loop–helix (bHLH) domain, and a C-terminal aspartokinase, chorismate mutase, TyrA (ACT)-like domain (Feller et al., 2006; Lim et al., 2017). Furthermore, like other anthocyanin biosynthesis-related bHLH TFs, 19 amino acid residues were well conserved in the bHLH domain. However, CmbHLH2^{Short} contained a partial MIR domain due to alternative splicing, leading to a premature stop codon (**Supplementary Figure S1**). These results suggest that the generation of an alternative splicing

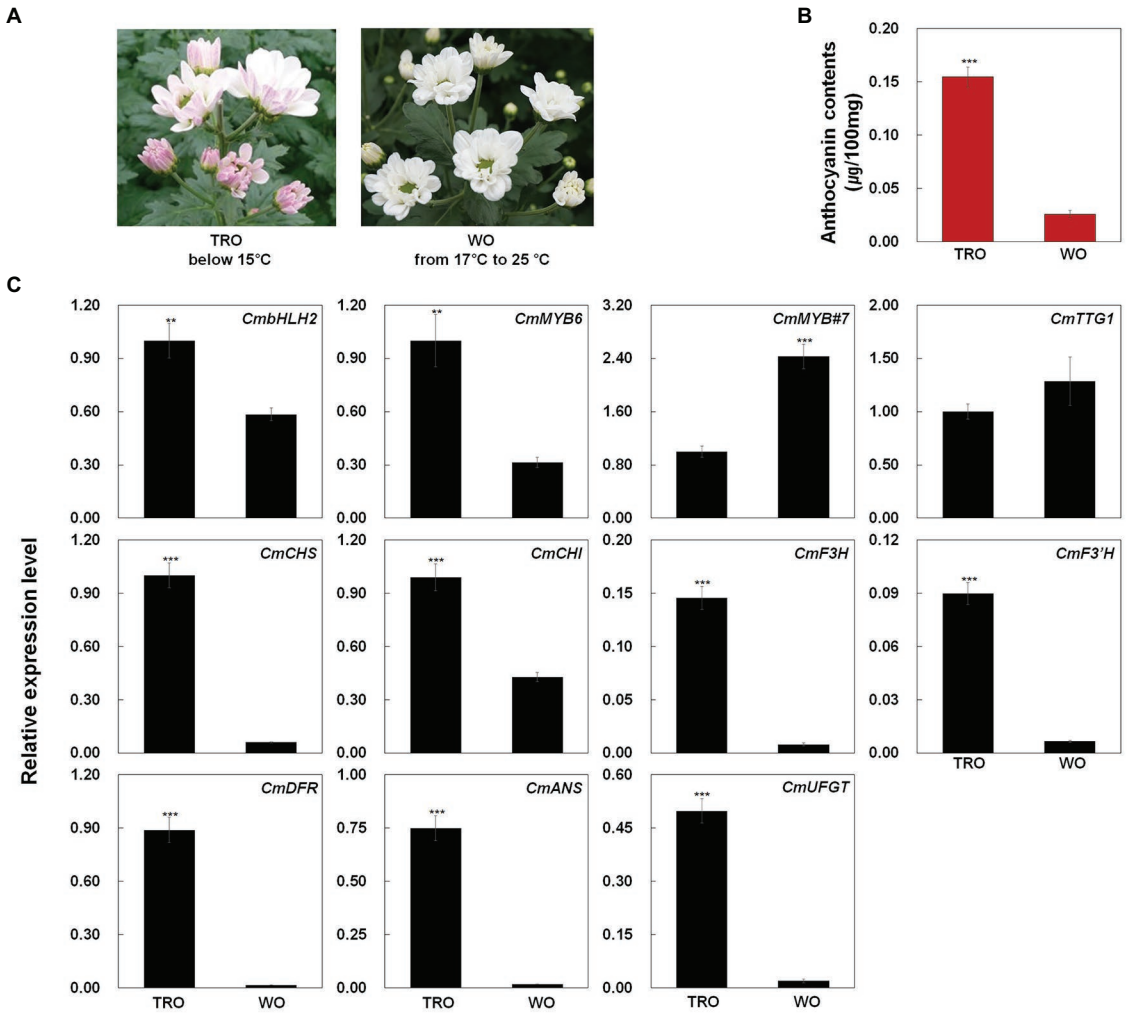

FIGURE 1 | Phenotypes, anthocyanin contents, and anthocyanin biosynthetic gene expression levels in turning red "OhBlang" (TRO) or white "OhBlang" (WO) ray florets. **(A)** Photographs of TRO and WO flowers at the full-bloom stage. **(B)** Anthocyanin level in TRO and WO ray florets. Error bars are the SE of three independent experiments. **(C)** Expression levels of anthocyanin biosynthetic genes and regulatory genes in TRO and WO ray florets. Results represent mean values ± SD from three biological replicates. *CmEF1*α was used as the reference gene. ** and *** indicate values that differ significantly from the control at $p < 0.01$ and $p < 0.001$, respectively, according to Student's paired *t*-test.

variant of *CmbHLH2* affects anthocyanin biosynthesis in the chrysanthemum cultivar "OhBlang."

Unlike *CmbHLH2* transcripts, *CmMYB6* transcripts isolated from TRO and WO ray florets shared 100% identity. *CmMYB6* was 765 bp, with an ORF encoding a deduced 254 amino acid. We designated this gene *CmMYB6* (GenBank accession number MW553912). *CmMYB6* encoded a protein with a one-amino acid difference in the R2 domain from the previously reported CmMYB6 protein (GenBank accession number KR002097). In a phylogenetic tree generated from R2R3 MYB proteins from various plant species (**Figure 2B**), CmMYB6 was grouped into the AN2 clade. Similar to previous findings, sequence alignment revealed that CmMYB6 shared the conserved R2R3 domains in the N-terminal region and the common KPRPR[S/T]F motif in the C-terminal region are conserved in all anthocyanin-promoting MYBs (**Supplementary Figure S2**).

Alternative Splicing of *CmbHLH2* in TRO and WO Ray Florets

To further explore the identity of the splicing variants derived from *CmbHLH2*, we amplified the genomic region of *CmbHLH2*, which was 4,951 bp. Multiple sequence alignment between *CmbHLH2* genomic DNA and the splicing variants revealed that *CmbHLH2*^{Full} consisted of 8 exons and 7 introns, but *CmbHLH2*^{Short} consisted of only 5 exons and 4 introns (**Figure 3A**; **Supplementary Figure S3**). Thus, the deduced CmbHLH2^{Full} polypeptide (predicted from the amplified genomic sequence of mature mRNA) was derived from the primary transcript of *CmbHLH2* containing splicing sites that were indeed recognized by the splicing machinery. By contrast, in *CmbHLH2*^{Short} pre-mRNA, an alternative 3' splicing sites in the third exon and 5' splicing sites in the sixth exon) were recognized by the spliceosome producing shorter transcripts.

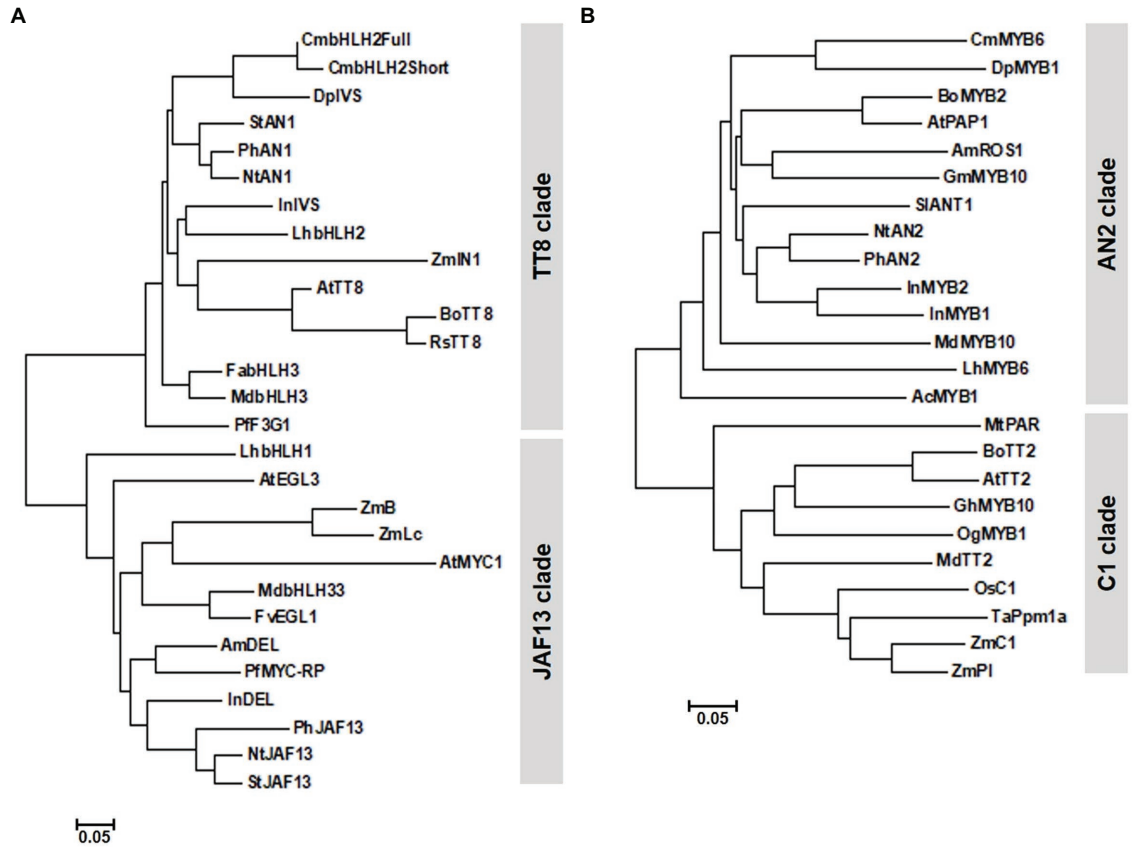

FIGURE 2 | Phylogenetic relationships of anthocyanin biosynthetic regulators in chrysanthemum and other species. **(A)** Phylogenetic tree of CmbHLH2 proteins from WO and TRO and subgroup IIIf bHLH proteins from other plants. **(B)** Phylogenetic tree of CmMYB6 and R2R3 MYB proteins from other plants. The phylogenetic tree was constructed using the neighbor-joining method with MEGA6 software.

Notably, in these transcripts, the loss of the fourth and fifth exons led to a frameshift, which produced an early stop codon at the beginning of the seventh exon. Thus, the corresponding proteins should show premature truncation, resulting in a much shorter protein than the theoretical one, with the loss of most residues downstream of the MIR domain.

To experimentally determine whether *CmbHLH2* undergoes alternative splicing, we performed RT-PCR using three sets of gene-specific primers targeting different regions of cDNA generated from RNA extracted from TRO and WO ray florets. The gene-specific primer sets (F1 + R1, F2 + R2, and F3 + R3) were designed from the first to second exon, from the fifth to seventh exon, and from the seventh to eighth exon, as shown in **Figure 3B**. RT-PCR assays with set 1 (F1 + R1) and set 3 (F3 + R3) primers showed no differences between TRO and WO ray florets, indicating that amplicons of the same size were generated (**Figure 3B**). However, RT-PCR assays with set 2 (F2 + R2) primers amplified products of the expected sizes in TRO ray florets but not in WO ray florets. These findings are consistent with cDNA data showing that the alternative donor site and alternative acceptor site result from the use of cryptic splice sites that might shorten an exon. In addition, these results indicate that *CmbHLH2* generates alternative transcripts in WO ray florets, resulting in the production of a nonfunctional protein with only a partial MIR domain.

Subcellular Localization Analysis of CmbHLH2 and CmMYB6

Since the subcellular localization of TFs in either the cytoplasm or nucleus regulates their biological activity, we assessed whether the abnormal subcellular distribution of the CmbHLH2 splicing variants could account for differences in anthocyanin biosynthesis. To address this hypothesis, we expressed individual GFP fusions of CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6 with a nucleus marker construct as a positive control (red fluorescent protein (RFP) fused to the nuclear localization signal (NLS) of the SV40 large T antigen) in Arabidopsis leaf protoplasts and examined their localization by immunofluorescence microscopy (**Figure 4A**). Protoplasts transfected with CmbHLH2^{Full}::GFP and CmMYB6::GFP showed strong GFP fluorescence in the nuclei, but those with CmbHLH2^{Short}::GFP showed GFP fluorescence in both the nuclei and cytoplasm (**Figure 4B**). These results indicate that the alternative splicing form CmbHLH2^{Short} exhibits an abnormal

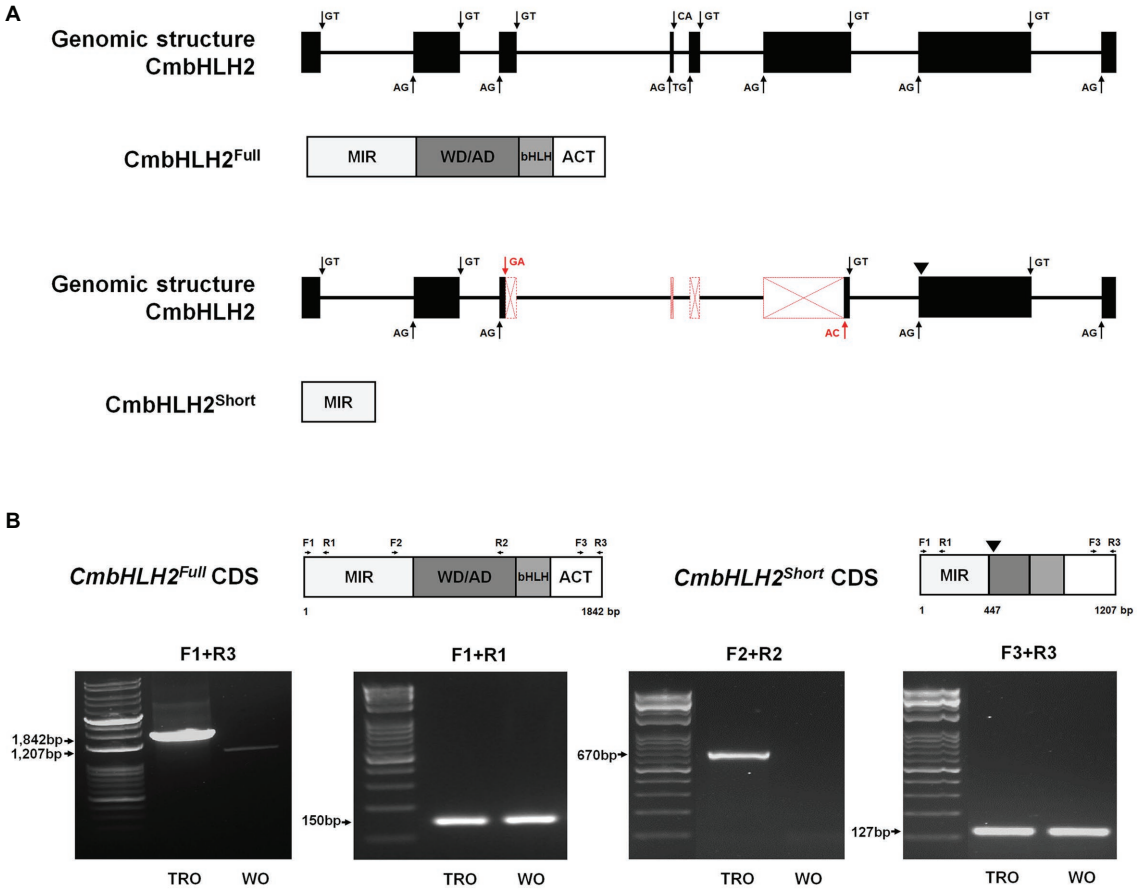

FIGURE 3 | Genomic structure and alternative splicing of *CmbHLH2*. **(A)** Diagram of the genomic structure of *CmbHLH2*. The *CmbHLH2*^{Full} and *CmbHLH2*^{Short} transcripts were generated by canonical splicing (black arrow) and alternative splicing (red arrow) of *CmbHLH2*. Protein from its mature mRNA is indicated below with functional domains. An inverted triangle indicates the premature stop codon. **(B)** The amplicons' representative gel image corresponds to *CmbHLH2*^{Full} and *CmbHLH2*^{Short} transcripts from TRO and WO ray florets. An inverted triangle indicates the premature stop codon. The primers used in this study are shown as arrows and listed in **Supplementary Table S1**.

subcellular distribution that could prevent or reduce its interaction with CmMYB6.

Transcriptional Activation Activity of CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6

To determine whether CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6, have transcriptional activation activity, we generated constructs harboring their gene sequences fused in-frame to the GAL4 DNA-binding domain (GAL4-BD) of the pGBKT7 vector. We assayed their transcriptional activation activity in yeast (**Figures 5A,B**). We transformed these constructs into yeast strain AH109 and screened the cells on synthetic dropout (SD) medium lacking tryptophan (SD/-T) and triple dropout medium, i.e., SD medium lacking tryptophan, histidine, and adenine (SD/-THA) supplemented with 10 mM 3AT and X-α-Gal.

Among the full and partially truncated clones of CmbHLH2^{Full} and CmbHLH2^{Short}, yeast transformants containing CmbHLH2_M, CmbHLH2_N, and CmbHLH2^{Short}, which lacked the C-terminal region of CmbHLH2, grew on selection medium and turned blue under the same conditions (**Figure 5C**). By contrast, yeast colonies (CmbHLH2^{Full} and CmbHLH2_C) containing the C-terminal region of CmbHLH2 failed to grow on the selection medium. These results suggest that the C-terminal region of CmbHLH2^{Full} represses protein autoactivation.

Among yeast transformants containing partial or full-length CmMYB6, yeast transformants containing CmMYB6_N1, CmMYB6_N2, and CmMYB6_N3, which lacked the C-terminal regions of the protein, failed to grow on selection medium and did not turn blue under the same conditions (**Figure 5D**). By contrast, yeast transformants containing CmMYB6_C, which lacked the N-terminal region of the protein, grew on selection medium, suggesting that the C-terminal region of CmMYB6 is required for its transcriptional activity. Taken together, these results indicate that the full-length CmMYB6 protein possesses strong autoactivation activity and that the C-terminus shows autoactivation activity, but the N-terminus does not (**Figure 5D**). Therefore, the autoactivation activity of CmMYB6 could be attributed to the C-terminal region.

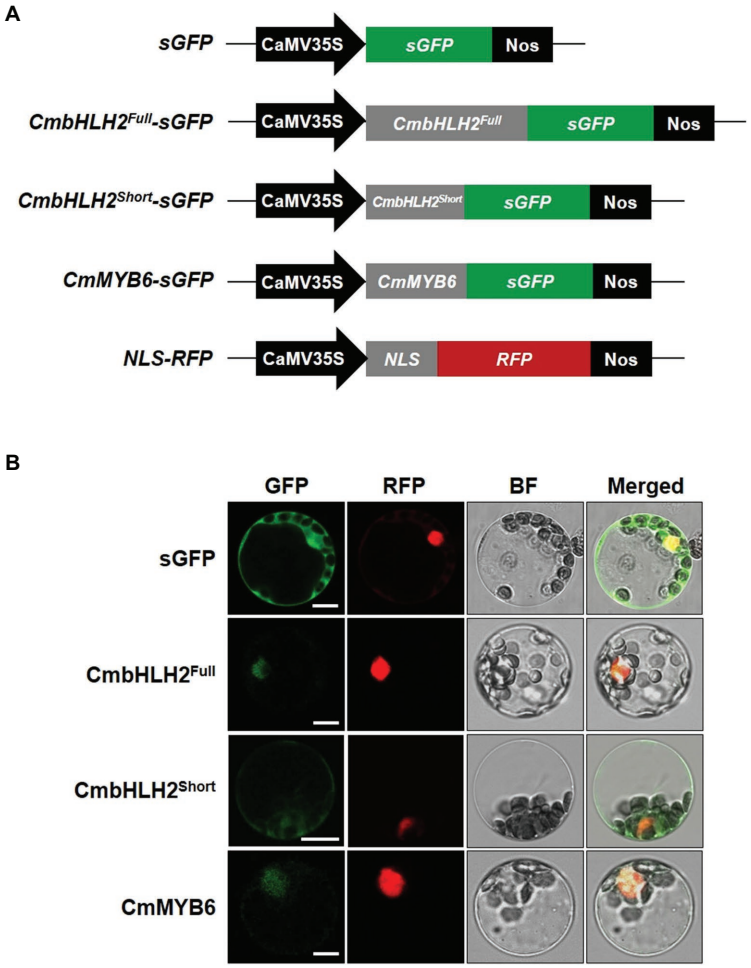

FIGURE 4 | Subcellular localization of CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6 in Arabidopsis leaf protoplasts. **(A)** Five constructs were used in this experiment: sGFP, soluble GFP; CmbHLH2^{Full}::GFP, CmbHLH2^{Full} fused to GFP; CmbHLH2^{Short}::GFP, CmbHLH2^{Short} fused to GFP; CmMYB6::GFP, CmMYB6 fused to GFP; and NLS::RFP, nuclear localization signal fused with RFP. **(B)** *In vivo* targeting of CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6 in Arabidopsis protoplasts. Data are representative of protoplasts expressing fusion proteins at 16 h after transformation. Bar = 10 μm.

Splicing Isoforms of CmbHLH2 Affect Its Interaction with CmMYB6

As shown in **Figure 1**, CmbHLH2 and CmMYB6 together regulate anthocyanin biosynthetic gene expression in chrysanthemum. To examine the interaction between CmbHLH2 and CmMYB6, we constructed bait vectors encoding full-length CmbHLH2^{Short} and complete and partially truncated CmbHLH2^{Full} fragments fused to pGADT7 harboring the GAL4 activation domain (GAL4-AD; **Figure 6A**) and performed a yeast two-hybrid (Y2H) assay (**Figure 6B**). As a result, CmbHLH2^{Full}_L strongly interacted with CmMYB6_N2 and CmMYB6_N3 but failed to interact with CmMYB6_N1, which harbored the R2 domain. Notably, CmbHLH2^{Full}_M and CmbHLH2^{Full}_N, which contained the common MIR domain, were able to interact with CmMYB6_N2 and CmMYB6_N3, but not CmMYB6_N1, like the results obtained with CmbHLH2^{Full}_L. However, CmbHLH2^{Short}, which possessed the partial MIR domain, failed to interact with any proteins with different CmMYB6 regions. These results indicate that the intact MIR domain of CmbHLH2^{Full} and the R3 domain of CmMYB6 are indispensable for interacting with these two proteins.

CmbHLH2 and CmMYB6 Cooperatively Regulate *CmCHS* and *CmDFR* Promoter Activity

Various TFs positively and negatively regulate anthocyanin biosynthetic genes. The MYB-recognizing element (MRE) and bHLH-recognizing element (BRE) at the proximal promoter regions of anthocyanin biosynthetic genes are thought to be the targets of MYB and bHLH TFs, respectively (Zhu et al., 2015). To test the activating effects of these TFs on their target genes, we performed a transient transactivation activity assay using *CmCHS* and *CmDFR*, which encode enzymes that catalyze the committed step in flavonoid and anthocyanin biosynthesis, respectively, and contain several MREs and BREs in their promoters (**Figure 7A**; **Supplementary Figure S4**). CmbHLH2^{Full} and CmMYB6 were independently infiltrated or co-infiltrated

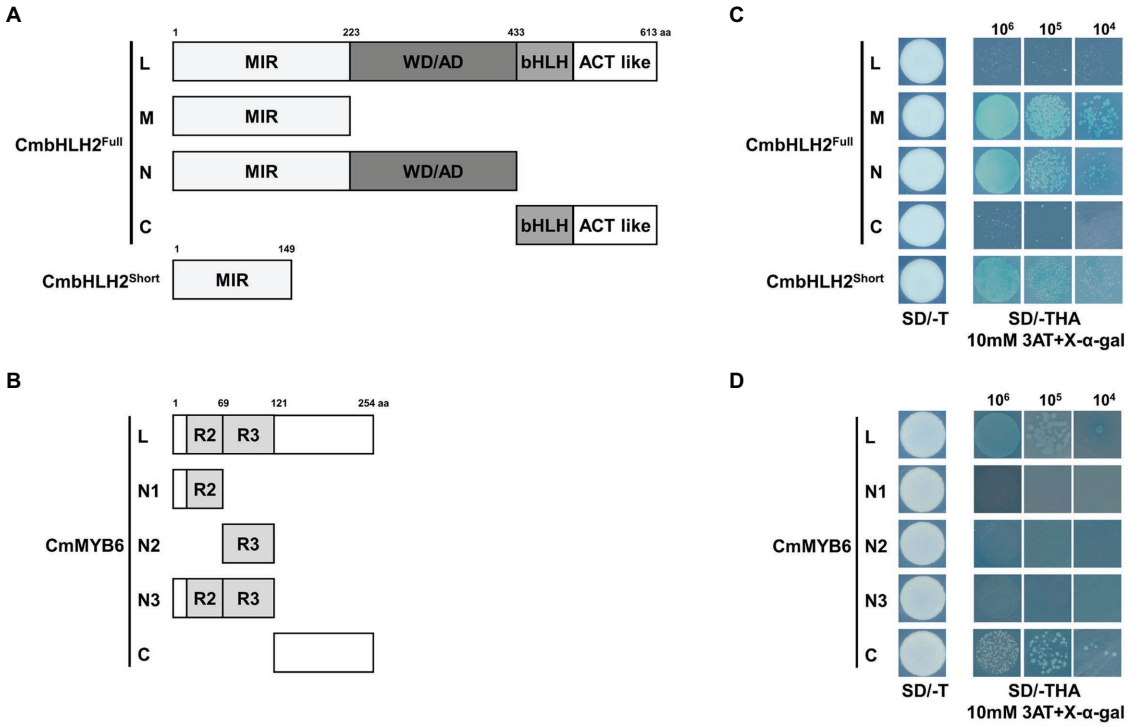

FIGURE 5 | Transactivation assay of CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6. **(A)** Schematic diagrams of CmbHLH2^{Full} and CmbHLH2^{Short} constructs used in the transcriptional activity assay. **(B)** Schematic diagrams of CmMYB6 constructs used in the transcriptional activity assay. **(C)** Transactivation assay of full-length CmbHLH2^{Full} and relevant truncated proteins and CmbHLH2^{Short} in yeast. **(D)** Transactivation analysis of full-length CmMYB6 and relevant truncated proteins in yeast. Amino acid positions are labeled in the diagrams. SD, minimal medium; AD, activation domain only; BD, binding domain only; 3AT, 3-amino-1,2,4-triazole; SD/−T, minimal medium lacking Trp; SD/−THA + X-α-Gal, minimal medium lacking Trp, His, and Ade but containing 20 mg/ml X-α-Gal; 10^6, 10^5, and 10^4 indicate the number of cells dotted on each plate of the medium.

into tobacco leaves along with modified pTr-GUS constructs containing the target *CmCHS* and *CmDFR* promoters driving *GUS* expression (**Figure 7B**). Neither CmbHLH2^{Full} nor CmMYB6 activated the *CmCHS* or *CmDFR* promoter when expressed separately in tobacco leaves. Furthermore, CmbHLH2^{Full} and CmMYB6 were co-infiltrated into tobacco leaves, GUS activity driven by the *CmCHS* and *CmDFR* promoters increased by approximately 5-fold and 26-fold, respectively. However, in a transient transactivation assay in which CmbHLH2^{Short} and CmMYB6 were expressed separately or together, the *CmCHS* and *CmDFR* promoters were not activated. These results suggest that CmbHLH2^{Short} and CmMYB6 cannot interact with each other, thus failing to activate the *CmCHS* and *CmDFR* promoters. Taken together, these results indicate that a transcriptional complex of CmbHLH2^{Full} and CmMYB6 activates the expression of the *CmCHS* and *CmDFR* promoters more effectively than CmbHLH2 or CmMYB6 alone.

Simultaneous Expression of *CmbHLH2* and *CmMYB6* Enhances Anthocyanin Accumulation in Tobacco Leaves

To further explore the role of CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6 in anthocyanin biosynthesis, we conducted a transient expression assay in tobacco leaves (**Figure 8A**). Similar to the promoter activation assay results, expressing CmbHLH2^{Full}, CmbHLH2^{Short}, or CmMYB6 alone failed to trigger anthocyanin accumulation in tobacco leaves. However, co-expressing CmbHLH2^{Full} and CmMYB6 resulted in red pigmentation in infiltrated tobacco leaves, whereas co-expressing CmbHLH2^{Short} and CmMYB6 did not. We detected red pigmentation at 5 days post-infiltration (dpi), which gradually increased until 7 dpi. We measured anthocyanin contents in leaf discs collected at 7 dpi. As expected, anthocyanin was barely detectible in leaves infiltrated separately with mock control, CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6, and in leaves co-infiltrated with CmbHLH2^{Short} and CmMYB6. By contrast, anthocyanin contents in leaves co-expressing CmbHLH2^{Full} and CmMYB6 was a 5-fold increment compared with those in leaves individually expressing CmbHLH2^{Full} or CmMYB6 (**Figure 8B**).

To further test the roles of CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6 as regulators of anthocyanin biosynthesis, we analyzed the transcript levels of nine structural genes involved in anthocyanin biosynthesis in infiltrated tobacco leaves by qPCR (**Figure 8C**). Like anthocyanin contents, the expression levels of anthocyanin biosynthetic genes were similar in leaves infiltrated separately with mock control, CmbHLH2^{Full}, CmbHLH2^{Short}, or CmMYB6 and in leaves co-infiltrated with CmbHLH2^{Short} and CmMYB6. However, the transcript levels of *NtCHS*, *NtDFR*, and *NtANS* were significantly higher in leaves co-expressing CmbHLH2^{Full} and CmMYB6 vs. the control.

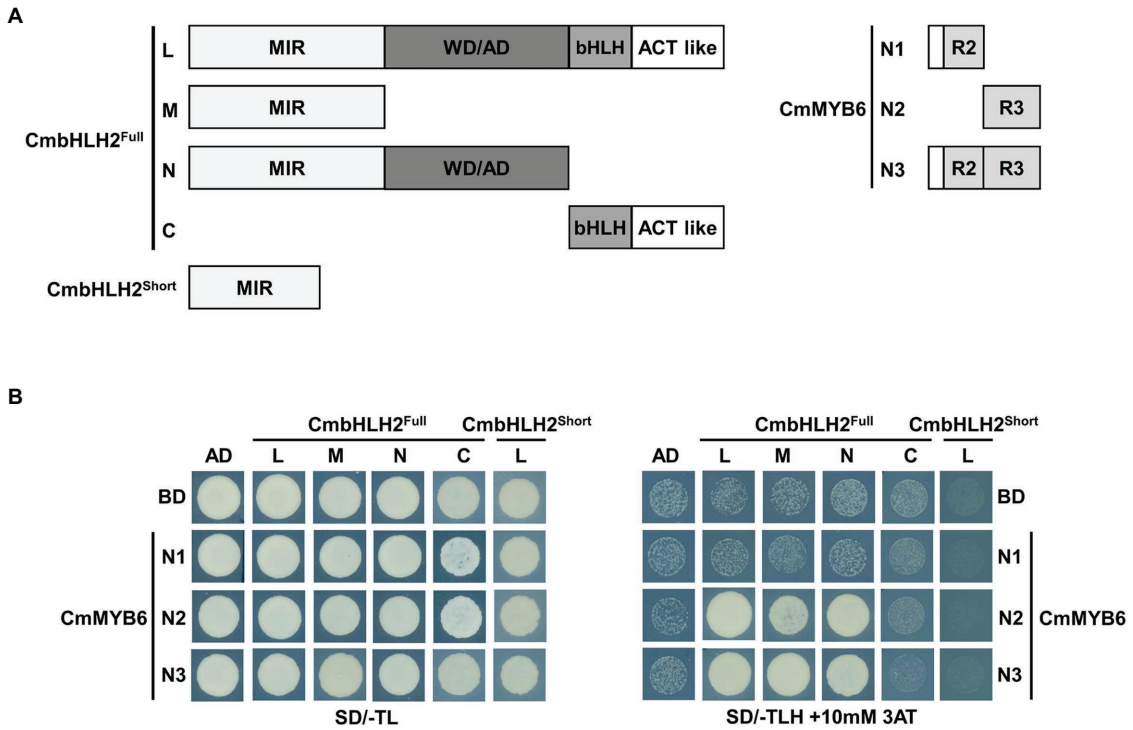

FIGURE 6 | Physical interactions among CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6. **(A)** Diagram of the constructs used in the Y2H experiment. The amino acid positions of the fragments are numbered. **(B)** Protein–protein interactions among CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6, as revealed by Y2H analysis. SD/-TL, minimal medium lacking Trp and Leu; SD/-TLH + 3AT, minimal medium lacking Trp, Leu, and His but containing 10-mM 3-amino-1,2,4-triazole (AT). AD and BD indicate the activation domain and binding domain, respectively.

In addition, the levels of these gene transcripts were consistent with the anthocyanin levels, and the significant red coloration observed in leaves simultaneously expressing CmbHLH2^{Full} and CmMYB6. Taken together, these results indicate that CmbHLH2^{Full} is essential for controlling anthocyanin biosynthesis together with CmMYB6 in chrysanthemum.

CmbHLH2 Restores Anthocyanin and Proanthocyanidin Biosynthesis in Arabidopsis *tt8-1* Mutant

As described above, distinct *CmbHLH2* transcripts were expressed in TRO and WO ray florets. Based on the results of Y2H, promoter activation, and transient expression assays, CmbHLH2^{Full} plays a crucial role in anthocyanin biosynthesis in cooperation with CmMYB6. To investigate the possible roles of CmbHLH2 in flavonoid biosynthesis in planta, we individually expressed *CmbHLH2^{Full}* and *CmbHLH2^{Short}* under the control of the CaMV 35S promoter in the Arabidopsis mutant *tt8-1*, which lacks anthocyanin in the junction between the stem and rosette leaves and fails to accumulate proanthocyanidin, resulting in yellow seeds (**Figure 9A**). Following selection with Basta, seeds of T_2 transgenic lines harboring CmbHLH2^{Full} were brown like wild-type Arabidopsis. By contrast, T_2 transgenic lines harboring CmbHLH2^{Short} still produced yellow seeds like those of *tt8-1*. We examined the anthocyanin levels in leaves of the transgenic lines harboring CmbHLH2^{Full} or CmbHLH2^{Short} at 4-week-old (**Figure 9B**). At that stage, we also performed the RT-PCR with primers designed to check for the presence and expression of *CmbHLH2*, determining that the *CmbHLH2* transgene was successfully expressed in all of the transgenic Arabidopsis plants (**Figure 9C**). Furthermore, anthocyanin levels in leaves from transgenic lines harboring CmbHLH2^{Full} were similar to those of the wild type, whereas anthocyanin levels were not restored to wild-type levels in plants harboring CmbHLH2^{Short}. These results demonstrate that the protein encoded by *CmbHLH2^{Full}* regulates proanthocyanidin biosynthesis in seeds and anthocyanin biosynthesis in leaves.

DISCUSSION

The biosynthesis of flavonoids, including anthocyanins and proanthocyanidins, is regulated by ternary MBW (MYB-bHLH-WDR) complexes composed of R2R3-MYB, bHLH, and WDR proteins (Hichri et al., 2011; Xu et al., 2015). Several studies have shown that the differences in anthocyanin levels in the same tissues of certain plant cultivars are mainly attributable to the MYB component of the MBW complex. For examples, in wild-type tomato, alternative splicing of the *SlAN2-like* gene (encoding an active R2R3 MYB) leads to the production of a nonfunctional protein due to the presence of a premature stop codon, resulting in a uniform red color in peels due to the

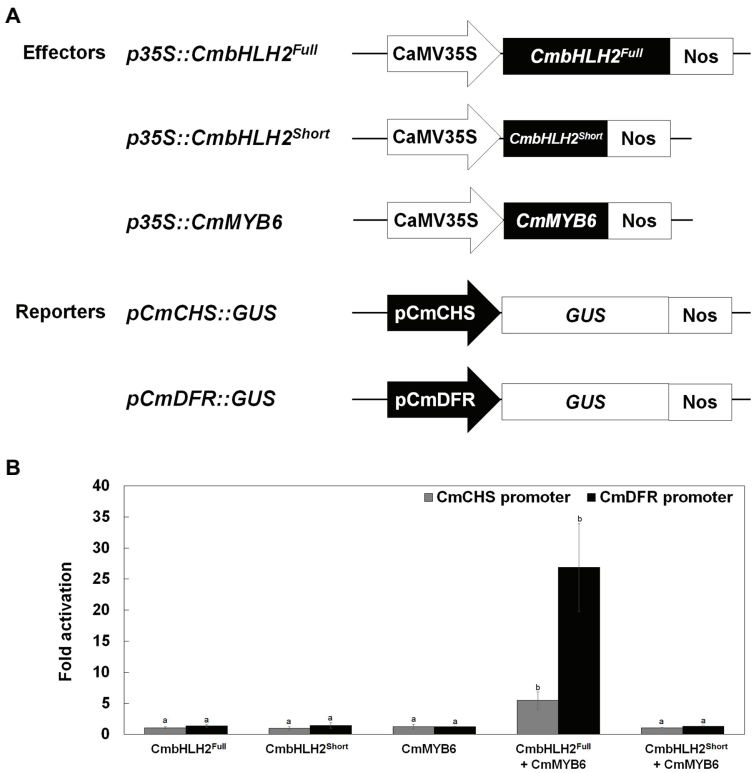

FIGURE 7 | Transcriptional activation assay of the roles of *CmbHLH2*^{Full}, *CmbHLH2*^{Short}, and *CmMYB6* in activating the *CmCHS* and *CmDFR* promoters. **(A)** Effector and reporter constructs used in the transcriptional activation assay. The effector construct contained *CmMYB6*, *CmbHLH2*^{Full}, and *CmbHLH2*^{Short} driven by the CaMV 35S promoter. The reporter constructs contained the *GUS* reporter gene driven by the *CmCHS* and *CmDFR* promoters. **(B)** The effects of regulatory proteins CmbHLH2^{Full}, CmbHLH2^{Short}, and CmMYB6 on *CmCHS* and *CmDFR* promoter activity. Results represent mean values ± SD from three biological replicates. Different letters above the bars indicate significantly different values ($p < 0.0001$) calculated using one-way ANOVA followed by Duncan's multiple range test.

accumulation of only carotenoids in this tissue (Colanero et al., 2020). However, the anthocyanin fruit (*aft*) gene of tomato cultivar Indigo Rose is a functional *SlAN2-like* gene, and mutation of SlMYBATV (a repressive R3 MYB), which acts as a negative repressor by interacting with the MBW complex, results in the accumulation of high levels of anthocyanin in peels (Cao et al., 2017). Similarly, Arabidopsis plants accumulate high levels of anthocyanin in leaves due to the low expression of *AtMYBL2*, the first reported R3 repressor, under both high light and high sucrose levels (Dubos et al., 2008; Matsui et al., 2008). Therefore, anthocyanin production is both positively and negatively regulated by MBW complex formation.

In the current study, we demonstrated that CmMYB6 and CmbHLH2 participate in the regulation of anthocyanin biosynthesis and that the gene products of *CmbHLH2* differ between TRO and WO ray florets. The *CmbHLH2* transcript has changed in splice sites resulting in premature stop codons in the CmbHLH2^{Short} alleles of WO ray florets. *CmbHLH2* transcript levels were correlated with early (EBG) and late (LBG) anthocyanin biosynthetic genes. Additionally, anthocyanin accumulation was only detected following the simultaneous expression of CmbHLH2^{Full} and CmMYB6, but not CmbHLH2^{Short} and CmMYB6 in a transient assay in tobacco leaves. Despite the presence of the *CmMYB6* gene product in TRO and WO, our findings indicate that the gene products of *CmbHLH2* play a crucial role in anthocyanin biosynthesis by functioning as key regulators of ray floret coloration.

Group IIIf bHLH proteins contain several domains, including the MIR, acidic, bHLH, and ACT-like domains, which function in trichome and anthocyanin biosynthesis (Hichri et al., 2011). In *A. thaliana*, the presence of IIIf bHLH proteins harboring only the N-terminal region (MIR and the acidic domains) was sufficient for trichome and non-hair root cell differentiation, but not for anthocyanin biosynthesis (Tominaga-Wada et al., 2012). *Anthocyanin*-less tomato contains a group IIIf bHLH TF with a single nucleotide polymorphism (G to T) that causes G184 to be converted to a stop codon, resulting in the loss of the bHLH domain and the ACT-like domain at the C-terminal region (Qiu et al., 2016). Our results suggest that, like in Arabidopsis and tomato, mutation of group IIIf bHLH proteins modulates anthocyanin biosynthesis in chrysanthemum.

Alternative splicing is when multiple mRNAs from the same gene are generated through the variable selection of splice sites during pre-mRNA splicing. It plays a key regulatory role in modulating gene expression during development and responding to environmental signals (Syed et al., 2012). Under low temperature, it revealed that variation in alternative splicing events might contribute to rapid changes in gene expression

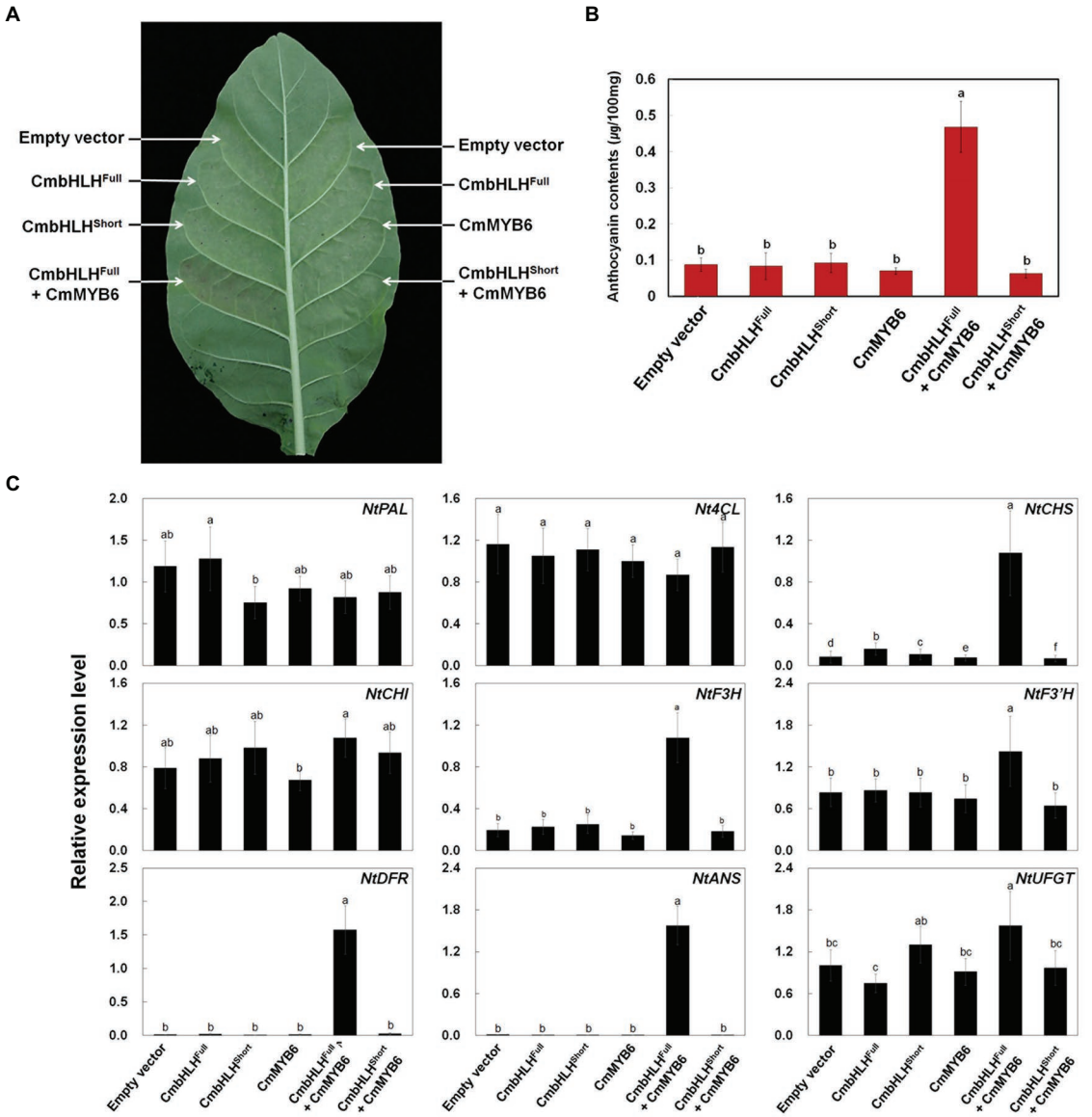

FIGURE 8 | Anthocyanin contents in transiently transformed tobacco leaves infiltrated with Agrobacterium strains carrying *CmMYB6*, *CmbHLH2*^{Full}, and *CmbHLH2*^{Short}. **(A)** Images of transiently transformed tobacco leaves at 5 days after agroinfiltration. **(B)** Anthocyanin content was extracted from each patch after agroinfiltration. **(C)** Relative expression levels of endogenous anthocyanin biosynthetic genes in tobacco plants, as determined by qPCR analysis. Results represent mean values ± SD from three biological replicates. Different letters above the bars indicate significantly different values ($p < 0.0001$) calculated using one-way ANOVA followed by Duncan's multiple range test.

and metabolite profile on sugar and antioxidant metabolism (Li et al., 2020). These changes may confer to improve adaptation in plants against environmental stress.

Alternative splicing of structural or TF genes related to anthocyanin biosynthesis affects pigment accumulation in plants. For example, in peach flowers, the *ANS* gene of plants with white flowers exhibits intron retention of a spare 129 bp sequence that is lacking in pink flowers, which leads to the dysfunction of the protein, ultimately blocking the accumulation of colored pigments in white petals (Yin et al., 2019). In tomato, splicing mutations of *R2R3 MYB SlAN2-like*, which control anthocyanin pigmentation, lead to the production of a dysfunctional protein (Colanero et al., 2020). In addition, alternative splicing of the group IIIf bHLH gene *BnTT8* was detected in the yellow seeds of rapeseed (Lin et al., 2020). Interestingly, a recent study reported that the weak allele of *Delila* (*del*^{23}), the group IIIf bHLH gene from snapdragon, generated by the alternative splicing due to the insertion in the intron located in the proximal region of the cryptic site resulted in reduced anthocyanin pigmentation in flower (Albert et al., 2021). These findings provide insight into why domesticated crops including tomato, rapeseed, peach, and snapdragon display

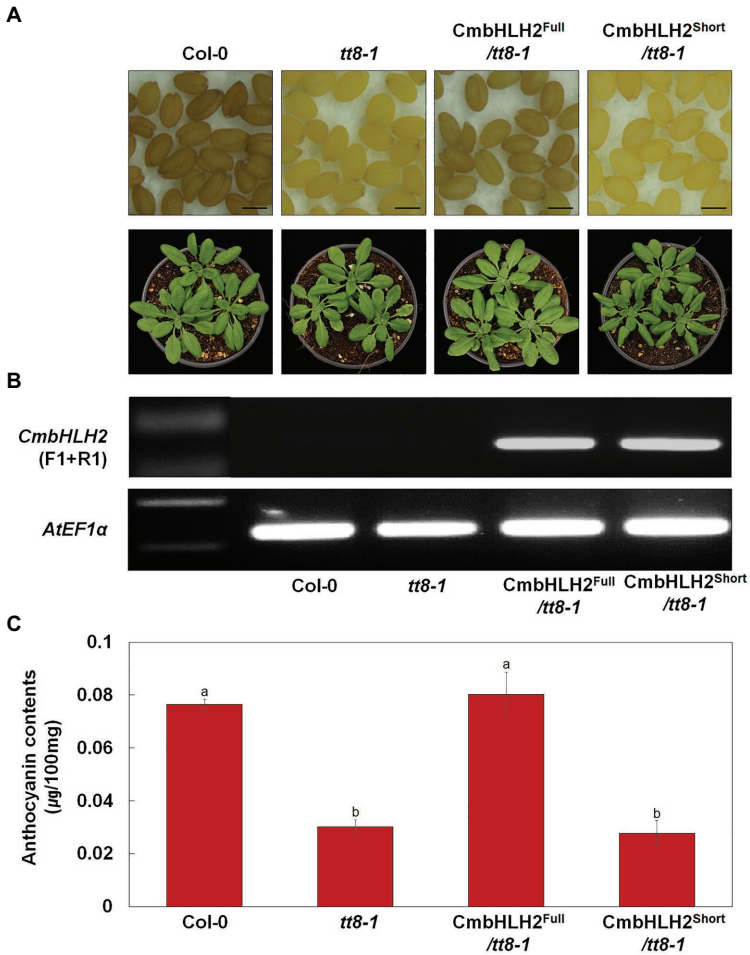

FIGURE 9 | Overexpression of *CmbHLH2*^{Full} and *CmbHLH2*^{Short} in Arabidopsis *tt8-1* mutants. **(A)** Seeds (top) and 4-week-old seedlings (bottom) of wild-type Col-0, *tt8-1*, and representative T_2 progeny of homozygous Arabidopsis *tt8-1* lines transformed with *CmbHLH2*^{Full} and *CmbHLH2*^{Short}, respectively. Scale bar indicates 0.1 mm. **(B)** Expression analysis of *CmbHLH2* in the leaves of wild-type, *tt8-1*, and transgenic plants by RT-PCR. *AtEF1*α was used as the reference gene. **(C)** Anthocyanin content of four-week-old Arabidopsis plants. Results represent mean values ± SD from three biological replicates. Different letters above the bars indicate significantly different values ($p < 0.0001$) calculated using one-way ANOVA followed by Duncan's multiple range test.

different pigment accumulation patterns in fruits, seeds, and flowers, respectively. In the current study, *CmbHLH2*^{Short}, an alternative splicing variant of *CmbHLH2*, was only detected in WO ray florets and generated a dysfunctional protein inactivation of the *CmCHS* and *CmDFR* promoters and the failure to produce anthocyanin pigments in tobacco leaves and Arabidopsis *tt8-1* plants. Taken together, these results point to an underlying regulatory mechanism for anthocyanin biosynthesis that involves the modification of anthocyanin biosynthetic or TF genes *via* alternative splicing.

This study demonstrated that different *CmbHLH2* transcripts encode truncated proteins lacking the acidic region, bHLH, and ACT domains in WO ray florets. Flavonoid-related bHLH2 TFs contain a MIR motif involved in interactions with the R3 domain of anthocyanin-promoting R2R3 MYB proteins. CmbHLH2^{Short}, with a partial MIR motif, was mislocalized to the nucleus and cytoplasm and lost the ability to bind to CmMYB6 (**Figures 4, 6**). As expected, the simultaneous expression of both CmbHLH2^{Short} and CmMYB6 did not induce anthocyanin accumulation or activate the promoters of *CmCHS* and *CmDFR*. These results suggest that the alternative splicing form of CmbHLH2 prevents the formation of MBW complexes that regulate the expression of anthocyanin biosynthetic genes and thus does not induce anthocyanin production.

Another possible explanation is that the alternative splicing form of *CmbHLH2* might significantly affect anthocyanin levels and anthocyanin biosynthetic gene expression by functioning as a repressor. Indeed, many studies have shown that truncated transcripts can interact with normal transcripts and function as repressors. For example, different alleles of *DvIVS*, a bHLH regulatory gene, affect the regulation of anthocyanin biosynthesis in dahlia (Ohno et al., 2011). The *DvIVS* transcript of the orange ray floret mutant *MJOr* encodes a full-length bHLH regulatory protein, leading to complete orange florets due to anthocyanin accumulation, whereas the *DvIVS* transcript of the yellow ray floret mutant *MJY* encodes a truncated protein

lacking the C-terminal region including the bHLH domain due to a Tdv1 insertion into the fourth intron, thereby leading to yellow florets. Morning glory produces different transcripts of *ItIVS*, a bHLH regulatory gene, due to an intragenic tandem duplication, leading to pale flower color (Park et al., 2007). Taken together, these results suggest that truncated bHLH protein not only impedes the conformation of the active MBW complex by directly interfering with the interaction between intact bHLH and MYB, but it also promotes the formation of an inactive MBW complex by replacing normal bHLH protein.

In addition, abnormal *bHLH* transcripts can have altered levels and stability. In dahlia and morning glory, the levels of the aberrant transcripts of *DvIVS* and *ItIVS* (containing premature termination codons) are very low compared with the normal transcripts of wild-type plants, suggesting that the aberrant transcripts might undergo nonsense-mediated decay, leading to degradation of the same mRNAs (Syed et al., 2012; Sibley et al., 2016). Indeed, CmbHLH2^{Short} has retained partial MIR domains but lacks DNA-binding domains and/or transcriptional regulatory domains and cannot interact with the R2R3 MYB CmMYB6. It is possible that the aberrant form competitively interacts with the full-length TF to inhibit its activity.

Overall, we conclude that the TF *CmbHLH2* plays an important role in anthocyanin biosynthesis, depending on the splicing form of its gene. Indeed, alternative splicing of the *CmbHLH2* gene produces at least two transcripts: one encoding a functional TF and an alternative form encoding a dysfunctional TF that negatively regulates the functional form (Seo et al., 2011). Thus, these findings provide insight into a molecular mechanism in which the alternative splicing of *CmbHLH2* modulates the protein interaction network involved in anthocyanin biosynthesis.

CONCLUSION

In this study, we demonstrated that anthocyanin biosynthesis in TRO and WO ray florets is determined by the presence of *CmbHLH2*, *CmbHLH2^{Full}*, and *CmbHLH2^{Short}* transcripts. In the ray florets of chrysanthemum flowers, anthocyanin contents and anthocyanin biosynthetic gene expression strongly depend on the splicing form of *CmbHLH2*. A transient expression assay in tobacco and a promoter activation assay indicated that co-expressing CmbHLH2^{Full} and CmMYB6 activated anthocyanin biosynthesis and the expression of *CmCHS* and *CmDFR* genes expressing CmbHLH2^{Short} and CmMYB6 did not. In addition, a complementation assay of Arabidopsis *tt8-1* revealed that CmbHLH2^{Full} functions in proanthocyanidin and anthocyanin biosynthesis CmbHLH2^{Short} fails to function in these processes. Taken together, these results suggest that the alternative splicing of *CmbHLH2* plays a key role in modulating anthocyanin biosynthesis in the ray florets of chrysanthemum flowers.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The name of the repository and accession numbers can be found at: National Center for Biotechnology information (NCBI) GenBank, https://www.ncbi.nlm.nih.gov/genbank/, MW532125 and MW532126.

AUTHOR CONTRIBUTIONS

S-HL and J-YL conceived, designed the research, and drafted the manuscript. D-HK and J-AJ conducted experiments. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the grants from the Next-Generation BioGreen 21 Linked Program (PJ015734), Rural Development Administration, and from Basic Science Research Program through the National Research Foundation of Korea (NRF-2020R1F1A1072559), the Ministry of Education, Korea.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.669315/full#supplementary-material

REFERENCES

Ahmed, N. U., Park, J. I., Jung, H. J., Yang, T. J., Hur, Y., and Nou, I. S. (2014). Characterization of dihydroflavonol 4-reductase (DFR) genes and their association with cold and freezing stress in *Brassica rapa*. *Gene* 550, 46–55. doi: 10.1016/j.gene.2014.08.013

Albert, N. W., Butelli, E., Moss, S. M. A., Piazza, P., Waite, C. N., Schwinn, K. E., et al. (2021). Discrete bHLH transcription factors play functionally overlapping roles in pigmentation patterning in flowers of *Antirrhinum majus*. *New Phytol.* 8, 1–15. doi: 10.1111/nph.17142

Brugliera, F., Tao, G. Q., Tems, U., Kalc, G., Mouradova, E., Price, K., et al. (2013). Violet/blue chrysanthemums—metabolic engineering of the anthocyanin biosynthetic pathway results in novel petal colors. *Plant Cell Physiol.* 54, 1696–1710. doi: 10.1093/pcp/pct110

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

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress responses. *Photochem. Photobiol.* 70, 1–9. doi: 10.1111/j.1751-1097.1999.tb01944.x

Chen, S. M., Li, C. H., Zhu, X. R., Deng, Y. M., Sun, W., Wang, L. S., et al. (2012). The identification of flavonoids and the expression of genes of anthocyanin biosynthesis in the chrysanthemum flowers. *Biol. Plant.* 56, 458–464. doi: 10.1007/s10535-012-0069-3

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

Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G., Lanet, E., Debeaujon, I., et al. (2008). MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *Plant J.* 55, 940–953. doi: 10.1111/j.1365-313X.2008.03564.x

Feller, A., Hernandez, J. M., and Grotewold, E. (2006). An ACT-like domain participates in the dimerization of several plant basic-helix-loop-helix transcription factors. *J. Biol. Chem.* 281, 28964–28974. doi: 10.1074/jbc.M603262200

Grotewold, E. (2006). The genetics and biochemistry of floral pigments. *Annu. Rev. Plant Biol.* 57, 761–780. doi: 10.1146/annurev.arplant.57.032905.105248

Harborne, J., and Williams, C. (2000). Advances in flavonoid research since 1992. *Phytochemistry* 55, 481–504. doi: 10.1016/S0031-9422(00)00235-1

Heim, M. A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., and Bailey, P. C. (2003). The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* 20, 735–747. doi: 10.1093/molbev/msg088

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J. Exp. Bot.* 62, 2465–2483. doi: 10.1093/jxb/erq442

Hoballah, M. E., Gübitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., et al. (2007). Single gene–mediated shift in pollinator attraction in Petunia. *Plant Cell* 19, 779–790. doi: 10.1105/tpc.106.048694

Kumar, S., Tamura, K., Jakobsen, I. B., and Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17, 1244–1245. doi: 10.1093/bioinformatics/17.12.1244

Li, Y., Mi, X., Zhao, S., Zhu, J., Guo, R., Xia, X., et al. (2020). Comprehensive profiling of alternative splicing landscape during cold acclimation in tea plant. *BMC Genomics* 21:65. doi: 10.1186/s12864-020-6491-6

Lim, S. H., and Ha, S. H. (2013). Marker development for identification of rice seed coat color. *Plant Biotechnol.* 7, 391–398. doi: 10.1007/s11816-013-0276-1

Lim, S. H., Kim, D. H., Kim, J. K., Lee, J. Y., and Ha, S. H. (2017). A radish basic helix-loop-helix transcription factor, RsTT8 acts a positive regulator for anthocyanin biosynthesis. *Front. Plant Sci.* 8:1917. doi: 10.3389/fpls.2017.01917

Lim, S. H., Park, B., Kim, D. H., Park, S., Yang, J. H., Jung, J. A., et al. (2020). Cloning and functional characterization of dihydroflavonol 4-reductase gene involved in anthocyanin biosynthesis of Chrysanthemum. *Int. J. Mol. Sci.* 21:7960. doi: 10.3390/ijms21217960

Lin, A., Ma, J., Xu, F., Xu, W., Jiang, H., Zhang, H., et al. (2020). Differences in alternative splicing between yellow and black-seeded rapeseed. *Plan. Theory* 9:977. doi: 10.3390/plants9080977

Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, A., McGhie, T. K., et al. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. *BMC Plant Biol.* 10:50. doi: 10.1186/1471-2229-10-50

Liu, X. F., Xiang, L. L., Yin, X. R., Grierson, D., Li, F., and Chen, K. S. (2015). The identification of a MYB transcription factor controlling anthocyanin biosynthesis regulation in Chrysanthemum flowers. *Sci. Hortic.* 194, 278–285. doi: 10.1016/j.scienta.2015.08.018

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

Nakabayashi, R., Yonekura-Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., et al. (2014). Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. *Plant J.* 77, 367–379. doi: 10.1111/tpj.12388

Ohno, S., Hosokawa, M., Hoshino, A., Kitamura, Y., Morita, Y., Park, K., et al. (2011). A bHLH transcription factor, DvIVS, is involved in regulation of anthocyanin synthesis in dahlia (Dahlia variabilis). *J. Exp. Bot.* 62, 5105–5116. doi: 10.1093/jxb/err216

Park, C. H., Chae, S. C., Park, S. Y., Kim, J. K., Kim, Y. J., Chung, S. O., et al. (2015). Anthocyanin and carotenoid contents in different cultivars of chrysanthemum (*Dendranthema grandiflorum* Ramat.) flower. *Molecules* 20, 11090–11102. doi: 10.3390/molecules200611090

Park, K. I., Ishikawa, N., Morita, Y., Choi, J. D., Hoshino, A., and Iida, S. (2007). A bHLH regulatory gene in the common morning glory, *Ipomoea purpurea*, controls anthocyanin biosynthesis in flowers, proanthocyanidin and phytomelanin pigmentation in seeds, and seed trichome formation. *Plant J.* 49, 641–654. doi: 10.1111/j.1365-313X.2006.02988.x

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

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

Schulz, E., Tohge, T., Zuther, E., Fernie, A. R., and Hincha, D. K. (2016). Flavonoids are determinants of freezing tolerance and cold acclimation in *Arabidopsis thaliana*. *Sci. Rep.* 6:34027. doi: 10.1038/srep34027

Seo, P., Kim, M., Ryu, J. Y., Jeong, E. Y., and Park, C. M. (2011). Two splice variants of the IDD14 transcription factor competitively form nonfunctional heterodimers which may regulate starch metabolism. *Nat. Commun.* 2:303. doi: 10.1038/ncomms1303

Sibley, C. R., Blazquez, L., and Ule, J. (2016). Lessons from non-canonical splicing. *Nat. Rev. Genet.* 17, 407–421. doi: 10.1038/nrg.2016.46

Spelt, C., Quattrocchio, F., Mol, J., and Koes, R. (2002). ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. *Plant Cell* 14, 2121–2135. doi: 10.1105/tpc.003772

Syed, N. H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J. W. (2012). Alternative splicing in plants-coming of age. *Trends Plant Sci.* 17, 616–623. doi: 10.1016/j.tplants.2012.06.001

Tominaga-Wada, R., Iwata, M., Nukumizu, Y., Sano, R., and Wada, T. (2012). A full-length R-like basic-helix-loop-helix transcription factor is required for anthocyanin upregulation whereas the N-terminal region regulates epidermal hair formation. *Plant Sci.* 183, 115–122. doi: 10.1016/j.plantsci.2011.11.010

Xiang, L. L., Liu, X. F., Li, X., Yin, X. R., Grierson, D., Li, F., et al. (2015). A novel bHLH transcription factor involved in regulating anthocyanin biosynthesis in chrysanthemums (*Chrysanthemum morifolium* Ramat.). *PLoS One* 10:e0143892. doi: 10.1371/journal.pone.0143892

Xiang, L., Liu, X., Li, H., Yin, X., Grierson, D., Li, F., et al. (2019). CmMYB#7, an R3 MYB transcription factor, acts as a negative regulator of anthocyanin biosynthesis in chrysanthemum. *J. Exp. Bot.* 70, 3111–3123. doi: 10.1093/jxb/erz121

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001

Yin, P., Zhen, Y., and Li, S. (2019). Identification and functional classification of differentially expressed proteins and insight into regulatory mechanism about flower color variegation in peach. *Acta Physiol. Plant.* 41:95. doi: 10.1007/s11738-019-2886-x

Yoo, S. D., Cho, Y. H., and Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2, 1565–1572. doi: 10.1038/nprot.2007.199

Zhu, Z., Wang, H., Wang, Y., Guan, S., Wang, F., Tang, J., et al. (2015). Characterization of the cis elements in the proximal promoter regions of the anthocyanin pathway genes reveals a common regulatory logic that governs pathway regulation. *J. Exp. Bot.* 66, 3775–3789. doi: 10.1093/jxb/erv173

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership