Plant Cell Reports
https://doi.org/10.1007/s00299-021-02818-x

ORIGINAL ARTICLE

Identification and characterization of unique 5-hydroxyisoflavonoid biosynthetic key enzyme genes in *Lupinus albus*

Jinyue Liu[1,2,3] · **Wenbo Jiang**[2]

Received: 27 October 2021 / Accepted: 19 November 2021

Abstract

Key message **5-Hydroxyisoflavonoids, no 5-deoxyisoflavonoids, in *Lupinus* species, are due to lack of CHRs and Type II CHIs, and the key enzymes of isoflavonoid biosynthetic pathway in white lupin were identified.**

Abstract White lupin (*Lupinus albus*) is used as food ingredients owing to rich protein, low starch, and rich bioactive compounds such as isoflavonoids. The isoflavonoids biosynthetic pathway in white lupin still remains unclear. In this study, only 5-hydroxyisoflavonoids, but no 5-deoxyisoflavonoids, were detected in white lupin and other *Lupinus* species. No 5-deoxyisoflavonoids in *Lupinus* species are due to lack of CHRs and Type II CHIs. We further found that the *CHI* gene cluster containing both Type I and Type II CHIs possibly arose after the divergence of *Lupinus* with other legume clade. LaCHI1 and LaCHI2 identified from white lupin metabolized naringenin chalcone to naringenin in yeast and tobacco (*Nicotiana benthamiana*), and were *bona fide* Type I CHIs. We further identified two isoflavone synthases (LaIFS1 and LaIFS2), catalyzing flavanone naringenin into isoflavone genistein and also catalyzing liquiritigenin into daidzein in yeast and tobacco. In addition, LaG6DT1 and LaG6DT2 prenylated genistein at the C-6 position into wighteone. Two glucosyltransferases LaUGT1 and LaUGT2 metabolized genistein and wighteone into its 7-*O*-glucosides. Taken together, our study not only revealed that exclusive 5-hydroxyisoflavonoids do exist in *Lupinus* species, but also identified key enzymes in the isoflavonoid biosynthetic pathway in white lupin.

Keywords CHI · Glucosyltransferase · IFS · Isoflavonoid · Prenyltransferase · White lupin

Introduction

Isoflavonoids constitute a large subfamily of flavonoids, playing critical roles in plant growth and development and plant physiology (Mierziak et al. 2014). They have been involved in several health-promoting effects, including reduced cancer risk and alleviating effects of hormone replacement therapy (Birt et al. 2001; Boue et al. 2011). Contrary to other subgroups of flavonoids, isoflavonoids are distributed predominantly in leguminous plants. It is sporadic for the occurrence of isoflavonoids in other plant families (Lapcik 2007; Reynaud et al. 2005).

Of the flavonoid biosynthesis pathway, chalcone synthase (CHS; EC 2.3.1.74) catalyzes the first committed step to form naringenin chalcone (6′-hydroxychalcone), by the stepwise condensation of three molecules of malonyl-CoA with one molecule of *p*-coumaroyl-CoA (Oguro et al. 2004). Meanwhile, chalcone reductase (CHR; EC 2.3.1.170) catalyzes the production of isoliquiritigenin (6′-deoxychalcone) through its effects on the CHS-catalyzed reaction, indicating that CHR is the essential enzyme that catalyzes the first committed step of 5-deoxyisoflavonoid biosynthesis (Bomati et al. 2005; Mameda et al. 2018; Welle and Grisebach 1988). As shown in Fig. S1, naringenin chalcone and isoliquiritigenin can be further metabolized into naringenin (5-hydroxyflavanone) and liquiritigenin (5-deoxyflavanone) by chalcone isomerase (CHI; EC 5.5.1.6), respectively (Jiang et al. 2015; Morita et al. 2014; Ralston et al. 2005; Shimada et al. 2003). The general distribution in higher plants suggests the

Communicated by Maike Petersen.

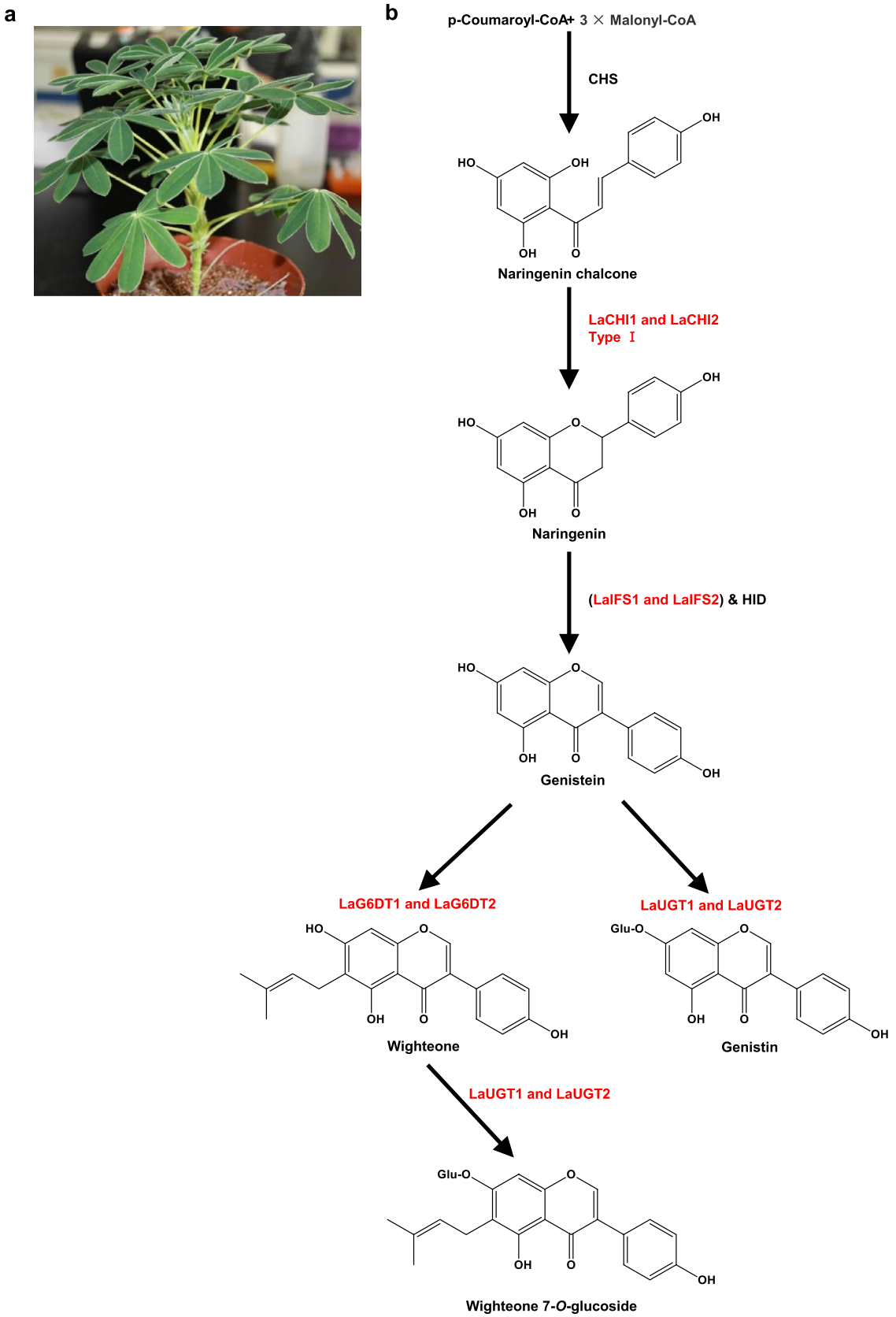
✉ Wenbo Jiang
jiangwenbo@caas.cn

[1] Center for Plant Biology, School of Life Sciences, Tsinghua University, Beijing 100084, China

[2] Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China

[3] Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang, Jiangxi 332900, China

Published online: 01 December 2021


a
b
p-Coumaroyl-CoA + 3 × Malonyl-CoA
CHS
Naringenin chalcone
LaCHI1 and LaCHI2
Type Ⅰ
Naringenin
(LaIFS1 and LaIFS2) & HID
Genistein
LaG6DT1 and LaG6DT2
LaUGT1 and LaUGT2
Wighteone
Genistin
LaUGT1 and LaUGT2
Wighteone 7-O-glucoside

◂**Fig. 1** The unique 5-hydroxyisoflavonoid pathway in white lupin. **a** The white lupin (*L. albus*) plants used in this study are showed. **b** White lupin unique 5-hydroxyisoflavonoids pathway. The enzymes characterized from white lupin, involved in 5-hydroxyisoflavonoids metabolism, are deposited in red (color figure online)

significance of CHI in the flavonoid and isoflavonoid pathway (Shimada et al. 2003). The plant CHI family can be classified into four subfamilies (Type I to Type IV) (Ralston et al. 2005). Type I and II proteins are *bona fide* catalysts having CHI enzymatic activity. Type I CHIs are ubiquitous in vascular plants and are responsible for plant flavonoids and isoflavonoids synthesis (Dixon et al. 1988; Ralston et al. 2005). Type II CHI enzymes appear to be legume-specific and are involved in isoflavonoid production (Ralston et al. 2005; Shimada et al. 2003). Type III CHIs are widely present in land plants and green algae, and those of *Arabidopsis thaliana* were recently shown to be fatty-acid-binding proteins involved in fatty acid metabolism in planta (Ngaki et al. 2012). Type IV CHIs are only found in land plants. Although AtCHIL, as the exclusive member of Type IV CHI proteins in *Arabidopsis*, promotes flavonoid production (Jiang et al. 2015); and loss-of-function mutation of Japanese morning glory (*Ipomoea nil*) Type IV CHI leads to low amounts of anthocyanin (Morita et al. 2014), the underlying mechanism remains unclear.

Flavanones (naringenin and liquiritigenin) as common flavonoid precursors are converted into 2-hydroxyisoflavanones, catalyzed by isoflavone synthase (IFS; EC 1.14.13.136), indicating that IFS is the key enzyme of isoflavonoid biosynthesis pathway in plants. The *IFS* genes have been identified from a variety of legumes such as *Glycine max* (L.) Merrill, *Pisum sativum* (L.), *Medicago sativa* (L.), *Trifolium pretense* (L.), white lupine and others (Akashi et al. 1999; Jung et al. 2000; Kim et al. 2003; Shimada et al. 2000). 2-Hydroxyisoflavanones are subsequently dehydrated, either spontaneously or catalyzed by 2-hydroxyisoflavanone dehydratase (HID; EC 4.2.1.105), to the corresponding isoflavones (genistein and daidzein) (Akashi et al. 1999, 2005; Jung et al. 2000; Shimada et al. 2000).

In planta, most isoflavonoids are exist as conjugates, not as aglycones. Isoflavone aglycones are converted into conjugates by glucosyltransferases, prenyltransferases, malonyltransferases and others, affecting pharmacokinetics and biological activities. Therefore, isoflavone conjugation is important for plant growth and human health (Noguchi et al. 2007). For example, glycosylation modification improves the solubility, stability and detoxification of flavonoids for defense and various changes in environment (Vogt and Jones 2000). UDP-glucose: isoflavone 7-*O*-glucosyltransferase in *G. max* (GmUGT88E3; EC 2.4.1.170) plays a key role in isoflavone conjugate synthesis, which leads to more accumulation in vacuoles and acts as an isoflavone pool involved in interaction with microorganisms (Noguchi et al. 2007). Compared with non-prenylated analogs, prenylated flavonoids have the increased lipophilicity and membrane permeability, which enhances biological activities and binding affinity of flavonoids for target biomolecules (Botta et al. 2005; Wang et al. 1997). Prenylation, catalyzed by prenyltransferases, is a rate-limiting step in the biosynthesis of prenylated flavonoids (Sasaki et al. 2011).

Lupins belong to a single genus, *Lupinus*, in a legume clade—the genistoids, one of the earliest diverging lineages of the Papilionoideae which evolved about 50–60 million years ago (Cannon et al. 2015). The whole genome triplicated at round 25 million years ago in the genistoid lineage, leading to *Lupinus* (Hane et al. 2017). *Lupinus* comprises an estimated 267 species (Drummond et al. 2012), and newly bred cultivars of domesticated *Lupinus* species consist of *Lupinus albus*, *Lupinus angustifolius*, *Lupinus luteus* and *Lupinus mutabilis*, which are edible (Karamac et al. 2018). Contrary to most genera of legumes, a variety of prenylated and glycosylated isoflavonoids are accumulated in lupins (Aisyah et al. 2016). Moreover, lupin seeds are rich in protein with very little starch compared to other major grain legumes such as chickpea and soybean (Villarino et al. 2016). Seeds of white lupin (*L. albus*) are used in snack foods in the Middle East and Europe, and owing to rich bioactive compounds such as isoflavonoids, rich protein and very low starch, recently lupin seeds are also used as food ingredients for baked goods, dairy products and fermented foods to reduce the risk of dyslipidaemia, diabetes, obesity, hypertension and bowel dysfunction (Arnoldi et al. 2015; Villarino et al. 2016). Although there are rich in prenylated and glycosylated isoflavonoids in white lupin, the isoflavonoids biosynthesis pathway has not been fully resolved.

In the present study, isoflavonoid compounds were determined in several *Lupinus* species by HPLC method, and only 5-hydroxyisoflavonoids (genistein and its conjugates) were detected in every tested *Lupinus* species. Furthermore, *Lupinus* plants contained no 5-deoxyisoflavonoids due to lack of CHRs and Type II CHIs. Isoflavone synthases LaIFS1 and LaIFS2 of white lupin converted substrates naringenin and liquiritigenin to 5-hydroxyisoflavone (genistein) and 5-deoxyisoflavone (daidzein), although there are no 5-deoxyisoflavonoids in white lupin. In addition, LaG6DT1 and LaG6DT2 were characterized and prenylated genistein at the C-6 position into wighteone. Two glucosyltransferases LaUGT1 and LaUGT2 metabolized genistein and wighteone into its 7-*O*-glucosides. Our study not only revealed that the exclusive 5-hydroxyisoflavonoids do exist in *Lupinus* species, but also identified key enzymes in the biosynthesis pathway (Fig. 1).

Materials and methods

Plant materials and growth conditions

Seeds of *Lupinus albus* (white lupin, cultivar "Sweet Lupini"), *Lupinus succulentus*, *Lupinus texensis*, *Lupinus popsicle*, *Lupinus hartwegii*, and *Glycine max* cultivar Williams 82, were obtained from the United States Department of Agriculture Soybean Germplasm Collection. Seeds were germinated and grown on sterile vermiculite pot in a controlled environment with 16 h of light and 8 h of dark at 22 °C, and watered with half-strength Murashige and Skoog (MS) liquid solution every week. The roots, stems, and leaves were collected, immediately frozen in liquid nitrogen, and stored at − 80 °C for further use.

Chemicals

Standard Chemicals, including naringenin chalcone, isoliquiritigenin, naringenin, liquiritigenin, genistein, genistin, and wighteone, were purchased from Shanghai yuanye Bio-Technology Co., Ltd (Shanghai, China). DMAPP was purchased from Sigma-Aldrich Corp. (St. Louis, MO). Chromatographic methanol and acetonitrile reagents were purchased from Fisher Chemical. Yeast synthetic drop-out (SD) medium was purchased from Sangon Biotech (Shanghai, China).

Identification and phylogenetic analysis of flavonoid pathway genes

The white lupin transcriptome library (Secco et al. 2014) was used to search the potential flavonoid-related genes using a BLAST program based on soybean (iso)flavonoid pathway genes as queries. Yielded contigs were used to predict the presence of a complete or incomplete ORFs. Only ORFs, which were predicted to encode enzymes in the (iso)flavonoid pathway, were used to conduct a second BLAST search to avoid misassembled contigs and remove redundant contigs. Primers (Table S1) were designed based on ORF sequences and polymerase chain reaction (PCR) was done to prove the existence of ORF sequences obtained from transcriptome library. Amino acid alignments of (iso)flavonoid-related proteins in white lupin and previous reported proteins from other plant species were conducted using ClustalX2 algorithm with default parameters. Based on these alignments, phylogenetic trees were constructed using a maximum-likelihood algorithm in MEGA6. A bootstrap resampling analysis with 1000 replicates was performed to evaluate the topology of the phylogeny with JTT substitution model.

Heterologous expression and characterization of CHI

The full-length ORFs of *CHI* genes from white lupin and soybean as positive control were amplified with Q5 High-Fidelity DNA Polymerase (New England Biolabs). The PCR product was digested with *Spe* I and *Xho* I (New England Biolabs), and subcloned into yeast vector pDR196GW, a yeast expression vector carrying the constitutive yeast PAM1 promoter, was kindly provided by Professor Kazafumi Yazaki (Kyoto University). The resulting vectors containing the *CHI* gene were individually transformed into yeast strain DD104 (Li et al. 2015), using lithium chloride-mediated transformation method as described as in the Clontech manual. The transformants were selected on SD (-uracil) plates, and further confirmed by PCR. The transformed yeast colony was inoculated into SD (-uracil) medium, and incubated at 30 °C with shaking at 200 rpm. At an OD_{600} of 1.0, naringenin chalcone or isoliquiritigenin were added to a final concentration of 100 μM, and incubation was continued for 20 h. The culture medium was collected and centrifuged to remove cells and debris, and the supernatant mixed with equal volume of methanol was directly injected for naringenin and liquiritigenin determination.

Heterologous expression and characterization of cytochrome P450 genes and prenyltransferase genes

For cytochrome P450 genes, the full-length ORFs of isoflavone synthase genes (*IFS*), were amplified with Q5 High-Fidelity DNA Polymerase (New England Biolabs). The PCR product was digested with *Sma* I and *EcoR* I (New England Biolabs), and subcloned into yeast vector pYeDP60 (BioVector NTCC, Beijing, China). The resulting vectors containing the *IFS* gene were individually transformed into yeast strain WAT11 (BioVector NTCC, Beijing, China), which was designed to express eukaryotic cytochrome P450 proteins (Pompon et al. 1996), using lithium chloride-mediated transformation method as described as in the Clontech manual. The transformants were selected on SD (-adenine, -uracil) plates, and further confirmed by PCR. The transformed yeast colony was inoculated into 20 mL SD (-adenine, -uracil) medium, and incubated at 30 °C with shaking at 200 rpm for overnight. Ten milliliters of the culture was transferred to 500 mL fresh SD (-adenine, -uracil) medium, and incubation was continued for overnight until no glucose was detectable with glucose test strip. Then galactose was added to a final concentration of 2% (w/v) to induce *IFS* expression. After

inducing for 20 h, yeast cells were collected and ground into power under liquid nitrogen condition.

For flavonoid-related prenyltransferase genes, the prenyltransferase genes were subcloned into yeast vector pDR196GW by *Spe* I and *Xho* I sites to produce pDR196GW-LaG6DT1 and pDR196GW-LaG6DT2, which were individually transformed into an engineered yeast strain DD104 that had accumulated DMAPP (Fischer et al. 2011; Li et al. 2015), using lithium chloride-mediated transformation method as described as in the Clontech manual. The transformants were selected on SD (-uracil) plates, and further confirmed by PCR. The transformed yeast colony was first cultured in 20 mL SD (-uracil) medium, and incubated at 30 °C with shaking at 200 rpm for overnight. Ten milliliters of the culture was transferred to 500 mL fresh SD (-uracil) medium, and incubation was continued for overnight until no glucose was detectable with glucose test strip. The yeast cells were collected and ground into power under liquid nitrogen.

The yeast microsomal proteins were extracted as described previously (Liu et al. 2018; Yazaki et al. 2002). In brief, the yeast powder was dissolved in chilled extraction buffer (50 mM Tris–HCl pH 7.5, 0.6 M sorbitol, 1% bovine serum albumin, 2 mM 2-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride). Yeast debris was removed by centrifugation at 10,000*g* for 15 min, and the supernatant was carefully transferred to 30 mL tube (Beckman). Microsomes were isolated from the supernatant by ultracentrifugation for 90 min at 100,000*g* with optima XPN-100 Ultracentrifuge (Beakman). The supernatant was carefully discarded and the microsomes were washed once with 1 mL TEG-M buffer (50 mM Tris–HCl pH7.5, 20% glycerol, and 2 mM 2-mercaptoethanol). Then the microsomal fraction containing the recombinant protein was re-suspended and homogenized in 1.5 mL of TEG-M buffer with glass homogenizer. The microsomal protein concentration was determined by Bradford method (Bradford 1976). All microsomal proteins were stored at − 80 °C. All steps for yeast microsomal preparation were performed at 4 °C.

The IFS activity was assayed in a total reaction volume of 200 µL containing 50 mM sodium phosphate buffer (pH 7.5), 1 mM reduced glutathione, 1 mM NADPH, 50 µg of crude microsomal proteins. The reaction was initiated by adding 100 µM substrates and incubated at 30 °C for 30 min. The reaction was stopped with adding an equal volume of methanol containing 5% acetic acid. Microsomal proteins from yeast cells harboring an empty vector were used as the negative control. Product quantification was done using a standard curve of genistein or daidzein. For kinetic parameters of IFS, a range change of naringenin or liquiritigenin substrates (5–400 µM) with a constant NADPH concentration (2 mM) were reacted in a total volume of 50 µL with three independent replications. For kinetic parameters of prenyltransferase, a range change of genistein substrates (5–400 µM) with a constant DMAPP concentration (330 µM) were reacted in a total volume of 50 µL with three independent replicates.

To test substrate specificity, the reaction mixtures, containing 100 mM Tris–HCl (pH 7.5), 1 mM DTT, 25 mM MOPS (pH 7.0), 10 mM Mg^{2+}, 100 µM isoflavonoid as prenyl acceptor, 200 µM DMAPP (Sigma-Aldrich) as prenyl donor, and 20 µg microsomal proteins, were incubated in a total volume of 100 µL at 30 °C for 30 min. The reaction was terminated by adding an equal volume of methanol. Product quantification was done using a standard curve of substrates for their similar UV absorption characteristic.

Heterologous expression and characterization of isoflavonoid glucosyltransferases

The full-length ORFs of glucosyltransferase sequences were amplified with Q5 High-Fidelity DNA Polymerase (New England Biolabs). The corresponding PCR products were digested with *BamH* I and *Sal* I (New England Biolabs), and then ligated to a pMAL-c2X vector (New England BioLab Inc.) digested with the same restriction enzymes. The resulting pMAL-c2X-LaUGT vectors were sequenced and transformed into *Escherichia coli* strain BL21(DE3). Recombinant proteins were induced as previously described (Yin et al. 2017). In brief, a single colony was inoculated into 10 mL LB medium (100 mg/L ampicillin and 2 g/L glucose) at 37 °C with 200 rpm for overnight, and then five milliliters of the culture was transferred to 500 mL fresh LB medium (100 mg/L ampicillin and 2 g/L glucose) and incubation was continued until OD600 = 1.0. Then isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.35 mM, and incubation was continued at 16 °C with 200 rpm for 16 h. Recombinant proteins were purified as previously described (Yin et al. 2017; Feng et al. 2021). Enzyme activity was performed with 10 µg purified recombinant proteins, 100 µM isoflavonoid substrate, and 500 µM UDP-Glucose in Tris–HCl buffer (pH 7.5) in a total volume of 100 µL. After adding isoflavonoid substrates, the reaction mixture was incubated at 30 °C for 30 min, and stopped by the addition of an equal volume of methanol.

Functional characterization of (iso)flavonoid pathway genes in tobacco

To test the in vivo function of these (iso)flavonoid-related genes, they were subcloned into the binary pCAMBIA2300-35S-OCS vector between the sites of *Sma* I and *Xba* I. The resulting vectors were individually transformed into *Agrobacterium tumefaciens* strain GV3101. A single colony confirmed by PCR was inoculated into 5 mL LB culture (containing 50 mg/L kanamycin, 25 mg/L rifampicin), and incubated at 28 °C with shaking at 220 rpm for overnight.

Then one milliliter of the culture was transferred to 10 mL LB medium containing the same antibiotics for about 3–5 h. At the OD_{600} value of 0.6–0.8, *Agrobacterium* cells were centrifuged for 15 min at 3600*g*, washed with 5 mL infiltration buffer (10 mM MES, 10 mM $MgCl_2$, and 100 µM acetosyringone), and then re-suspended in the infiltration buffer to a final OD_{600} value of 0.5. The *Agrobacterium* cells containing (iso)flavonoid pathway genes and P19 gene were mixed with an equal volume and incubated for an hour at room temperature prior to infiltration. The leaves of 4–6-week-old tobacco (*Nicotiana benthamiana*) were infiltrated with the *Agrobacterium* cells, and grown under 16/8 h light/dark rhythms for 3 days. Then the infiltrated leaves were re-infiltrated with 200 µM substrates, and grown for 24 h at the same condition as before. Tobacco leaves infiltrated with the *Agrobacterium* cells carrying the vector pCAMBIA2300-35S-OCS as the negative control. These leaves were directly harvested, frozen and stored at − 80 °C for further analysis.

Analysis of (iso)flavonoid metabolites

Tobacco leaves were ground into a fine power, and extracted with a five-fold volume of methanol (w/v) at room temperature for overnight. After centrifugation at 16,000*g* for 15 min the supernatant was transferred to a new microtube, and then mixed with 0.3 fold volume of ultrapure water and incubated at − 20 °C for overnight. The mixed solution was centrifuged again at 16,000*g* for 15 min to remove chlorophyll, and the supernatant was filtrated through 0.22 µm springe filters. Then the (iso)flavonoid extract was further determined by HPLC or UPLC-MS/MS.

For specific flavonoid metabolomics analysis, 20 µL of the final extract was analyzed by HPLC (Agilent 1260) as previously described with minor modifications (Liu et al. 2021). In brief, chromatographic separation of flavonoid metabolomics was performed on an Eclipse XDB-C18 column (4.6 × 150 mm, 5 µm; Agilent) with column temperature maintained at 35 °C. The linear gradient program of binary mobile phases (solvent A: water containing 0.1% formic acid; solvent B: acetonitrile) was as follows: 0 min, 95% A to 30 min, 30% A at flow rate of 1 ml. A photodiode array detector was used for the detection of UV–visible absorption from 190 to 600 nm. For quantification of wighteone in tobacco, an Agilent 1290 infinity UPLC hyphenated with 6470 triple quadrupole mass spectrometry with electrospray ionization source (LC-ESI-QQQ-MS/MS) was used. Chromatographic separation was performed using a ZORBAX Eclipse Plus C18 column (Agilent, 2.1 × 50 mm, 1.8 µm) at 30 °C. 5 µL of sample was injected, and the linear gradient program of binary mobile phases (solvent A: water with 0.1% formic acid; solvent B: acetonitrile) was as follows: 0 min, 70% A; 5 min, 40% A; 7 min, 10% A at flow rate of 0.3 ml. The quantitative method was established based on multiple reaction monitoring (MRM) mode with detailed description in Table S2. The optimum operating ESI condition was as follows: gas temperature, 330 °C; gas flow rate, 13 L/min; nebulizer pressure, 40 psi; sheath gas temperature, 300 °C; sheath gas flow, 12 L/min. The capillary voltages were optimized to 4000 V in positive mode.

RNA extraction and gene expression analysis

For gene expression analysis, leaves, stems and roots were harvested separately from white lupin plants and total RNAs were extracted using TRIzol reagent (Takara, Japan). Total RNAs were first treated with DNase I to digest genomic DNA and then were used to synthesize the first-strand cDNA using Moloney Murine Leukemia Virus Reverse Transcriptase (Promega, WI, USA). The cDNA was used as template for quantitative real-time PCR (qRT-PCR) analysis with SYBR Green PCR Kit (Kangwei Biotech, China). The qRT-PCR experiments were performed on ABI 7500 real-time Detection System (Applied Biosystems, USA) using gene-specific primer pairs listed in Table S1. The *Actin* gene, described as previous study (Shen et al. 2012), was used as a housekeeping gene. The reaction was carried out as follows: 94 °C for 30 s, followed by 40 cycles of 94 °C, 5 s; 60 °C, 20 s; and 72 °C, 20 s. Data were calculated from three biological replicates, and each biological replicate was examined in triplicate. Relative transcript abundance was calculated based on the $2^{-\Delta\Delta CT}$ method (Pfaffl 2001). The transcript abundance was presented as relative value, and the transcript abundance in the leaves was set as value of 1: LaCHI1 in Fig. 2b, LaIFS1 in Fig. 3b, LaG6DT1 in Fig. 4b, LaUGT1 in Fig. 5b.

Statistical analyses

The SPSS 20.0 statistical package (IBM SPSS Statistics for Windows, Version 20.0; IBM Corp) was used for the statistical data analysis. Statistical analysis was done with Student's t test. Samples included in the analysis were arranged based on at least three biological replicates.

Results

Lupinus plants biosynthesized 5-hydroxyisoflavonoids, but not 5-deoxyisoflavonoids

Isoflavones are classified into 5-hydroxyisoflavones and 5-deoxyisoflavones based on presence and absence of C–OH at C-5 position of isoflavone skeleton (Fig. S1). We summarized the data of detailed isoflavonoid compounds of white lupin, which had been clearly characterized over

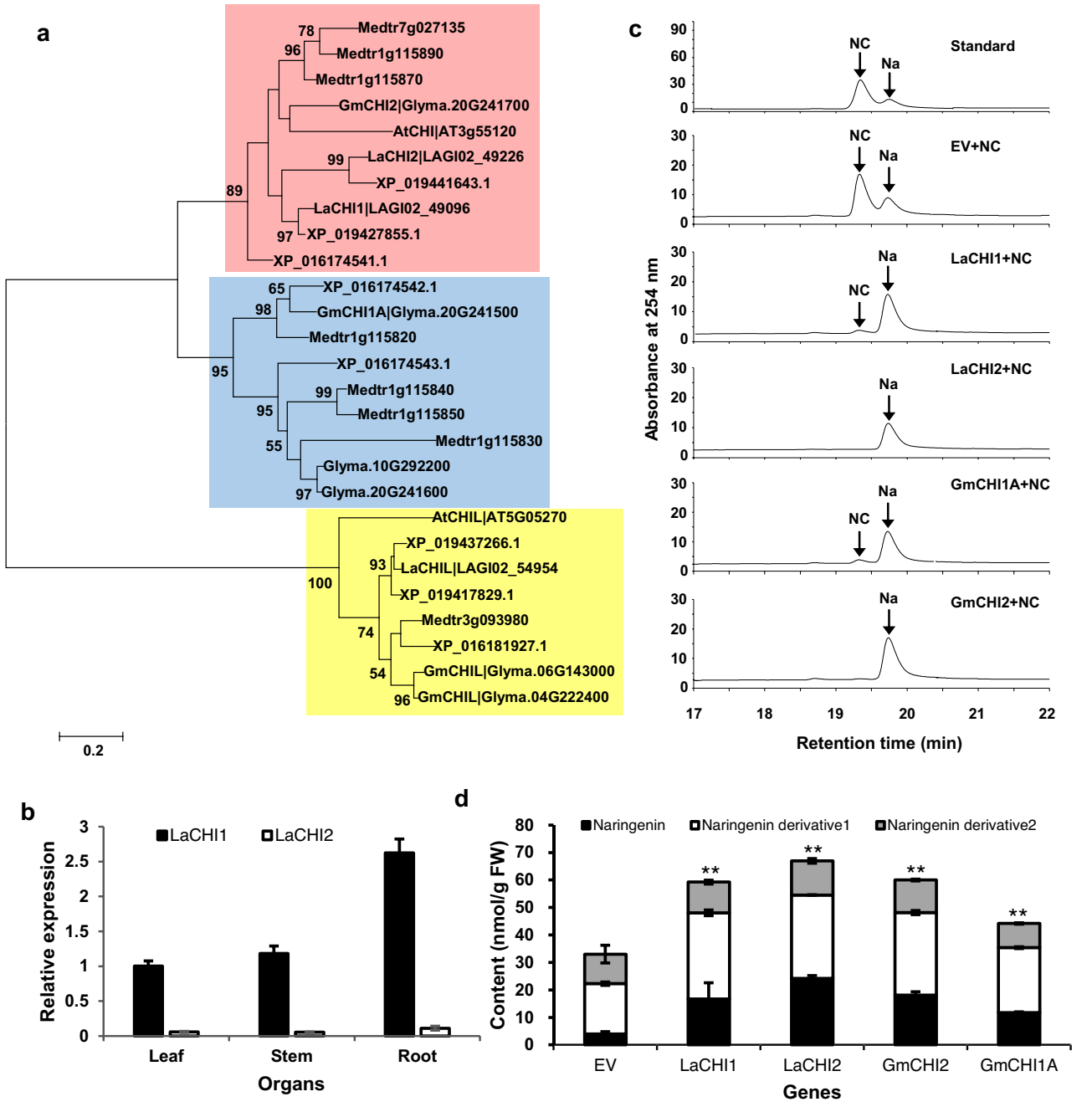


Fig. 2 Identification and functional characterization of CHIs in white lupin. **a** A maximum-likelihood phylogenetic tree of CHI family. The three groups are Type IV CHI (CHI Like, yellow), Type I (red) and Type II CHI (blue). Protein sequences used in the phylogenetic tree construction are from *G. max*, *M. truncatula*, *A. thaliana*, *A. ipaensis*, *L. albus* and *L. angustifolius*. **b** Expression pattern of *LaCHI1* and *LaCHI2* in different organs of 2-month-old white lupin. **c** Chromatograms of reaction product naringenin (Na) of CHI enzymes with naringenin chalcone (NC) as substrates in yeast. The yeast containing the empty vector (EV) is conducted as the negative control, and GmCHI1A and GmCHI2 are conducted as the positive control in assays. **d** Naringenin content in transgenic CHI tobacco in comparison with assays conducted using empty vector (EV) as the negative control, and GmCHI1A and GmCHI2 as the positive control. The identification of products was verified by comparison with assays that authentic standard naringenin was incubated with tobacco. Data are means ± SD of three biological replicates. Asterisks indicate significant differences by Student's *t* test (**$P < 0.01$) (color figure online)

the past two decades (Bednarek et al. 2001; Kachlicki et al. 2005; Katagiri et al. 2000), and discovered that all isoflavonoid compounds consisting of isoflavonoid aglycones and conjugates such as glycosylation, methylation and prenylation, were derived from genistein, which represents a main isoflavone category containing C–OH at C-5, C-7 and C-4′

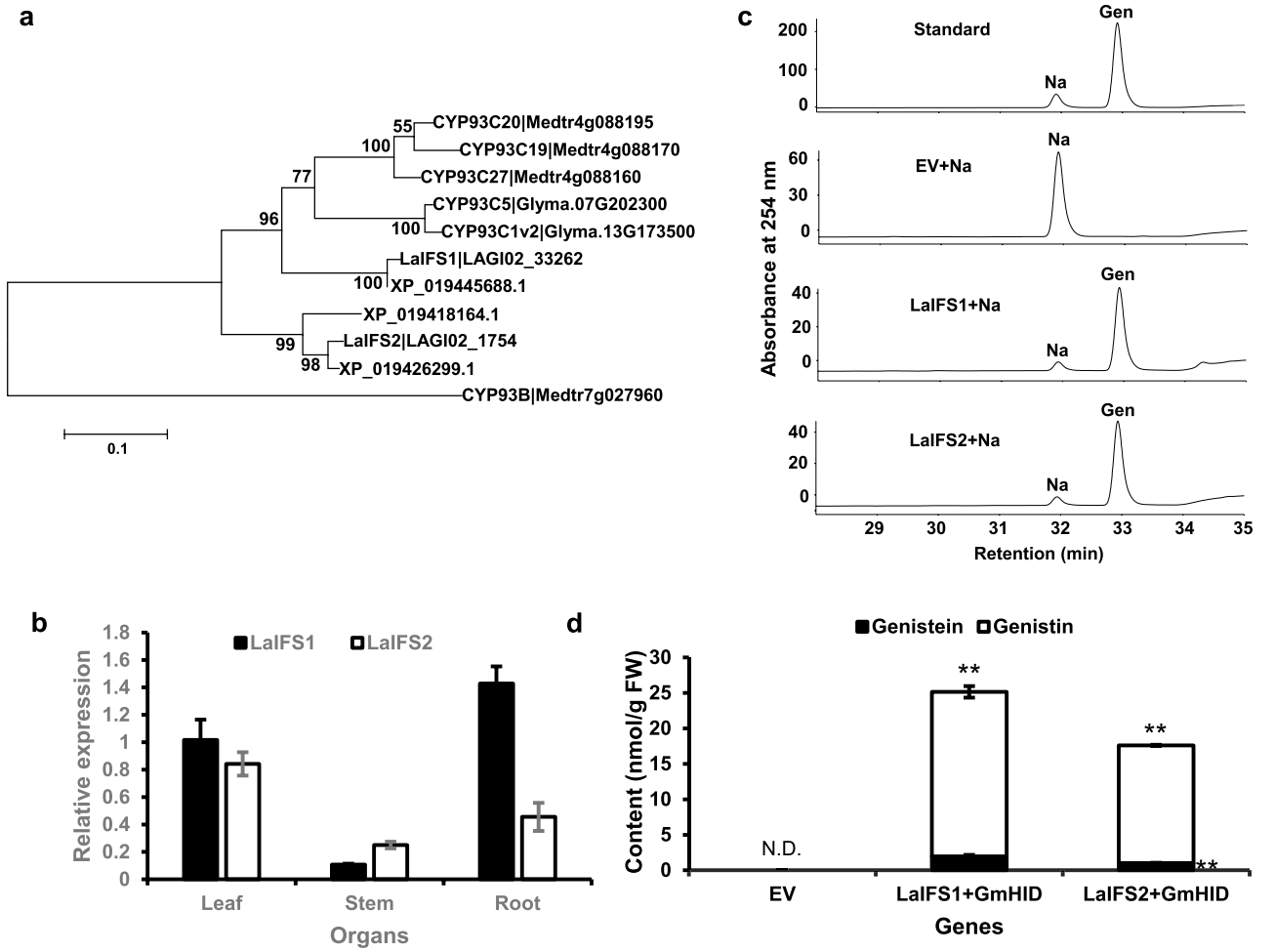


Fig. 3 Identification and functional characterization of LaIFS. **a** A maximum-likelihood phylogenetic tree of legume isoflavone synthase. Protein sequences used in the phylogenetic tree construction are from *G. max*, *M. truncatula, L. albus* and *L. angustifolius*. The outgroup of the phylogenetic tree is a member of cytochrome P450 93B subfamily which functions as a flavone synthase. **b** Expression pattern of *LaIFS1* and *LaIFS2* in different organs of 2-month-old white lupin. **c** Chromatograms of reaction product genistein (Gen) of LaIFS enzymes with naringenin (Na) as substrates in yeast. The recombinant IFS proteins were from yeast expression systems. The microsomal protein of yeast containing the empty vector (EV) is conducted as the negative control in assays. **d** The content of genistein and genistin (genistein 7-*O*-glucose) in transgenic *LaIFS* tobacco in comparison with assays conducted using empty vector (EV) as the negative control. The identification of the products was verified by comparison with assays that authentic standard genistein was incubated with tobacco. Data are means ± SD of three biological replicates. Asterisks indicate significant differences by Student's *t* test (**$P < 0.01$)

position of isoflavone skeleton (Fig. S1). To experimentally confirm it, we reinvestigated isoflavonoid components of white lupin. Based on the isoflavone standards, genistein, genistin and wighteone were identified from white lupin (Fig. S2), whereas the 5-deoxyisoflavone was not detectable, as described in previous studies (Bednarek et al. 2001; Kachlicki et al. 2005; Katagiri et al. 2000). These data showed that only 5-hydroxyisoflavonoid compounds (genistein and its conjugates), not 5-deoxyisoflavonoids, were biosynthesized in white lupin.

Previous studies reported that there are only 5-hydroxyisoflavonoids, not 5-deoxyisoflavonoids, in several other species of *Lupinus* (Aisyah et al. 2016; Wojakowska et al. 2013). Therefore, we hypothesized that all species of *Lupinus* biosynthesize 5-hydroxyisoflavonoids as the main isoflavonoid components. To prove our hypothesis, we collected other five species of *Lupinus*: *L. micranthus*, *L. succulentus*, *L. texensis*, *L. popsicle*, *L. hartwegii*, and the isoflavonoid category was characterized using HPLC method. These data clearly showed that 5-hydroxyisoflavones were detected in every tested species of *Lupinus* (Fig. S3), but not 5-deoxyisoflavone components. To rule out the influences from artificial factors in our experiments, soybean (*Glycine max*) was added as the control. 5-hydroxyisoflavonoids were detected from soybean leaves, and 5-hydroxyisoflavonoids and 5-deoxyisoflavonoids were detected from soybean roots

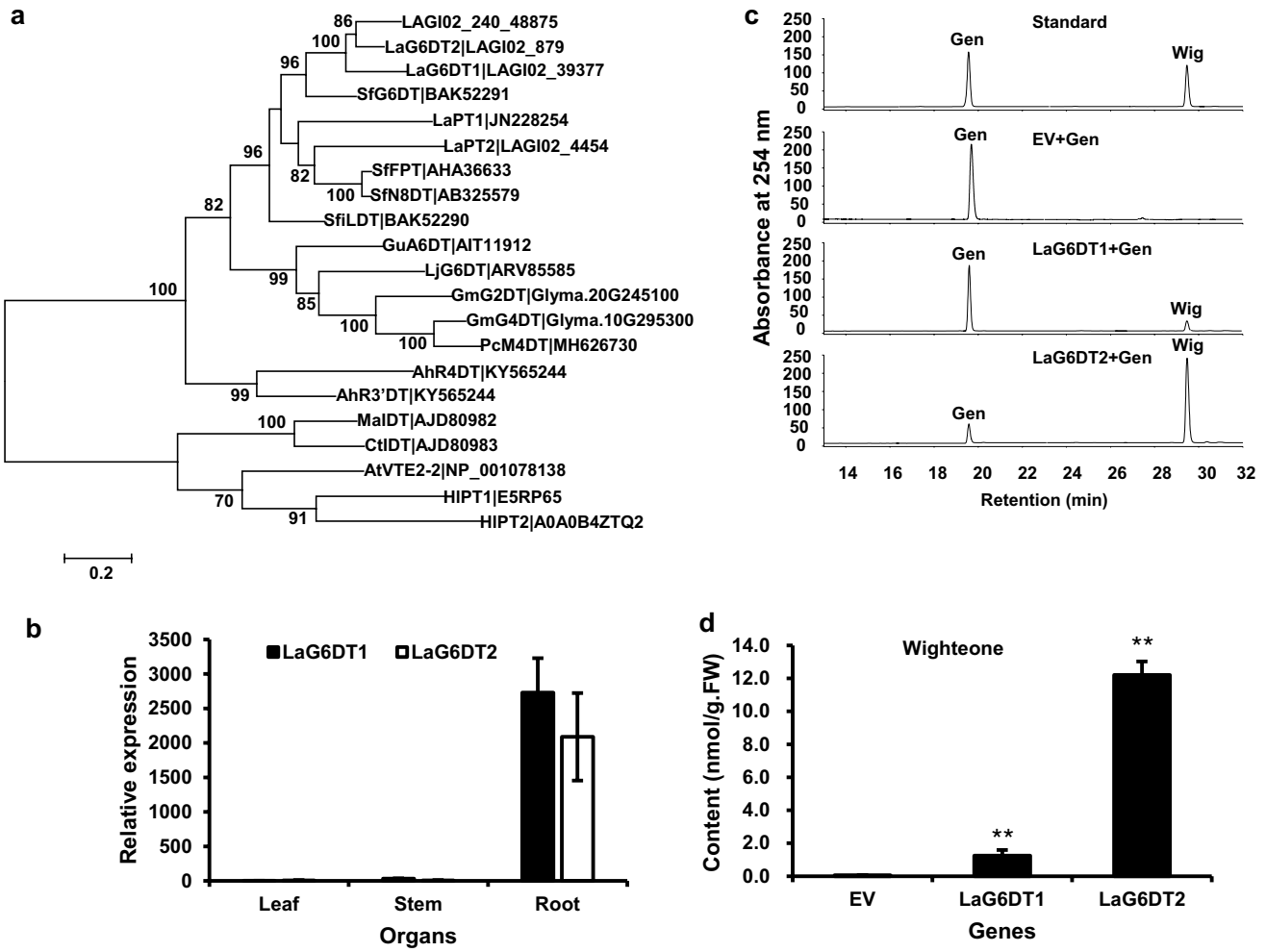


Fig. 4 Identification and functional characterization of LaG6DT. **a** A maximum-likelihood phylogenetic tree of prenyltransferases. Protein sequences used in the phylogenetic tree construction are from *G. max*, *L. albus*, *A. ipaensis* and other species, which function were undoubted characterized in previous research. **b** Expression pattern of *LaG6DT1* and *LaG6DT2* in different organs of 2-month-old white lupin. **c** Chromatograms of reaction product wighteone (Wig) of LaG6DT enzymes with genistein (Gen) as substrates in yeast. The recombinant G6DT proteins were from yeast expression systems. The microsomal protein of yeast containing the empty vector (EV) is conducted as the negative control in assays. **d** The content of wighteone in transgenic *LaG6DT* tobacco in comparison with assays conducted using empty vector (EV) as the negative control. Data are means ± SD of three biological replicates. Asterisks indicate significant differences by Student's *t* test (**$P < 0.01$)

(Fig. S2), consistent with the distribution of isoflavonoids in *G. max* (Aisyah et al. 2013). These results strongly suggested that *Lupinus* plants biosynthesized unique 5-hydroxyflavonoid components, but the model plants of legume, such as *G. max* and *Medicago trunctula*, produce both 5-hydroxyisoflavonoids and 5-deoxyisoflavonoids components (Aisyah et al. 2013; Farag et al. 2008),

No CHR orthologs were identified in *Lupinus* genome

No 5-deoxyisoflavonoids were observed from *Lupinus* species in this (Figs. S2 and S3) and previous studies (Aisyah et al. 2016; Wojakowska et al. 2013). CHR is a key enzyme that catalyzes the first committed step of 5-deoxyisoflavonoid biosynthesis (Bomati et al. 2005; Mameda et al. 2018; Welle and Grisebach 1988), indicating that the exist of CHR is a prerequisite of 5-deoxyflavonoids biosynthesis. Therefore, we hypothesized that no 5-deoxyisoflavones in *Lupinus* plants were due to the lack of *CHR* gene. To prove our hypothesis, we used the identified *CHR* sequence as a bait to search the genome of *L. angustifolius* that is the first sequenced species of *Lupinus* (Hane et al. 2017). No orthologs were identified from *L. angustifolius* in our screen, and a similar result was also obtained by searching white lupin transcript library (Fig. S4). These results clearly indicated that the absence of *CHR* gene led to no 5-deoxyisoflavonoids in *Lupinus* species.

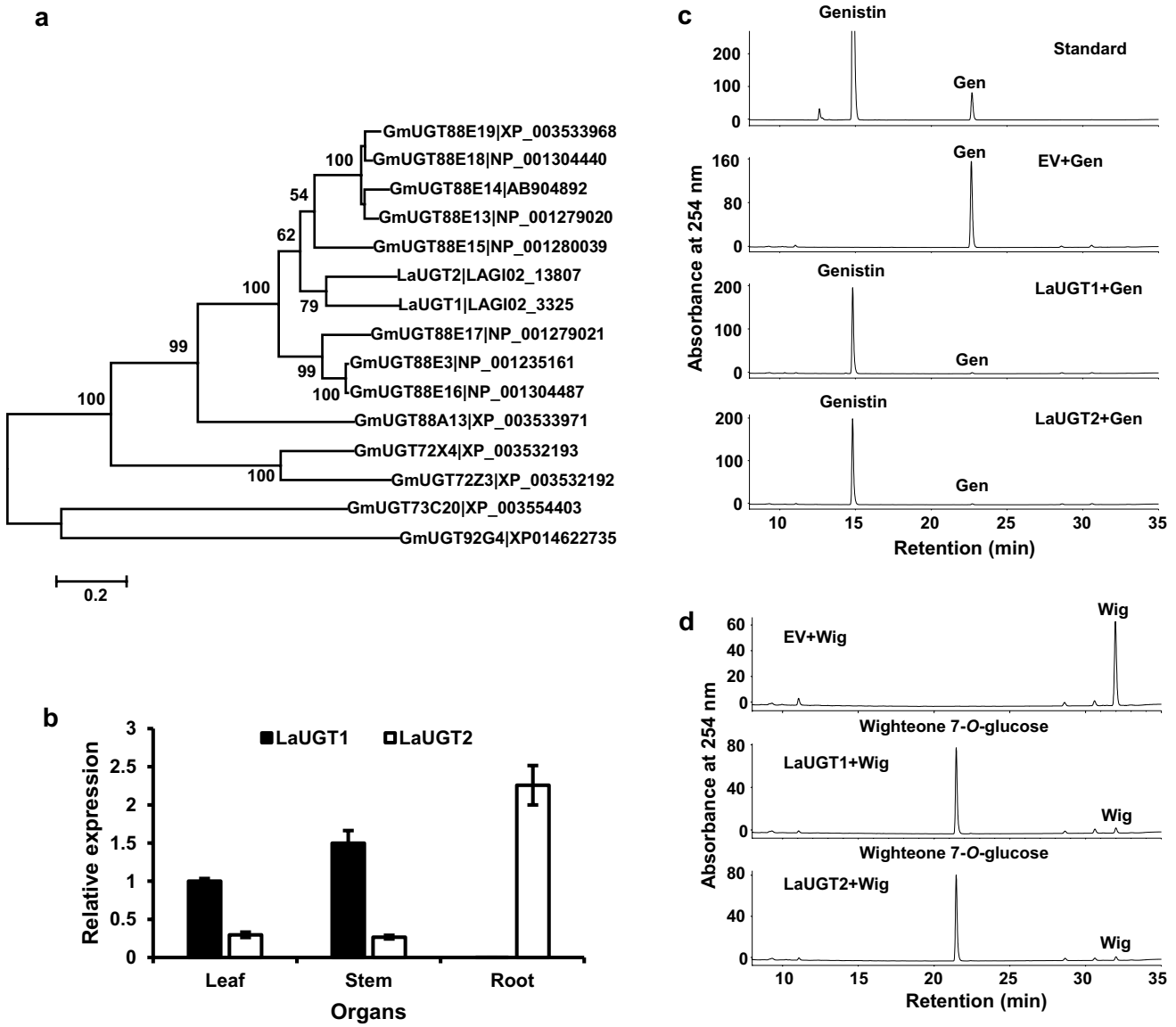


Fig. 5 Identification and functional characterization of isoflavone 7-*O*-glucosyltransferases. **a** A maximum-likelihood phylogenetic tree of glucosyltransferases. Protein sequences used in the phylogenetic tree construction are mainly from *G. max*, except the two involved in this study from *L. albus*. The glucosyltransferases from *G. max* were all functional characterized as flavonoid glucosyltransferases in previous research. **b** Expression pattern of *LaUGT1* and *LaUGT2* in different organs of 2-month-old white lupin. **c** Chromatograms of reaction product genistin (genistein 7-*O*-glucose) of LaUGT enzymes with genistein (Gen) as substrates in *Escherichia coli*. **d** Chromatograms of reaction product wighteone 7-*O*-glucose of LaUGT enzymes with wighteone (Wig) as substrates in *E. coli*. The recombinant UGT proteins were expressed and purified from *E. coli* expression systems. The purified MBP protein (EV) is conducted as the negative control in assays

Type I CHI, not Type II CHIs were characterized in white lupin

In silico screening of white lupin transcript library with a probe based on *CHI* genes from *G. max*, we identified three putative *CHI* contigs numbered as LAGI02_49096, LAGI02_49226 and LAGI02_54954. These three contigs contained full-length open reading frames (ORFs). Among them, LAGI02_49096 and LAGI02_49226 had higher sequence similarity to GmCHI2 (Glyma.20G241700, accession number: NP_001236768), Type I CHI, than to GmCHI1A (Glyma.20G241500, NP_001235219), Type II CHI (Ralston et al. 2005). LAGI02_54954 displayed more sequence similarity to putative GmCHIL (Glyma.06G143000 and Glyma.04G222400) than to CHI2 and CHI1A (Fig. S5). Phylogenetic analysis showed that LAGI02_49096 and LAGI02_49226 were clustered together with Type I CHI, obviously distinguished from

Type II CHI; and LAGI02_54954 was clade with CHIL (Type IV CHI) proteins (Fig. 2a). These results showed that LAGI02_54954 as a Type IV CHI (CHIL), did not directly catalyze chalcone substrates and LAGI02_49096 (referred to as *LaCHI1*, GenBank accession No. MT427390) and LAGI02_49226 (*LaCHI2*, GenBank accession no. MT427391) may function as Type I CHI. Therefore, we focused on these two *LaCHIs* to prove whether they functioned as Type I CHI.

The qRT-PCR analysis showed both *LaCHI1* and *LaCHI2* were expressed in roots, stems and leaves, and *LaCHI1* had higher expression levels compared with *LaCHI2* in the three tested organs (Fig. 2b). Previous studies show that yeast is a suitable system to test CHI enzymatic activity (Ralston et al. 2005). Therefore, the two *LaCHIs* genes, and *GmCHI2* and *GmCHI1A* as positive controls, were introduced into yeast, respectively. Transformed yeast cultures were fed with chalcone naringenin, and the metabolites were determined with HPLC method. Yeast transformed by *LaCHI1* or *LaCHI2* metabolized naringenin chalcone to naringenin (Figs. 2c), but it failed to convert isoliquiritigenin to liquiritigenin (Fig. S6). In addition, we tried to use the tobacco transient expression system to further confirm the two LaCHIs function *in planta*. *LaCHI1*, *LaCHI2*, *GmCHI2* or *GmCHI1A* were transiently expressed in tobacco leaves containing supplied exogenous substrates. Due to the endogenous *CHI* genes, tobacco containing the empty plasmid could convert naringenin chalcone to naringenin (Figs. 2d and S7), and the leaves of expressed *LaCHI1*, *LaCHI2* or *GmCHI2* metabolize naringenin chalcone to naringenin more efficiently, and then more naringenin and naringenin derivatives were accumulated (Figs. 2d and S7). Furthermore, our data showed that LaCHI1 or LaCHI2 failed to convert isoliquiritigenin to liquiritigenin, and further supported the fact that GmCHI1A is able to metabolize isoliquiritigenin to liquiritigenin *in planta* (Fig. S8). Taken together, these evidences from phylogenetic analysis and enzyme activities in vivo suggested that LaCHI1 and LaCHI2 were the *bona fide* Type I CHIs in white lupin.

LaIFS1 or LaIFS2 functioned as isoflavone synthases

Isoflavone synthase (IFS) metabolizes flavanones (naringenin and liquiritigenin) into 2-hydroxyisoflavanone, which is the key step in isoflavonoid biosynthesis pathway. A previous study (Jung et al. 2000) reported an isoflavone synthase EST (AF195813) sequence from white lupin. However, when we used this EST sequence as a query to BLAST against white lupin transcriptional library, no identical sequences were yielded. Here, our study focused on two *IFS* genes: *LaIFS1* (GenBank accession no. MT427392) and *LaIFS2* (GenBank accession no. MT427393) identified from contigs LAGI02_33262 and LAGI02_1754 in white lupin, respectively (Fig. 3a). These two LaIFSs shared high sequence similarities with those of *L. angustifolius* (Narozna et al. 2017). The qRT-PCR analysis showed that the transcripts of *LaIFS1* and *LaIFS2* were detected in roots, stems and leaves, and the transcript level in leaves and roots was higher than that in stems (Fig. 3b), which is consistent with the fact that the amounts of isoflavonoids were much higher in leaves and roots compared with those in stems.

LaIFS1 or LaIFS2 was heterologously expressed in yeast, and then microsomal proteins contained LaIFS1 or LaIFS2 were incubated with naringenin and NADPH in each individual reaction. The experiments showed that LaIFS1 or LaIFS2 catalyzed flavanone naringenin into isoflavone genistein in vitro (Fig. 3c), and also catalyzed liquiritigenin into daidzein (Fig. S9) although there are no 5-deoxyisoflavonoids in white lupin. The kinetic parameters shown in Table 1 suggested that LaIFS1 or LaIFS2 had an order higher of V_{max}/K_m values for liquiritigenin than naringenin, which is consistent with the previous report for other IFS (Jung et al. 2000).

To determine the function of LaIFS1 or LaIFS2 in vivo, these two genes were transiently expressed in tobacco leaves. The isoflavone compounds could not be always detected from the extracts of tobacco leaves supplied with flavanone naringenin or liquiritigenin. Isoflavanones, the products of IFS, are dehydrated spontaneously or catalyzed by 2-hydroxyisoflavanone dehydratase (HID) to isoflavone (Akashi et al. 2005). Therefore, LaIFS and soybean HID were transiently co-expressed in tobacco leaves, and then isoflavones were

Table 1 Kinetic properties of LaIFS1 and LaIFS2 against substrate flavanones

Substrate	K_m (μM)		V_{max} (nmol/min)		V_{max}/K_m ($\times 10^{-6}$)	
	LaIFS1	LaIFS2	LaIFS1	LaIFS2	LaIFS1	LaIFS2
Liquiritigenin	59.4 ± 10.24	15.32 ± 1.71	0.55 ± 0.15	0.15 ± 0.052	9.25 ± 1.49	9.89 ± 2.45
Naringenin	219.3 ± 13.87	8.27 ± 0.40	0.13 ± 0.011	0.0065 ± 0.00041	0.61 ± 0.013	0.79 ± 0.013

Data were calculated using Lineweaver–Burk plots method based on triplicate assays. The yeast microsomes using in the enzyme assays were crude proteins in which the amound of CYP450 was not quantified. The kinetic constants of V_{max} of LaIFS1 and LaIFS2 against the same substrate cannot be compared each other. Instead, the V_{max} paramater of the same microsomal proteins on differential substrate could compared each other

clearly detected (Figs. 3d, S10 and S11). However, we found that the higher level of genistin and daidzin was accumulated in tobacco leaves than that of genistein and daidzein (Figs. 3d, S10 and S11), suggesting that LaIFS1 or LaIFS2 catalyzed naringenin and liquiritigenin into isoflavones genistein and daidzein, respectively, and then the native glycosyltransferases in tobacco catalyzed them into corresponding glucosylated compounds. These results indicated that LaIFS1 or LaIFS2 functioned as isoflavone synthases to catalyze naringenin into genistein and also catalyze liquiritigenin into daidzein although white lupin produces no 5-deoxyisoflavonoids.

Two new isoflavone-specific prenyltransferases were characterized in white lupin

Various prenylated isoflavones are accumulated in white lupin, but only LaG3'DT (LaPT1) was reported to prenylate genistein at the B-ring C-3' position to produce isowighteone (Shen et al. 2012). Here, LaG3'DT was used as a query to perform BLAST program. This search yielded six other putative flavonoid-specific prenyltransferase transcripts. Four transcripts contained full-length ORFs (Fig. S12). Among them, the protein encoded by LAGI02_4454 is a flavonol-specific prenyltransferase that prenylates flavonol aglycones at the C-8 position to produce corresponding C-8 prenylflavonols (Liu et al. 2021). Phylogenetic analysis showed the other three putative prenyltransferases were clustered together with legume flavonoid prenyltransferases, and differed from identified stilbenoid prenyltransferases in peanut and flavonoid prenyltransferases from non-legume plants (Fig. 4a). Expression pattern showed that the prenyltransferase genes were mainly expressed in roots (Fig. 4b), which is very similar with the previously identified *LaG3'DT* (Shen et al. 2012). These evidences suggested that the three new prenyltransferases possibly involved in catalyzing isoflavones into corresponding prenylisoflavones at a new carbon position (C-6 or C-8 position), which is different from LaG3'DT.

The ORFs of these three putative isoflavone-prenyltransferase genes were cloned and subsequently transformed into yeast strain. Five isoflavones (genistein, genistin, biochanin A, daidzein and formononetin) were used as putative accepters and DMAPP as donor substrates for enzymatic activity assays. The proteins encoded by LAGI02_39377 or LAGI02_879 could catalyze genistein into wighteone based on the authentic standard wighteone (Fig. 4c). No enzymatic activity of the protein encoded by LAGI02_240_48875 was observed in the in vitro assays. Therefore, these two prenyltransferases LAGI02_39377 and LAGI02_879 were renamed as LaG6DT1 (GenBank accession no. MT427394) and LaG6DT2 (GenBank accession no. MT427395), respectively. Sequence analysis showed that LaG6DT1 contains a putative transit peptide at its N-terminus, but LaG6DT2 has no transit peptide. Although there was no significant difference of the apparent V_{max}/K_m between LaG6DT1 and LaG6DT2 in our in vitro enzymatic assays using crude yeast microsomal proteins (Table 2), LaG6DT2 could convert nearly all genistein to wighteone in the reaction compared to a few of LaG6DT1 (Fig. 4c), which further confirmed the fact that the presence of N-terminal transit peptide in GmG4DT and LaG3'DT affect their enzymatic activities significantly in vitro (Akashi et al. 2009; Shen et al. 2012).

Further, tobacco transient expression system was used to analyze their in vivo function, and wighteone was not detected from tobacco leaves, overexpressed *LaG6DT1* or *LaG6DT2* and infiltrated with genistein, by HPLC method. Then wighteone was detectable when the samples were subjected to determine using UPLC-MS/MS technology (Fig. S13). The wighteone content from tobacco samples overexpressed *LaG6DT1* and *LaG6DT2* was significantly higher than that of negative control (Fig. 4d). Our results suggested that both LaG6DT1 and LaG6DT2 are the *bona fide* prenyltransferases in white lupin.

Two glucosyltransferases catalyzing isoflavonoids to 7-*O*-glucoside were characterized in white lupin

Previous studies showed that the glycosylation, which affects metabolic activity and bioavailability and contributes to the extraordinary diversity of flavonoid compounds, is the most frequent chemical modification in plant flavonoid metabolism (Vogt and Jones 2000), and in *Lupinus* plants, isoflavonoids have common glucosylation modification (Kachlicki et al. 2005). To uncover uridine diphosphate-dependent glycosyltransferase (*UGT*) genes involved in isoflavonoid metabolism in white lupin, we searched and cloned two glucosyltransferase genes using genistein 7-*O*-glucosyltransferase UGT88E3 (Noguchi et al. 2007) as a bait. Phylogenetic analysis showed that these two glucosyltransferases (encoded by LAGI02_3325 and LAGI02_13807) were clustered together with soybean UGT88 family (Fig. 5a) that glucosylated flavonoid aglycones (Funaki et al. 2015). The qRT-PCR analysis showed that LAGI02_3325 (referred as *LaUGT1*, GenBank accession No. MT427396) and LAGI02_13807 (*LaUGT2*, GenBank accession No. MT427397) transcripts were detected in roots, stems and leaves, but they had a different expression pattern that *LaUGT1* was expressed higher in leaves and stems, and *LaUGT2* was higher in roots (Fig. 5b). The recombinant proteins of these two glucosyltransferases were incubated with flavonoid aglycones and UDP-glucose, and catalyzed genistein into genistin (Fig. 5c). In addition, they also used wighteone as substrate to produce wighteone 7-*O*-glucoside (Fig. 5d). To further determine whether these two glucosyltransferases have the same function in vivo. LaUGT1 or

LaUGT2, and genistein or wighteone were infiltrated into tobacco leaves. Genistin or wighteone 7-*O*-glucoside was detected from extracts of inoculated leaves. However, no more genistin or wighteone 7-*O*-glucose was accumulated in tobacco leaves co-infiltrated with *LaUGTs* compared with the control (Figs. S14 and S15), suggesting that tobacco (*N. benthamiana*) has endogenous glucosyltransferase(s) metabolized isoflavones into glycosylated isoflavones, which is further supported by our results that genistin was observed in the *LaIFS*-overexpressed tobacco leaves (Fig. 3d). These evidences suggested that these two glucosyltransferases catalyzed isoflavone aglycones into its glucosides in white lupin.

Discussion

No 5-deoxyisoflavonoids in *Lupinus* species were due to lack of CHRs and type II CHIs

Isoflavonoids are divided into two groups based on the presence or absence of C–OH at C-5 position of A-ring (Fig. S1). To date, most of isoflavonoids are found in legumes and more than half of leguminous flavonoids are 5-deoxyflavonoids (Hegnauer and Gpayer-Barkmeijer 1993). In general, 5-hydroxyisoflavonoids and 5-deoxyisoflavonoids are believed to widely distribute in all legume species, which is attributed to the acquirement of related genes in the isoflavonoid pathway before the divergence of the Fabaceae (Shimada et al. 2003).

However, our experimental evidences (Figs. S2 and S3) suggested that 5-deoxyisoflavonoids were not biosynthesized in *Lupinus* species. Our data are consistent with those from metabolomics analysis of isoflavonoids in *L. angustifolius* and *L. albus* (Bednarek et al. 2001; Katagiri et al. 2000; Wojakowska et al. 2013). CHR is the first committed enzyme to produce 5-deoxyflavanone (Welle and Grisebach 1988), and the *bona fide* CHRs are strictly defined in the subgroup A in soybean CHR family (Mameda et al. 2018). The phylogenetic tree analysis suggested that all the previously reported *bona fide* CHRs are in the same subgroup (Fig. S4) as previously described (Mameda et al. 2018). In this study, we found that there were no *bona fide* CHR members in *Lupinus* species. The absence of a true CHR provides the first evidence to support our experiment observation that *Lupinus* plants biosynthesize 5-hydroxyisoflavonoids, but not 5-deoxyisoflavonoids.

Legume-specific Type II CHIs is the second committed enzyme to produce 5-deoxyflavanone (Dixon et al. 1988; Ralston et al. 2005; Shimada et al. 2003). LaCHI1 and LaCHI2 in white lupin, and the two CHI homologs from *L. angustifolius* display more sequence identity with Type I CHI than Type II (Table S3). Further, our experiments showed that LaCHI1, LaCHI2, and Type I CHI GmCHI2 (Glyma.20G241700) (Ralston et al. 2005), catalyzed naringenin chalcone to naringenin, whereas Type II CHI GmCHI1A (Glyma.20G241500) converted naringenin chalcone and isoliquiritigenin to corresponding flavanone naringenin and liquiritigenin, respectively (Figs. 2, S6 and S8). Based on the fact that Type I CHIs convert naringenin chalcone to naringenin, whereas Type II CHIs catalyze naringenin chalcone and isoliquiritigenin into naringenin and liquiritigenin, respectively (Shimada et al. 2003), LaCHI1 and LaCHI2 functioned as Type I CHIs. These results suggested that there are no Type II CHI enzymes in white lupin.

In a previous study, four *CHI* genes of *Lotus japonicus* form a tandem cluster, and three of them are Type II CHIs, and they evolved from an ancestral CHI by gene duplication (Shimada et al. 2003). To further understand CHI clusters evolution, we searched other three representative genomes of legume species: *G. max*, *M. truncatula* and *Arachis ipaensis*, and found that *CHI* genes in these three species also formed a tandem cluster, although the structure of CHI clusters has fine distinction between each other. In contrast, no similar CHI cluster is presence in *L. angustifolius* genome and the recently published white lupin genome (Fig. S16) (Hufnagel et al. 2020). These data indicated that *CHI* cluster contained Type I and II CHI genes possibly evolved after the divergence of *Lupinus* species, which is in a legume clade-the genistoids, one of the earliest diverging lineages of the Papilionoideae about 50–60 million years ago (Cannon et al. 2015).

Table 2 Kinetic properties of LaG6DT1 and LaG6DT2 against substrate genistein

Substrate	K_m (μM)		V_{max} (nmol/min)		V_{max}/K_m ($\times 10^{-6}$)	
	LaG6DT1	LaG6DT2	LaG6DT1	LaG6DT2	LaG6DT1	LaG6DT2
Genistein	46.70 ± 7.21	3519.67 ± 207.33	0.025 ± 0.0030	2.36 ± 0.14	5.57 ± 0.18	6.72 ± 0.0034

Data were calculated using Lineweaver–Burk plots method based on triplicate assays. The yeast microsomes using in the enzyme assays were crude proteins in which the amound of prenyltransferase was not quantified. The kinetic constants of V_{max} of LaG6DT1 and LaG6DT2 against the same substrate cannot be compared each other

LaIFS1 and LaIFS2 had catalytic ability toward naringenin and liquiritigenin

LaIFS1 and LaIFS2 showed isoflavone synthase activity toward substrate naringenin (Fig. 3), the precursor of 5-hydroxyisoflavonoid pathway. Here, we found that they also had catalytic ability toward substrate liquiritigenin (Figs. S9 and S11), a precursor for 5-deoxyisoflavone biosynthesis, although liquiritigenin was not able to be biosynthesized in white lupin. More importantly, based on the result that an order higher of V_{max}/K_m values for liquiritigenin than naringenin were obseverd in the in vitro IFS assays (Table 1), LaIFSs prefer to use liquiritigenin rather than naringenin as substrates, which is consistent with soybean IFSs (Jung et al. 2000). Therefore, we hypothesized that for the species containing 5-deoxyisoflavonoid biosynthesis pathway, it is not very necessary to evolve a specific IFS for catalyzing liquiritigenin because all IFS showed enzymatic activities to liquiritigenin.

Two new prenyltransferases and two new glycosyltransferases are identified in white lupin

The isoflavone aglycones are catalyzed into isoflavone conjugates by glucosyltransferases, prenyltransferases, malonyltransferases and others. The glucosylation and prenylation in white lupin isoflavonoid profiles are the most frequent modifications. LaG3′DT is able to catalyze C-3′ prenylation of genistein (Shen et al. 2012). Recently, LaPT2 was identified and functioned as a flavonol prenyltransferase in our assays (Liu et al. 2021). In this study, additional three putative prenyltransferase genes were identified in white lupin, and two prenyltransferases LaG6DT1 and LaG6DT2 functioned as genistein C-6 prenyltransferases (Fig. 4), whereas another one putative prenyltransferase did not use any isoflavones as substrates. To date, althought wighteone, lupiwighteone and isowighteone (C-6, C-8, C-3′ prenylation of genistein) were detected in white lupin, only prenyltransferases corresponding to C-6 and C-3′ position are characterized in present (Fig. 4) and previous study (Shen et al. 2012). The genes involved in genistein C-8 prenyltransferases are not isolated either from white lupin or other species. We also retrieved several putative prenyltransferase genes in the published white lupin genomes and further investigation will be needed to determine whether these genes functioned as genistein C-8 prenyltransferases.

Glucosylation isoflavonoid compounds are detected in every tested tissues in *Lupinus* species (Aisyah et al. 2016; Bednarek et al. 2001; Kachlicki et al. 2005). Two glycosyltransferases (LaUGT1 and LaUGT2) were identified in white lupin, and converted genistein to genistein 7-*O*-glucosyltransferases, and also converted wighteone to 7-*O*-glucosylwighteone (Fig. 5). Based on these results, we speculated that prenylation modification of isoflavonoids may be prior to its glucosylation modification. However, we found that tobacco is not a suitable system to confirm the in vivo function of the two glycosyltransferases because native glycosyltransferases in tobacco catalyzed genistein and wighteone into corresponding glucosylated compounds. This finding shows that plant glycosyltransferases related to flavonoid metabolism may be promiscuous as observed in the tobacco system for studying other small molecular metabolism (Polturak et al. 2018; Torrens-Spence et al. 2018).

Conclusion

In the present study, only 5-hydroxyisoflavonoids (genistein and its conjugates), no 5-deoxyisoflavonoids, were detected in every tested *Lupinus* species, which is due to lack of CHRs and Type II CHIs. The isoflavone synthases LaIFS1 and LaIFS2 converted substrates naringenin to 5-hydroxyisoflavone (genistein). In addition, LaG6DT1 and LaG6DT2 were characterized and prenylated genistein at the C-6 position into wighteone. Two glucosyltransferases LaUGT1 and LaUGT2 metabolized genistein and wighteone into its 7-*O*-glucosides. Our study not only revealed that the exclusive 5-hydroxyisoflavonoids do exist in *Lupinus* species, but also identified key enzymes in the biosynthesis pathway.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00299-021-02818-x.

Acknowledgements This work was supported by the National Natural Science Foundation of China (31870281 and 32160060), and the Agricultural Science and Technology Innovation Program (ASTIP-IAS10). We are grateful to Professor Kazafumi Yazaki (Kyoto University) for yeast vector pDR196GW, Professor Guodong Wang (Institute of Genetics and Developmental Biology, the Chinese Academy of Sciences, Beijing, China) for providing yeast strain DD104 and an engineered yeast strain DD104 that had accumulated DMAPP.

Author contribution statement JL and WJ planned and designed the research. JL performed experiments. JL and WJ analyzed the data, and wrote the manuscript.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Aisyah S, Gruppen H, Madzora B, Vincken JP (2013) Modulation of isoflavonoid composition of *Rhizopus oryzae* elicited soybean (*Glycine max*) seedlings by light and wounding. J Agric Food Chem 61:8657–8667

Aisyah S, Vincken JP, Andini S, Mardiah Z, Gruppen H (2016) Compositional changes in (iso)flavonoids and estrogenic activity of three edible *Lupinus* species by germination and *Rhizopus*-elicitation. Phytochemistry 122:65–75

Akashi T, Aoki T, 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

Akashi T, Aoki T, Ayabe S (2005) Molecular and biochemical characterization of 2-hydroxyisoflavanone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis. Plant Physiol 137:882–891

Akashi T, Sasaki K, Aoki T, Ayabe S, Yazaki K (2009) Molecular cloning and characterization of a cDNA for pterocarpan 4-dimethylallyltransferase catalyzing the key prenylation step in the biosynthesis of glyceollin, a soybean phytoalexin. Plant Physiol 149:683–693

Arnoldi A, Boschin G, Zanoni C, Lammi C (2015) The health benefits of sweet lupin seed flours and isolated proteins. J Funct Foods 18:550–563

Bednarek P, Franski R, Kerhoas L, Einhorn J, Wojtaszek P, Stobiecki M (2001) Profiling changes in metabolism of isoflavonoids and their conjugates in *Lupinus albus* treated with biotic elicitor. Phytochemistry 56:77–85

Birt DF, Hendrich S, Wang WQ (2001) Dietary agents in cancer prevention: flavonoids and isoflavonoids. Pharmacol Ther 90:157–177

Bomati EK, Austin MB, Bowman ME, Dixon RA, Noel JP (2005) Structural elucidation of chalcone reductase and implications for deoxychalcone biosynthesis. J Biol Chem 280:30496–30503

Botta B, Delle Monache G, Menendez P, Boffi A (2005) Novel prenyltransferase enzymes as a tool for flavonoid prenylation. Trends Pharmacol Sci 26:606–608

Boue SM, Burow ME, Wiese TE, Shih BY, Elliott S, Carter-Wientjes CH, McLachlan JA, Bhatnagar D (2011) Estrogenic and antiestrogenic activities of phytoalexins from red kidney bean (*Phaseolus vulgaris* L.). J Agric Food Chem 59:112–120

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal Biochem 7:248–254

Cannon SB, McKain MR, Harkess A, Nelson MN, Dash S, Deyholos MK, Peng Y, Joyce B, Stewart CN Jr, Rolf M, Kutchan T, Tan X, Chen C, Zhang Y, Carpenter E, Wong GKS, Doyle JJ, Mack JL (2015) Multiple polyploidy events in the early radiation of nodulating and nonnodulating legumes. Mol Biol Evol 32:193–210

Dixon RA, Richard Blyden E, Robbins MP, Van Tunen AJ, Mol JN (1988) Comparative biochemistry of chalcone isomerases. Phytochemistry 27:2801–2808

Drummond CS, Eastwood RJ, Miotto ST, Hughes CE (2012) Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): testing for key innovation with incomplete taxon sampling. Syst Biol 61:443–460

Farag MA, Huhman DV, Dixon RA, Sumner LW (2008) Metabolomics reveals novel pathways and differential mechanistic and elicitor-specific responses in phenylpropanoid and isoflavonoid biosynthesis in *Medicago truncatula* cell cultures. Plant Physiol 146:387–402

Feng CY, Li SS, Taguchi G, Wu Q, Yin DD, Gu ZY, Wu J, Xu WZ, Liu C, Wang LS (2021) Enzymatic basis for stepwise C-glycosylation in the formation of flavonoid di-C-glycosides in sacred lotus (*Nelumbo nucifera Gaertn.*). Plant J 106:351–365

Fischer MJ, Meyer S, Claudel P, Bergdoll M, Karst F (2011) Metabolic engineering of monoterpene synthesis in yeast. Biotechnol Bioeng 108:1883–1892

Funaki A, Waki T, Noguchi A, Kawai Y, Yamashita S, Takahashi S, Nakayama T (2015) Identification of a highly specific isoflavone 7-*O*-glucosyltransferase in the soybean (*Glycine max* (L.) Merr.). Plant Cell Physiol 56:1512–1520

Hane JK, Ming Y, Kamphuis LG, Nelson MN, Garg G, Atkins CA, Bayer PE, Bravo A, Bringans S, Cannon S, Edwards D, Foley R, Gao LL, Harrison MJ, Huang W, Hurgobin B, Li S, Liu CW, McGrath A, Morahan G, Murray J, Weller J, Jian J, Singh KB (2017) A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant-microbe interactions and legume evolution. Plant Biotechnol J 15:318–330

Hegnauer R, Gpayer-Barkmeijer RJ (1993) Relevance of seed polysaccharides and flavonoids for the classification of the leguminosae: a chemotaxonomic approach. Phytochemistry 34:3–16

Hufnagel B, Marques A, Soriano A, Marques L, Divol F, Doumas P, Sallet E, Mancinotti D, Carrere S, Marande W, Arribat S, Keller J, Huneau C, Blein T, Aime D, Laguerre M, Taylor J, Schubert V, Nelson M, Geu-Flores F, Crespi M, Gallardo K, Delaux PM, Salse J, Berges H, Guyot R, Gouzy J, Peret B (2020) High-quality genome sequence of white lupin provides insight into soil exploration and seed quality. Nat Commun 11:492

Jiang WB, Yin QG, Wu RR, Zheng GS, Liu JY, Dixon RA, Pang YZ (2015) Role of a chalcone isomerase-like protein in flavonoid biosynthesis in *Arabidopsis thaliana*. J Exp Bot 66:7165–7179

Jung W, Yu O, Lau SM, O'Keefe DP, Odell J, Fader G, McGonigle B (2000) Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. Nat Biotechnol 18:208–212

Kachlicki P, Marczak L, Kerhoas L, Einhorn J, Stobiecki M (2005) Profiling isoflavone conjugates in root extracts of lupine species with LC/ESI/MSn systems. J Mass Spectrom 40:1088–1103

Karamac M, Orak HH, Amarowicz R, Orak A, Piekoszewski W (2018) Phenolic contents and antioxidant capacities of wild and cultivated white lupin (*Lupinus albus* L.) seeds. Food Chem 258:1–7

Katagiri Y, Ibrahim RK, Tahara S (2000) HPLC analysis of white lupin isoflavonoids. Biosci Biotechnol Biochem 64:1118–1125

Kim BG, Kim SY, Song HS, Lee C, Hur HG, Kim SI, Ahn JH (2003) Cloning and expression of the isoflavone synthase gene (IFS-Tp) from *Trifolium pratense*. Mol Cells 15:301–306

Lapcik O (2007) Isoflavonoids in non-leguminous taxa: a rarity or a rule? Phytochemistry 68:2909–2916

Li H, Ban Z, Qin H, Ma L, King AJ, Wang G (2015) A heteromeric membrane-bound prenyltransferase complex from hop catalyzes three sequential aromatic prenylations in the bitter acid pathway. Plant Physiol 167:650–659

Liu JY, Jiang WB, Xia YY, Wang XM, Shen GA, Pang YZ (2018) Genistein-specific *G6DT* gene for the inducible production of wighteone in *Lotus japonicus*. Plant Cell Physiol 59:128–141

Liu J, Xia Y, Jiang W, Shen G, Pang Y (2021) *LaPT2* gene encodes a flavonoid prenyltransferase in white lupin. Front Plant Sci 12:673337

Mameda R, Waki T, Kawai Y, Takahashi S, 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

Mierziak J, Kostyn K, Kulma A (2014) Flavonoids as important molecules of plant interactions with the environment. Molecules 19:16240–16265

Morita Y, Takagi K, Fukuchi-Mizutani M, Ishiguro K, Tanaka Y, Nitasaka E, Nakayama M, Saito N, Kagami T, Hoshino A, Iida S (2014) A chalcone isomerase-like protein enhances flavonoid production and flower pigmentation. Plant J 78:294–304

Narozna D, Ksiazkiewicz M, Przysiecka L, Kroliczak J, Wolko B, Naganowska B, Madrzak CJ (2017) Legume isoflavone synthase genes have evolved by whole-genome and local duplications yielding transcriptionally active paralogs. Plant Sci 264:149–167

Ngaki MN, Louie GV, Philippe RN, Manning G, Pojer F, Bowman ME, Li L, Larsen E, Wurtele ES, Noel JP (2012) Evolution of the chalcone-isomerase fold from fatty-acid binding to stereospecific catalysis. Nature 485:530–533

Noguchi A, Saito A, Homma Y, Nakao M, Sasaki N, Nishino T, Takahashi S, Nakayama T (2007) A UDP-glucose:isoflavone 7-*O*-glucosyltransferase from the roots of soybean (*Glycine max*) seedlings. purification, gene cloning, phylogenetics, and an implication for an alternative strategy of enzyme catalysis. J Biol Chem 282:23581–23590

Oguro S, Akashi T, Ayabe S, Noguchi H, Abe I (2004) Probing biosynthesis of plant polyketides with synthetic *N*-acetylcysteamine thioesters. Biochem Biophys Res Commun 325:561–567

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45

Polturak G, Heinig U, Grossman N, Battat M, Leshkowitz D, Malitsky S, Rogachev I, Aharoni A (2018) Transcriptome and metabolic profiling provides insights into betalain biosynthesis and evolution in *Mirabilis jalapa*. Mol Plant 11:189–204

Pompon D, Louerat B, Bronine A, Urban P (1996) Yeast expression of animal and plant P450s in optimized redox environments. Methods Enzymol 272:51–64

Ralston L, Subramanian S, Matsuno M, 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

Reynaud J, Guilet D, Terreux R, Lussignol M, Walchshofer N (2005) Isoflavonoids in non-leguminous families: an update. Nat Prod Rep 22:504–515

Sasaki K, Tsurumaru Y, Yamamoto H, Yazaki K (2011) Molecular characterization of a membrane-bound prenyltransferase specific for isoflavone from *Sophora flavescens*. J Biol Chem 286:24125–24134

Secco D, Shou HX, Whelan J, Berkowitz O (2014) RNA-seq analysis identifies an intricate regulatory network controlling cluster root development in white lupin. BMC Genom 15:230

Shen GA, Huhman D, Lei ZT, Snyder J, Sumner LW, Dixon RA (2012) Characterization of an isoflavonoid-specific prenyltransferase from *Lupinus albus*. Plant Physiol 159:70–80

Shimada N, Akashi T, Aoki T, Ayabe S (2000) Induction of isoflavonoid pathway in the model legume *Lotus japonicus*: molecular characterization of enzymes involved in phytoalexin biosynthesis. Plant Sci 160:37–47

Shimada N, Aoki T, Sato S, Nakamura Y, Tabata S, 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

Torrens-Spence MP, Pluskal T, Li FS, Carballo V, Weng JK (2018) Complete pathway elucidation and heterologous reconstitution of rhodiola salidroside biosynthesis. Mol Plant 11:205–217

Villarino CB, Jayasena V, Coorey R, Chakrabarti-Bell S, Johnson SK (2016) Nutritional, health, and technological functionality of lupin flour addition to bread and other baked products: benefits and challenges. Crit Rev Food Sci Nutr 56:835–857

Vogt T, Jones P (2000) Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci 5:380–386

Wang BH, Ternai B, Polya G (1997) Specific inhibition of cyclic AMP-dependent protein kinase by warangalone and robustic acid. Phytochemistry 44:787–796

Welle R, 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

Wojakowska A, Muth D, Narozna D, Madrzak C, Stobiecki M, Kachlicki P (2013) Changes of phenolic secondary metabolite profiles in the reaction of narrow leaf lupin (*Lupinus angustifolius*) plants to infections with *Colletotrichum lupini* fungus or treatment with its toxin. Metabolomics 9:575–589

Yazaki K, Kunihisa M, Fujisaki T, Sato F (2002) Geranyl diphosphate: 4-hydroxybenzoate geranyltransferase from *Lithospermum erythrorhizon*: cloning and characterization of a key enzyme in shikonin biosynthesis. J Biol Chem 277:6240–6246

Yin QG, Shen GA, Di SK, Fan CY, Chang ZZ, Pang YZ (2017) Genome-wide identification and functional characterization of UDP-glucosyltransferase genes involved in flavonoid biosynthesis in *Glycine max*. Plant Cell Physiol 58:1558–1572