





Evaluation of pathways to the C-glycosyl isoflavone puerarin in roots of kudzu (*Pueraria montana lobata*)

Laci M. Adolfo¹  | David Burks¹  | Xiaolan Rao²  |
 Anislay Alvarez-Hernandez¹  | Richard A. Dixon¹ 

¹BioDiscovery Institute and Department of Biological Sciences, University of North Texas, Denton, Texas, USA

²State Key Laboratory of Biocatalysis and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan, Hubei Province, China

Correspondence

Richard A. Dixon, Department of Biological Sciences, University of North Texas, 1155 Union Circle #305220, Denton, TX 76203-5017, USA.
 Email: richard.dixon@unt.edu

Funding information

Samuel Roberts Noble Foundation; Texas Department of Agriculture; Washington State University; University of North Texas

Abstract

Kudzu (*Pueraria montana lobata*) is used as a traditional medicine in China and Southeast Asia but is a noxious weed in the Southeastern United States. It produces both O- and C-glycosylated isoflavones, with puerarin (C-glucosyl daidzein) as an important bioactive compound. Currently, the stage of the isoflavone pathway at which the C-glycosyl unit is added remains unclear, with a recent report of direct C-glycosylation of daidzein contradicting earlier labeling studies supporting C-glycosylation at the level of chalcone. We have employed comparative mRNA sequencing of the roots from two *Pueraria* species, one of which produces puerarin (field collected *P. montana lobata*) and one of which does not (commercial *Pueraria phaseoloides*), to identify candidate uridine diphosphate glycosyltransferase (UGT) enzymes involved in puerarin biosynthesis. Expression of recombinant UGTs in *Escherichia coli* and candidate C-glycosyltransferases in *Medicago truncatula* were used to explore substrate specificities, and gene silencing of UGT and key isoflavone biosynthetic genes in kudzu hairy roots employed to test hypotheses concerning the substrate(s) for C-glycosylation. Our results confirm UGT71T5 as a C-glycosyltransferase of isoflavone biosynthesis in kudzu. Enzymatic, isotope labeling, and genetic analyses suggest that puerarin arises both from the direct action of UGT71T5 on daidzein and via a second route in which the C-glycosidic linkage is introduced to the chalcone isoliquiritigenin.

1 | INTRODUCTION

Kudzu, *Pueraria montana lobata*, is a creeping, climbing perennial vine. With its diversity, adaptation, and rapid growth rate, kudzu has become an invasive species in the Southeastern United States (EPPO, 2007; Jewett et al., 2003; Sun et al., 2005). It has been planted on over 13,700 ha for erosion control in the Southeastern United States, but its replacement of existing vegetation results in ecological and economic losses, with lost productivity being estimated at \$118 ha⁻¹ year⁻¹ (Britton et al., 2002; Sun et al., 2005).

Kudzu is classified in the legume family Fabaceae and belongs to the Glycininae subtribe, which also includes soybean (*Glycine max*) (Wong et al., 2011). The classification of various kudzu varieties is imprecise, with multiple scientific names being ascribed to the same plant. The morphology of these plants can be very different. However, characteristics such as leaf morphology can also vary on the same plant (Adolfo et al., 2022; Sun et al., 2005; Van der Maesen, 2002). Only *P. lobata*, sometimes referred to as *P. montana lobata* in the literature, produces a high content of puerarin in the roots (EPPO, 2007; Van der Maesen, 2002).

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Plant Direct* published by American Society of Plant Biologists and the Society for Experimental Biology and John Wiley & Sons Ltd.

The roots of kudzu (*Radix Puerariae*) have been used in Chinese herbal medicine for the treatment of fever, acute dysentery, diarrhea, diabetes, and cardiovascular diseases for more than 2000 years (Wong et al., 2011). The biological activity of *Radix Puerariae* has been attributed to its major isoflavonoid secondary metabolites, including daidzein and genistein, and their glycosides, especially daidzein 8-C-glucoside, also known as puerarin (Figure S1). The ability of these compounds to modulate blood glucose levels and treat alcoholism has been extensively studied (Lukas, 2002). Puerarin is rapidly absorbed from the intestine and exhibits beneficial effects for both cardiovascular and cerebrovascular health. It can improve blood circulation (Li et al., 2010; Meezan et al., 2005; Prasain et al., 2004; Zhang et al., 2008), cerebral microcirculation (Wu et al., 2014), and hemorheology (Shen et al., 1996). Puerarin can reduce calcification of the vascular tissue by reducing inflammation and the activity of reactive oxygen species (Liu et al., 2019) and may aid in neurological protection through similar mechanisms (Li et al., 2017; Zhang et al., 2014). The 7-O-glucoside of daidzein, daidzin (Figure S1), has been investigated as a treatment for alcoholism (Kaufman et al., 1997; Liang & Olsen, 2014).

Isoflavones are quantitatively the major secondary metabolites accumulating in kudzu roots, being present as daidzein, genistein, and formononetin and their corresponding glucosides, namely, puerarin, daidzin, genistin (genistein 7-O-glucoside), and ononin (formononetin 7-O-glucoside), respectively (He et al., 2011; Zhao et al., 2011) (Figure S1). Precursor labeling and competitive feeding studies (Inoue & Fujita, 1977), along with molecular biological studies in a range of other legume species (Akashi et al., 2005; Jung et al., 2000; Modolo et al., 2007; Shimamura et al., 2007; Steele et al., 1999), have suggested potential biosynthetic pathways to isoflavone glycosides in kudzu roots (He et al., 2011; Li et al., 2014). The critical step in isoflavone formation is the 2-hydroxylation of the C-ring of flavanone, associated with aryl migration to generate 2-hydroxyisoflavanone, catalyzed by a cytochrome P-450-dependent monooxygenase (2-hydroxyisoflavanone synthase [2-HIS], often erroneously named isoflavone synthase [IFS]) (Jung et al., 2000; Kochs & Grisebach, 1986; Steele et al., 1999) (Figure 1). The dehydration of 2-hydroxyisoflavanone catalyzed by a 2-hydroxyisoflavanone dehydratase (2-HID) then yields isoflavone (Akashi et al., 2005), although this reaction can also occur spontaneously (Figure 1).

A major unresolved question in puerarin biosynthesis is the stage at which the 8-C-glycosylation reaction occurs. A number of plant C-glycosyltransferases have been characterized. These include enzymes active in 2,4,6-trinitrotoluene detoxification (Gandia-Herrero et al., 2008) and C-glycosyl flavone formation in cereals and legumes (Brazier-Hicks et al., 2009; Hirade et al., 2015; Nagatomo et al., 2014). C-Glycosyl flavone formation involves glycosylation at the level of a 2-hydroxyflavanone intermediate, not at the level of the final product, as is usual for O-glycosylation. However, in 2015, a C-GT was characterized from *Gentiana triflora* that could directly C-glycosylate the flavone apigenin (Sasaki et al., 2015). Since then, other C-GTs have been characterized as acting on terminal pathway molecules (He et al., 2019), and a kudzu C-GT catalyzing the direct glycosylation of

daidzein at the 8-C position has been reported (Wang et al., 2017). This reaction, which parallels the formation of isoflavone 7-O-glycosides (He et al., 2011; Modolo et al., 2007), represents the simplest route for introduction of the 8-C-glycosyl group in puerarin (Figure 1, Route 3), but has yet to be proven genetically. Furthermore, early isotopically labeled precursor feeding studies had ruled out this route, instead suggesting that the C-glycosyl bond might be formed earlier in the pathway, at the level of chalcone (Figure 1, Route 1) (Inoue & Fujita, 1977). However, there is currently no enzymatic support for this model. Based on the mechanism for C-glycosylation of flavones in cereals, it is also theoretically possible that formation of puerarin may occur via C-glycosylation of 2,7,4'-trihydroxyisoflavanone (Figure 1, Route 2) or a ring-opened keto-enol tautomer derived from it (Brazier-Hicks et al., 2009; He et al., 2011; Kochs & Grisebach, 1986; Li et al., 2014). However, there are currently no data directly to support this hypothesis.

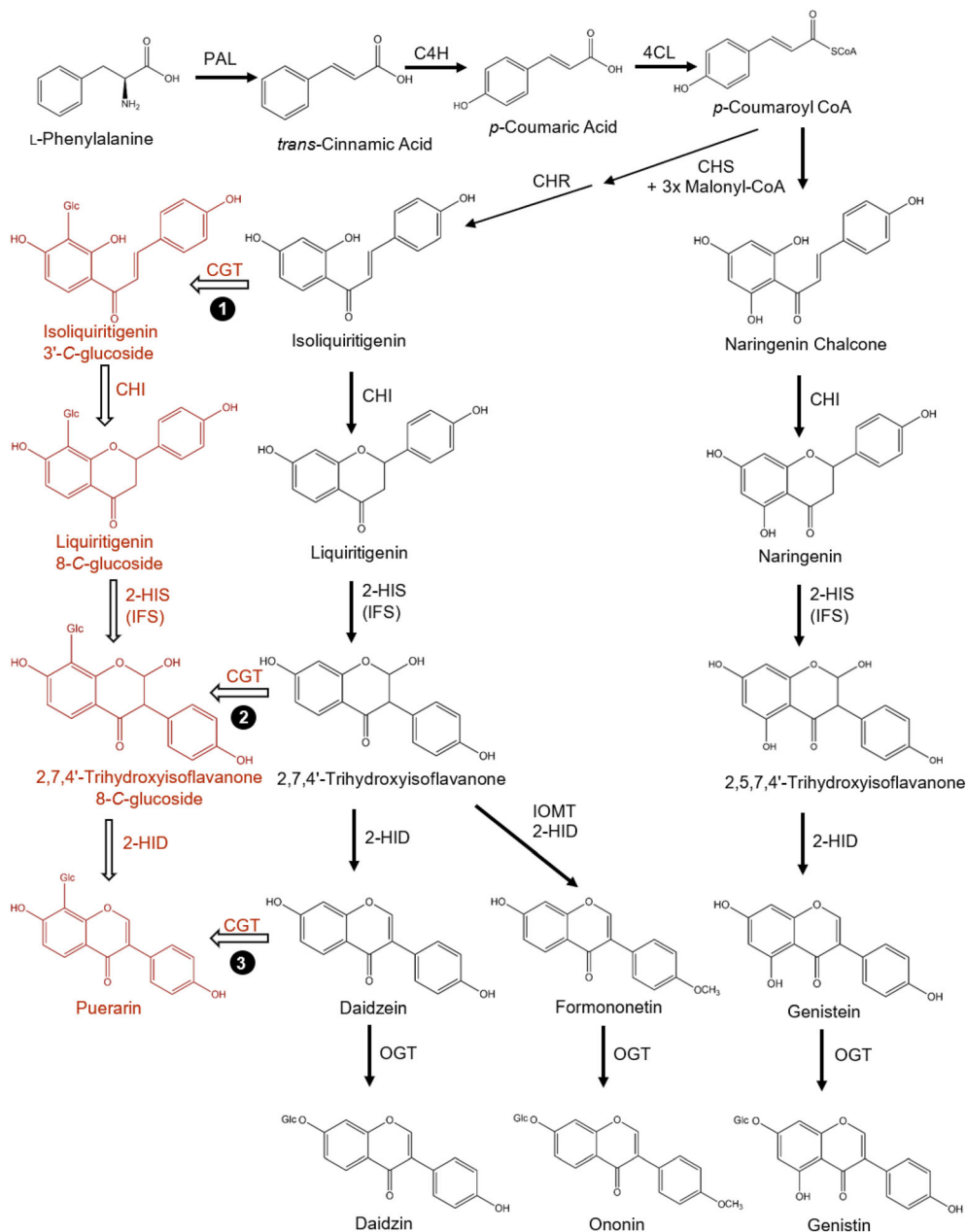
The pathways of isoflavone biosynthesis in kudzu remain uncertain in part because of the limited genomic resources in *Pueraria*. We previously reported an EST collection from a subtractive cDNA library from roots of a commercially obtained "kudzu" of unclear origin and a field-collected accession of *P. lobata*, which produced no puerarin and significant amounts of puerarin, respectively (He et al., 2011). The commercial accession was recently identified as *Pueraria phaseoloides* (Adolfo et al., 2022), and *P. lobata* has recently been renamed *P. montana lobata*, nomenclature that will be used here. Li et al. (2014) isolated 40 partial uridine diphosphate glycosyltransferase (UGT) cDNAs from kudzu and characterized one as encoding an isoflavone 7-O-glucosyltransferase; these studies were then extended to a broader RNA sequencing approach which resulted in the identification of a UGT, named UGT43, which can generate puerarin directly from daidzein in vitro (Wang et al., 2017). Through the application of de novo transcriptomics analysis and comparative transcriptomics methods, we here identify candidate genes involved in isoflavone glycosylation in kudzu and test hypotheses concerning the C-glycosylation reaction in puerarin biosynthesis using protein expression, [¹³C and ²H] precursor labeling, and RNA silencing in kudzu hairy root cultures. Our results confirm the direct C-glycosylation of daidzein by UGT71T5 as one route for puerarin biosynthesis, but also support an alternative pathway originating from C-glycosylation at the level of chalcone.

2 | RESULTS

2.1 | Flavonoid profiles of commercial and wild-collected kudzu roots

In previous studies, HPLC analysis indicated that young roots derived from field-collected *P. montana lobata* produced significant amounts of puerarin, whereas no puerarin was detected in roots of commercially obtained *Pueraria* material, subsequently identified as *P. phaseoloides* (Adolfo et al., 2022; He et al., 2011). Both accessions produced the isoflavone 7-O-glucosides daidzin, ononin, and genistin

FIGURE 1 Potential pathways leading to puerarin in kudzu roots. The enzymes are: CHS, chalcone synthase; CHR, chalcone reductase; CHI, chalcone isomerase; 2-HIS (IFS), 2-hydroxyisoflavanone synthase; IOMT, isoflavone O-methyltransferase; HID, trihydroxyisoflavanone dehydratase; GT, glucosyltransferase. (CHS and CHR co-act to generate 6'-deoxychalcone [isoliquiritigenin], rather than acting sequentially as depicted for ease of comparison with the pathway to genistein). Hollow arrows with numbers indicate possible steps for introduction of the C-glycosyl substituent of puerarin. Hypothetical reactions specific for puerarin biosynthesis are shown in red.



(Figure S1). The wild plant material from the original location (GPS coordinates given in Section 4) was reexamined for the content of isoflavones compared with that from the *P. phaseoloides* accession (Figure 2a,b). As before, puerarin was detected in root extracts from wild *P. montana lobata*, but not from commercial *P. phaseoloides*; the latter accumulated ononin as the major isoflavone conjugate.

2.2 | Identification of isoflavone biosynthetic genes in *Pueraria* species

Based on the previously conflicting conclusions regarding the origin of the C-glycosidic bond in puerarin biosynthesis, four steps in the pathway warrant additional study. The most important is the critical C-

glycosyltransferase reaction, but its place in the pathway may also have implications for the substrate specificities of chalcone isomerase (CHI), 2-HIS, and 2-HID (Figure 1). We therefore set out to obtain kudzu transcriptomes for comparative analysis, in *P. phaseoloides* and *P. montana lobata*, for potential C-GTs and contigs for CHI, 2-HIS, and 2-HID. Our previous study had identified isoflavone pathway genes in kudzu based on sequencing of a subtractive cDNA library that aimed to target only those genes that were preferentially expressed in *P. montana lobata* (He et al., 2011). The present approach was designed to also detect orthologous genes in *P. phaseoloides*.

To obtain *Pueraria* root transcriptomes for the identification of isoflavone pathway genes, roots of commercial *P. phaseoloides* and field-collected *P. montana lobata* were analyzed. Quality control and assembly of the transcriptomes for comparative genomic analyses

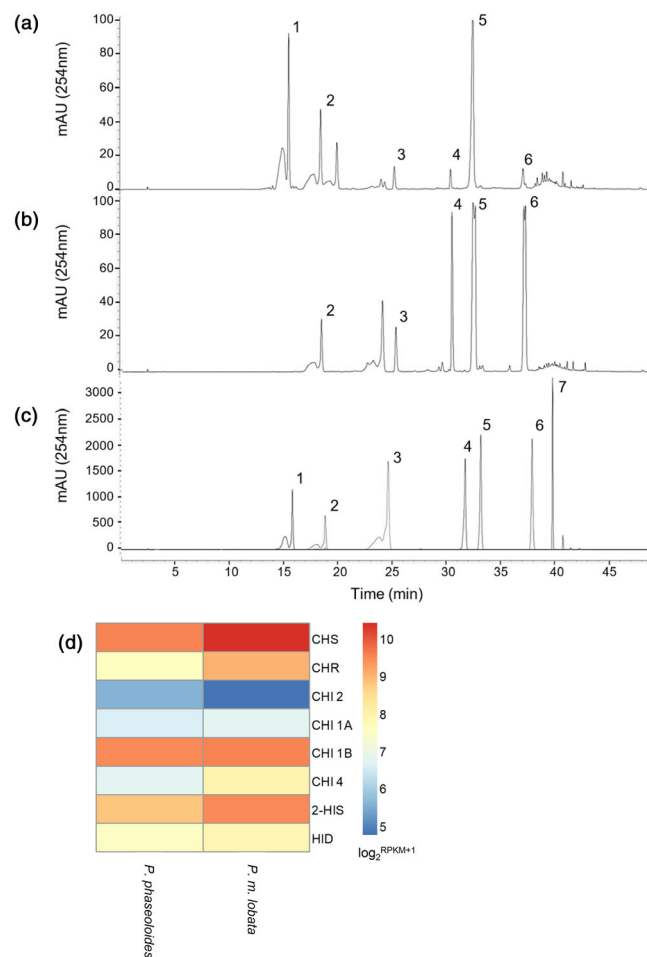


FIGURE 2 Levels of isoflavones and isoflavone pathway gene transcripts in mature root extracts from *P. m. lobata* and *P. phaseoloides*. (a–c), HPLC chromatograms showing isoflavone profiles in *P. m. lobata* (a) and *P. phaseoloides* (b) and standards (c). mAU, milli-absorbance units. 1, puerarin; 2, daidzin; 3, genistin; 4, ononin; 5, daidzein; 6, genistein; 7, formononetin. (d), transcript levels of CHS, CHR, CHI, 2-HIS and 2-HID genes in roots of *P. m. lobata* and *P. phaseoloides*, based on RNA sequencing (Adolfo et al., 2022).

between *Pueraria* species and other legumes have been presented elsewhere (Adolfo et al., 2022). Analysis of the transcriptomes led to the identification of four loci encoding CHI enzymes, with pairs of orthologs in *P. phaseoloides* and *P. montana lobata*. Two pairs of orthologs fell into the Type II class of CHIs, one pair in Type I and one pair in Type IV (Supplemental Figure S2). Only the type II CHIs had been identified by the subtraction library strategy (He et al., 2011). In contrast to CHI, only single 2-HIS (IFS) and 2-HID loci were identified. The 2-HIS proteins from *P. phaseoloides* and *P. m. lobata* were 93% identical (Supplemental Figure S3). There was a larger difference between the 2-HID proteins in *P. phaseoloides* and *P. m. lobata* (81% identical) than between the orthologs of the different CHI classes, with the 2-HID in *P. m. lobata* being more similar to a 2-HID protein recently reported from *P. candollei* var *mirifica* (Genbank accession number AFK64684.1) (Supplemental Figure S4). Tuberos roots of

P. candollei var *mirifica* also contain puerarin and have been used in Thai traditional medicine (Boonsongcheep et al., 2010).

The relative transcript levels of the various isoflavone biosynthetic enzyme genes in roots of *P. phaseoloides* and *P. m. lobata*, based on RNA sequencing analysis (Adolfo et al., 2022) are shown in Figure 2d. The entry-point enzymes CHS (for flavonoids) and 2-HIS (for isoflavonoids) were more strongly expressed in the roots of *P. m. lobata*.

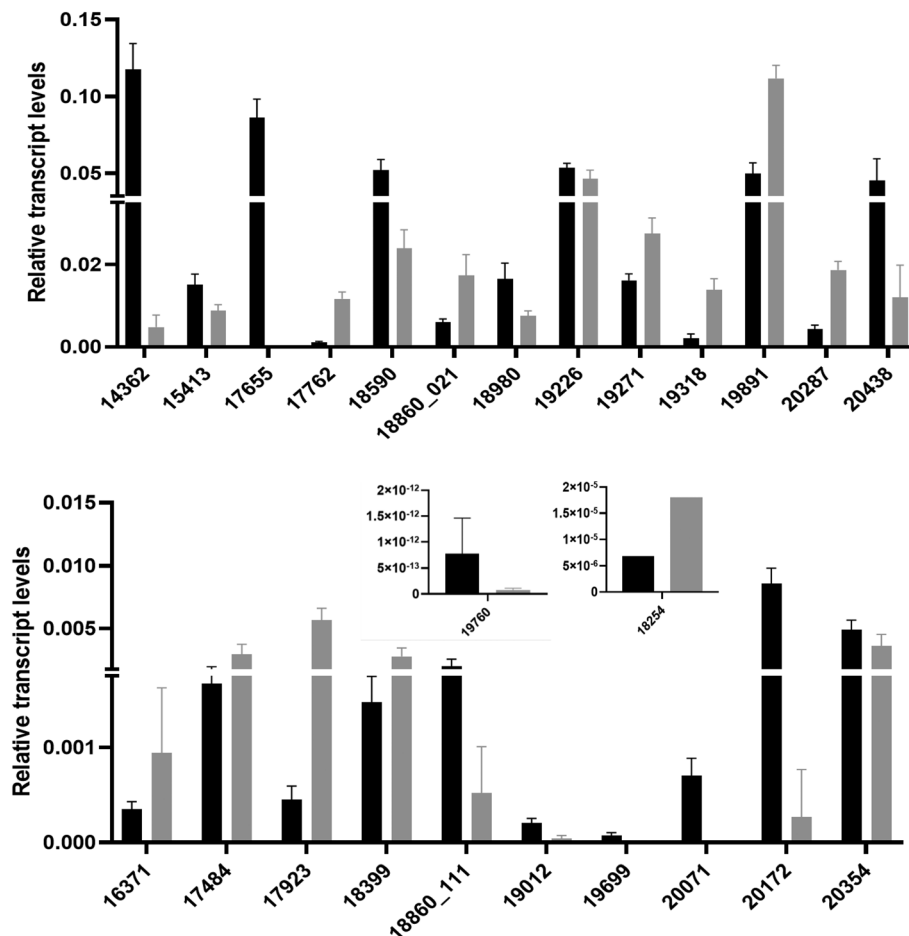
2.3 | Identification of C-glycosyltransferase candidates from *P. m. lobata*

For identification of C-glycosyltransferase candidates, a Trinity-derived *P. m. lobata* transcriptome (Grabherr et al., 2011) was processed by Transdecoder (Haas & Papanicolaou, 2018) and analyzed by Interproscan (Jones et al., 2014). Out of the sequences identified by Interproscan, 159 matched the Pfam (El-Gebali et al., 2019) ID PF00201 and 86 matched the PROSITE (Sigrist et al., 2013) ID PS00375 which contain the 44 amino acid signature (PSPG box) present in small molecule glycosyltransferases (Hughes & Hughes, 1994). The digital expression of these genes was analyzed and 25 were determined to meet at least one of two criteria we applied for potential involvement in puerarin biosynthesis: the percent difference between *P. m. lobata* and *P. phaseoloides* was higher than 250 or, if there was no expression found in *P. phaseoloides*, the expression in *P. m. lobata* was higher than ten counts (Supplemental Table S1). Of the 25 candidates, 8 were truncated and 17 were full length (1,398–1,617 bp). qPCR analysis revealed that, of the 25 candidates, 14 had higher transcript levels in roots compared to leaves of *P. m. lobata* (Figure 3).

A BLAST search on the nucleotide sequences (Morgulis et al., 2008) was next performed utilizing the NCBI nr/nt BLAST database to determine if the candidates had matches to other UGTs previously characterized from kudzu. Six had no matches to previously identified kudzu UGTs. Two (15,413 and 19,760) matched reported but unanalyzed genes from kudzu (He et al., 2011; Wang et al., 2016), and six had top BLAST hits to genes from kudzu that had previously been investigated as recombinant proteins. Of these, four had given negative results in enzyme assays, candidate 17,655 showed activity with 6,7,4-trihydroxyisoflavone (He et al., 2011; Li et al., 2014), and candidate 20,172 matched the C-GT with activity against daidzein (UGT43) (Wang et al., 2017). Candidates 19,012 and 20,071 were most closely related to 20,172. With the exception of 20,071, candidates that matched previously analyzed UGTs from kudzu with greater than 99% sequence identity, or had extremely low relative transcript levels (e.g. 19,760), were not analyzed further.

A phylogenetic analysis was then performed with previously characterized UGTs with known regiospecificity for sugar attachment to the substrate (Figure 4). Only UGTs transferring a glucose moiety to flavonoids were examined; their accession numbers are given in Supplemental Table S2. The phylogenetic tree resulted in seven clades. Candidates 20,354 and 20,438 clustered with TcCGT1 (a flavone

FIGURE 3 Relative transcript levels of UGT candidates in root and leaf tissue of *P. m. lobata*. The candidate glycosyltransferase transcript levels were adjusted to the housekeeping gene *elf5*. The black bars represent root tissue and the light gray bars represent leaf tissue. Bars show means plus standard deviations for at least three averaged technical replicates of three biological replicates. Numbers on X axis indicate UGT candidates. Note the different scale on the insets.



UGT with broad substrate- and regio-specificity, able to form C-, O-, N- and S-linked glycosides (He et al., 2019)) and PIUGT4, identified in vitro as having activity at the 7-O position with a variety of flavonoids except flavanones (Li et al., 2014). A verified C-GT (GtCGT), along with a 7-O UGT (PIUGT57), clustered with the group of UGTs shown to glycosylate flavonoids at the 5-OH position. Candidate 20,172 clustered with multi-site UGTs, and candidate 18,590 clustered with UGTs that attach a sugar to the 3-O position on the substrates. The other candidates clustered with either 7-O specific or multi-site UGTs.

The cumulative analysis of digital expression levels, relative transcript levels, BLAST results, and phylogenetic relationships narrowed the final glycosyltransferase candidates to eight that were full-length or near full-length (19,012, 18,590, 20,071, 15,413, 20,354, 19,226, 18860_11, and 20,172; 20,438 was incomplete). Their sequences were submitted for official classification, and their nomenclatures, previous history and subfamily assignments are shown in Table 1.

2.4 | Activities of recombinant UGT proteins

The UGTs in Table 1 were cloned into the protein expression vector pDEST15 with a GST tag and expressed in *E. coli*. After purification on Glutathione-Sepharose (see Figure 5a for UGT71T5), recombinant

proteins were assayed with the isoflavones daidzein, genistein and biochanin A, as well as the flavonoids isoliquiritigenin and liquiritigenin. After initial analysis with substrates from the isoflavone pathway, UGTs 72B61, 73C41, 73C42, 79B59, 88A40, 88F18, and 90A24 were analyzed with the additional flavonoids/polyphenols apigenin, luteolin, quercetin, kaempferol, coumestrol, esculetin, and resveratrol. Only three of the other recombinant proteins, UGTs 88A40, 73C42, and 72B61, showed activity against any of the above substrates (summarized in Supplemental Table S3). Among other activities, UGT88A40 catalyzed glucosylation of biochanin A (position not determined) and UGT73C42 catalyzed glucosylation of daidzein and genistein (both 7-O- and 4'-O-, but not 8-C) and biochanin A (position not determined).

UGT71T5 was confirmed as generating puerarin from daidzein (Figure 5b). The activity of recombinant UGT71T5 was then examined with the chalcone isoliquiritigenin, the flavanone liquiritigenin, and 2,7,4'-trihydroxyisoflavanone, as potential alternative substrates for C-glycosylation upstream of daidzein (Figure 1). LC-MS analysis revealed that UGT71T5 produced isoliquiritigenin glucoside(s) at a retention time between 15 and 15.5 min (Figure 5c). The product(s) were confirmed as C-glycosides following β -glucosidase digestion (Figure 5c). Enzymatic activity of UGT71T5 with daidzein and isoliquiritigenin using various substrate concentrations indicated that UGT71T5 was saturated with daidzein below 25 μ M whereas

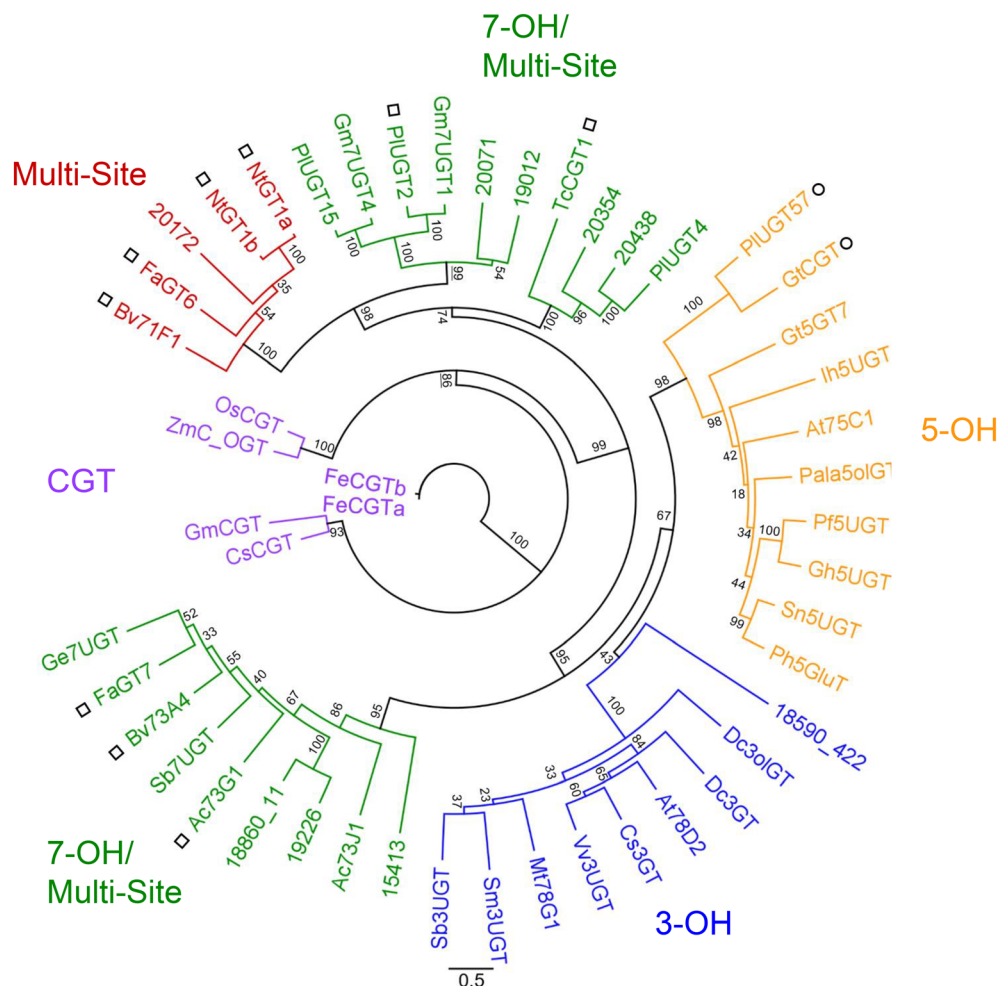


FIGURE 4 Phylogenetic relationship of candidate UGTs with known plant UGTs with different regiospecificities. *P. m. lobata* UGTs identified here are shown with number designations. A small square identifies multi-site UGTs. A small circle identifies UGTs that have been characterized as having activity different from the group with which they cluster. The tree was made from a protein alignment (MUSCLE) (Edgar, 2004) using ATGC PhymL 2.2.4 (Guindon et al., 2010) with 1,000 bootstrap replicates in Geneious prime 2022.0.1. The scale bar indicates the length of .5 substitutions. GenBank accession numbers are given in supplemental Table S2.

TABLE 1 Candidate UGTs identified in the present work

<i>P. m. lobata</i> trinity output	Nucleotide match	Query coverage (%)	Percent identity	Official UGT name	Phylogenetic group	Citation
18,590_c4_g2_i5	<i>P. lobata</i> UGT6	100	100%	UGT79B59	A	Wang et al., 2016
20,071	<i>P. lobata</i> KGT35	92	89.94%	UGT88F18	E	Li et al., 2014
	<i>P. lobata</i> GT19J14	92	89.78%			He et al., 2011
18860_c1_g1	Glycine max cDNA, clone: GMFL01-07-M14	98	92.90%	UGT73C41	D	N/A
19,012	PREDICTED: <i>Glycine soja</i> UGT 88A1-like	100	90.08%	UGT88A40	E	N/A
19,226	PREDICTED: <i>G. soja</i> UGT 73C6-like	99	93.72%	UGT73C42	D	N/A
20,354	<i>Phaseolus vulgaris</i> hypothetical protein	94	79.09%	UGT72B61	E	N/A
15,413	<i>P. lobata</i> UGT 42	100	99.22%	UGT90A24	C	Wang et al., 2016
20,172	<i>P. lobata</i> UGT 43	100	99.72%	UGT71T5	E	Wang et al., 2017
20,438	PREDICTED: <i>G. max</i> Hydroquinone glucosyltransferase	99%	92.05%	N/A	E	N/A

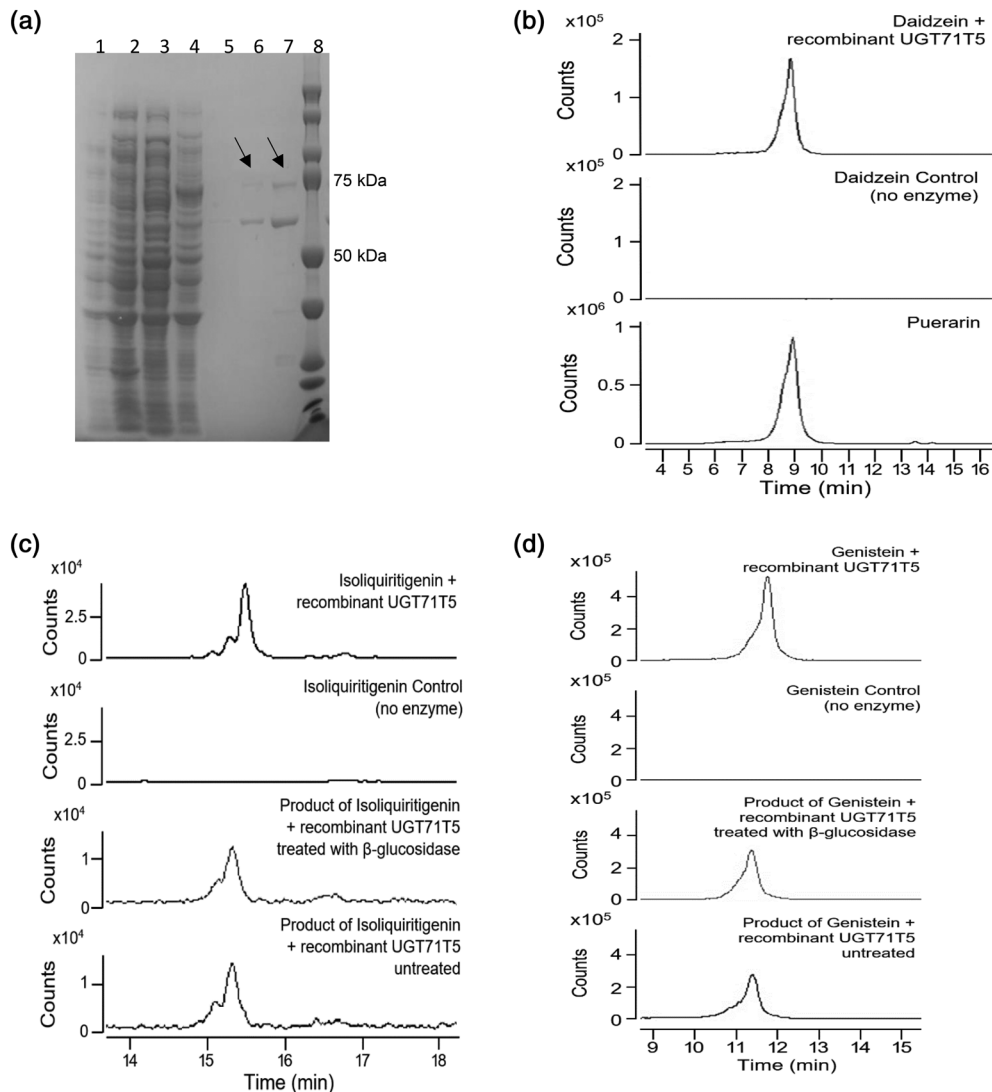
Relative transcript levels for *P. m. lobata* are shown in Figure 3 and digital expression levels are shown in Supplemental Table S1.

increased concentrations of isoliquiritigenin up to 200 μ M resulted in increased activity. UGT71T5 produced a single product from liquiritigenin, identified as an *O*-glucoside by its digestion with β -glucosidase (Supplemental Figure S5). The product had a slightly earlier retention

time than the standard liquiritin (liquiritigenin 4'-*O*-glucoside), indicating it is likely the 7-*O* glucoside.

2,7,4'-Trihydroxyisoflavanone was synthesized from liquiritigenin using recombinant 2-HIS (CYP93C20) from *M. truncatula* as described

FIGURE 5 In vitro activity of recombinant UGT71T5. (a), SDS-PAGE gel showing purification of UGT71T5. Lanes are 1, induced; 2, uninduced; 3, soluble; 4, insoluble; 5, elution 1; 6, elution 2; 7, elution 3; 8, protein MW markers. (b), extracted ion chromatogram (EIC) at m/z 415 for examination of the product of UGT71T5 with daidzein. (c), EIC at m/z 417 for examination of the product of UGT71T5 with isoliquiritigenin. (d), EIC at m/z 433 for examination of the product of UGT71T5 with genistein.



in Methods (Chang et al., 2018). On incubating the synthesized 2,7,4'-trihydroxyisoflavanone with recombinant UGT71T5, puerarin was detected (Supplemental Figure S6B, G); however, since the yield of puerarin was small, and daidzein was also detected (2,7,4'-trihydroxyisoflavanone can spontaneously dehydrate to daidzein) (Supplemental Figure S6E), the possibility that the puerarin produced was a result of the glycosylation of daidzein rather than 2,7,4'-trihydroxyisoflavanone could not be ruled out. This is consistent with our inability to detect a potential 2,7,4'-trihydroxyisoflavanone C-glycoside in these assays.

UGT71T5 generated a single product on incubation with genistein and UDP-Glc, with a molecular mass of 431 and an earlier retention time than the 7-O-glucoside genistin (Figure 5d). This product was unaffected by incubation with β -glucosidase, confirming that it was a C-glycoside (Figure 5d).

To test whether UGT71T5 could C-glycosylate a polyketide intermediate formed transiently during the CHS reaction (pre-chalcone), UGT71T5 was co-incubated with recombinant CHS in the presence of UDP-Glc and the CHS substrates malonyl CoA and 4'-coumaroyl CoA. For comparison, reactions were also set up in which UGT71T5

and UDP-Glc were added after the CHS enzyme reaction. Analysis of the reaction products on LC-MS revealed a single peak with a mass of 271 ([M-H] of naringenin and naringenin chalcone) (Supplemental Figure S7A). UV spectroscopy indicated that the product was naringenin rather than naringenin chalcone (Supplemental Figure S7B). Further analysis of the co-incubated reaction containing UGT71T5 and CHS at m/z 433.4 ([M-H] of mono-glycosylated naringenin chalcone/naringenin) revealed a cluster of peaks around 9–10 min that were not present in the single CHS reaction or the corresponding controls (Supplemental Figure S7C). When UGT71T5 was added following the CHS reaction, analysis of the EIC at m/z 433.4 showed a peak around 14–15 min that was not present in the control (Supplemental Figure S7F).

2.5 | Analysis of UGT71T5 function in *M. truncatula* hairy roots

M. truncatula is a legume rich in isoflavones (Frag et al., 2007), but does not make puerarin. To examine the activity of UGT71T5 in

planta, *Agrobacterium rhizogenes* strain A4TC24 was transformed with an expression vector carrying the UGT71T5 open reading frame and used to generate *M. truncatula* hairy roots (Figure 6a) by a previously published protocol (Boisson-Dernier et al., 2001). UGT71T5 transcripts were readily detected by qPCR in two independent *M. truncatula* hairy root lines (Figure 6b). The metabolite profiles of the roots were compared to those of control roots generated with an expression vector carrying the sequence for β -glucuronidase (GUS). LC-MS analysis showed the presence of puerarin, at a retention time of around 7 min and $m/z = 415$, in both hairy root lines (Figure 6c), with none in GUS control lines (Figure 6c). However, the amount of puerarin did not proportionally reflect the level of UGT71T5 transcripts in the two lines. Both UGT71T5-expressing hairy root lines produced more genistein 8-C-glycoside than puerarin (Figures 6d). Analysis of the same lines grown at a different time confirmed the production of puerarin (Supplemental Figure S8).

Several other compounds were differentially produced in the *M. truncatula* hairy root lines expressing UGT71T5. The molecular

mass of one was consistent with that of 3'-methoxy-puerarin (m/z 445), which is commonly found in *P. m. lobata* (Rong et al., 2002). These compounds were not analyzed further.

2.6 | Downregulation of UGT71T5 in *P. m. lobata* hairy root cultures confirms its role in puerarin biosynthesis

To date, there is no direct genetic evidence for involvement of a specific UGT in puerarin biosynthesis. To examine the in planta function of UGT71T5, we therefore developed a kudzu hairy root system. After several attempts and optimizations, the protocol described in Methods was adopted, transforming with an RNAi-construct for UGT71T5 and a comparable vector for a GUS control.

The transcript levels for UGT71T5 in the UGT71T5 knock-down lines were reduced by about 50% compared to the GUS control. The transcript levels of other flavonoid pathway genes were also reduced.

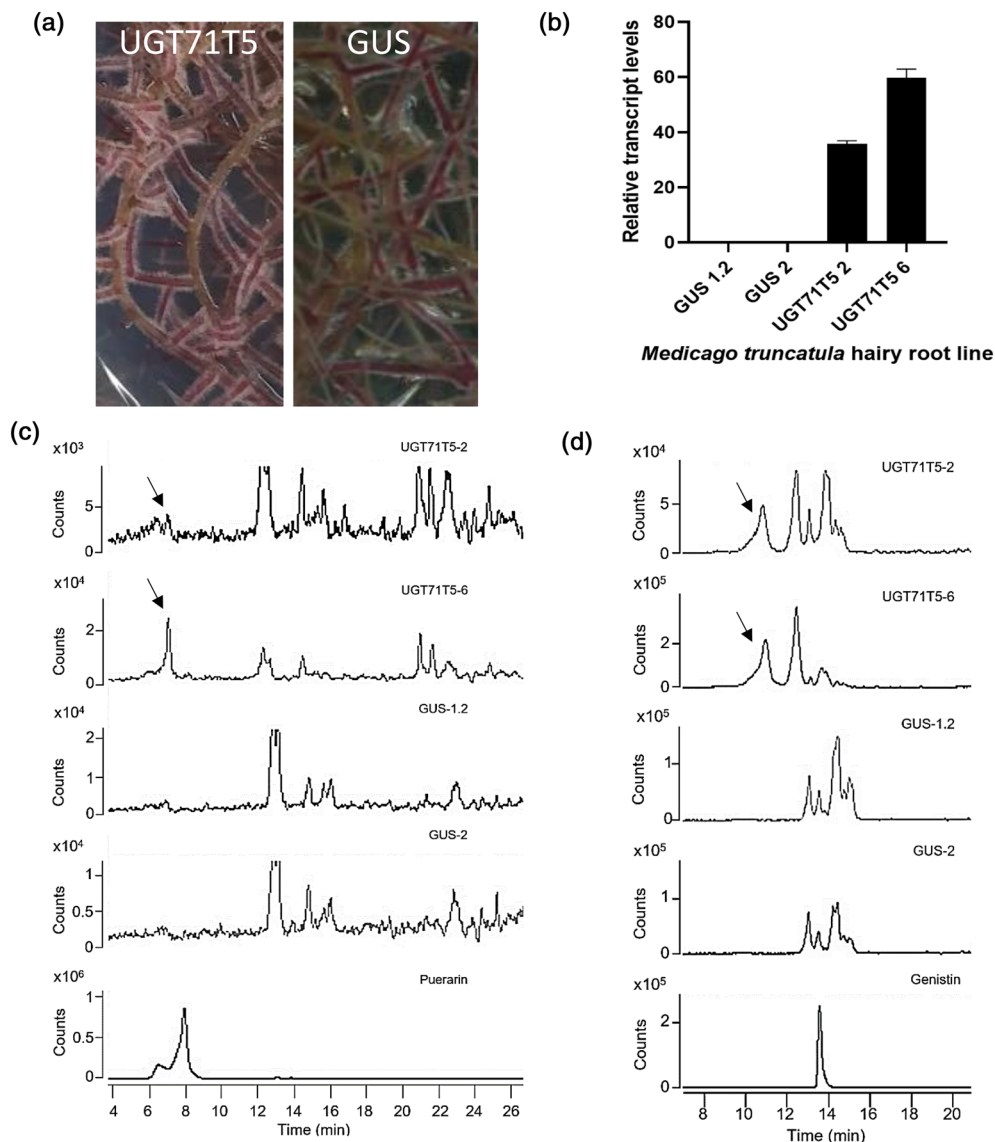


FIGURE 6 Expression of UGT71T5 in *Medicago truncatula* hairy roots. (a), appearance of *M. truncatula* hairy root cultures expressing UGT71T5. There were no observed phenotypic differences between *M. truncatula* hairy root cultures expressing UGT71T5 versus the GUS control gene. (b), relative transcript levels of UGT71T5 in *M. truncatula* hairy roots expressing UGT71T5 or GUS as compared to elf5. Bars show means and standard deviations from four technical replicates per hairy root line. (c), extracted ion chromatograms of extracts from *M. truncatula* hairy roots expressing UGT71T5 or GUS at m/z 415 for puerarin. (d), extracted ion chromatograms of extracts from *M. truncatula* hairy roots expressing UGT71T5 or GUS at m/z 431 for C-glycosyl genistein. Data for additional independent hairy root lines are given in supplemental Figure S10.

However, in the UGT71T5 knock-down lines the expression of CHI 1 was lower than CHI 2 whereas the reverse was true in the GUS controls (Figure 7a). The knock-down of UGT71T5 resulted in a strong reduction in the levels in puerarin (by approximately 3-fold, Figure 7b). Despite the reduced relative transcript levels for other flavonoid pathway genes there was no significant difference in the amount of daidzin or isoliquiritigenin; however, there was a small but significant difference in the amount of daidzein.

2.7 | Examination of the route to puerarin through precursor labeling experiments

The above results provide support for the involvement of UGT71T5 in the biosynthesis of puerarin and C-glucosyl genistein, potentially by direct glycosylation of the isoflavone aglycones. However, earlier labeling experiments did not support this pathway, instead suggesting that the C-glycosidic bond was formed at an earlier stage (Inoue & Fujita, 1977). In view of this, and the detectable activity of UGT71T5 with chalcone, flavanone and possibly 2-hydroxyisoflavanone precursors of puerarin, we decided to re-evaluate the pathway to puerarin using precursor feeding. Whereas earlier studies had examined incorporation of ^{14}C -labeled precursors, we utilized heavy isotope

[^{13}C - and ^2H]- labeled precursors to facilitate identification of labeled products.

$^{13}\text{C}_9$ -L-phenylalanine and $^2\text{H}_6$ -daidzein were obtained commercially, and $^{13}\text{C}_6$ -isoliquiritigenin was synthesized from $^{13}\text{C}_6$ -4-hydroxybenzaldehyde and 2',4'-dihydroxyacetophenone by alkaline condensation (Geissmann & Clinton, 1946) (Supplemental Figure S9A). One week-old *P. m. lobata* seedlings on float disks were exposed to labeled compounds for 3 weeks, flavonoids extracted from roots, and the relative proportions of unlabeled and newly synthesized compounds determined by mass spectrometry. Percentage of label incorporation was determined by measuring the amount of unlabeled and labeled compounds at the m/z of the native compound and m/z + 5 (for daidzein feeding), m/z + 6 (for isoliquiritigenin) and m/z + 9 (for L-Phe) (Figure S9B, C). Analysis of the labeling medium by LC-MS at the end of each experiment indicated that no labeled compounds remained in the medium after the 3-week incubation.

^{13}C -L-Phe was strongly incorporated into puerarin (60% labeled after 3 weeks), but surprisingly less into daidzin (40%), daidzein (40%), and ononin (36%) (Table 2). Labeling of isoliquiritigenin and liquiritigenin was around 30%, and formononetin only 16%. The incorporation of ^{13}C -isoliquiritigenin into upstream metabolites was surprisingly low considering its position in the pathway upstream from L-Phe (Table 2) although, as expected from either spontaneous isomerization

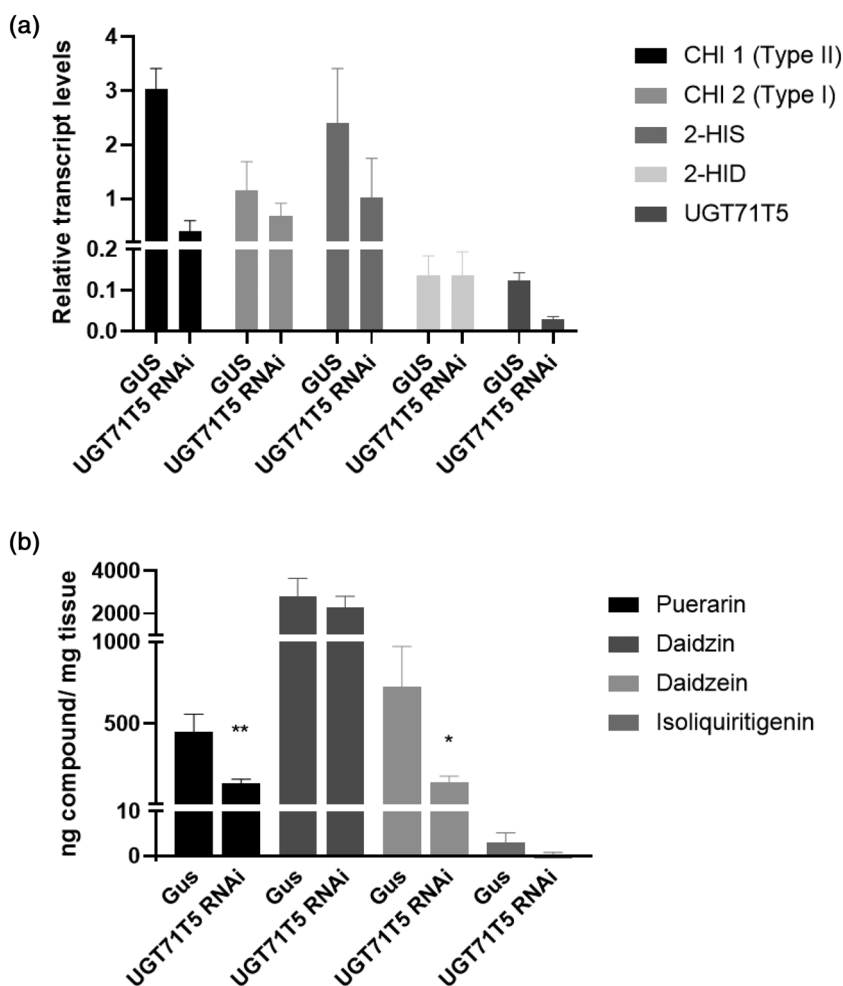


FIGURE 7 Down-regulation of UGT71T5 in *P. m. lobata* hairy roots. (a), relative transcript levels of UGT71T5 and other isoflavonoid pathway genes in UGT71T5-RNAi and GUS control lines. CHI 1, chalcone isomerase 1; CHI 2, chalcone isomerase 2; 2-HIS, 2-hydroxyisoflavanone synthase; 2-HID, 2-hydroxyisoflavanone dehydratase. Bars show means and standard deviations from at least three technical replicates. (b), (Iso)flavonoid levels in UGT71T5 knockdown lines as compared to the GUS control. Data are means from three biological replicates. * = significant at $p < .05$ and ** = significant at $p < .01$, unpaired t-test.

TABLE 2 Incorporation (%) of labeled L-phenylalanine, isoliquiritigenin, and daidzein into selected flavonoid compounds in kudzu roots

	¹³ C9-L-phenylalanine	¹³ C6-isoliquiritigenin	² H6-daidzein
Isoliquiritigenin	31.43 ± 20	80.45 ± 16.58	N/A
Liquiritigenin	31.61 ± 19.57	73.81 ± 24.61	N/A
Daidzein	41.68 ± 11.73	1.01 ± 1.14	N/D
Daidzin	42.88 ± 10.43	3.85 ± 3.78	7.25 ± 3.9
Puerarin	61.96 ± 9.16	13.30 ± 10.23	17.12 ± 4.52
Formononetin	16.42 ± 7.99	N/D	N/D
Ononin	36.84 ± 19.61	N/D	N/D
Genistein	57.39 ± 12.62	N/A	N/A
Genistin	48.42 ± 11.33	N/A	N/A

N/A indicates molecules that are not downstream of the labeled compound.

N/D indicates molecules with a % of labeled mass below the limit of quantification.

or the very high turnover number of chalcone isomerase (Jez et al., 2000), labeling of liquiritigenin was high (74%). The incorporation of ¹³C-isoliquiritigenin into daidzein and daidzin averaged only 1% and 4%, respectively, but incorporation into puerarin was paradoxically much higher at 13%. Incorporation of ¹³C-isoliquiritigenin into formononetin and ononin was below the limit of detection.

²H-Daidzein was incorporated into daidzin (7.25%) and, in contrast to the results of the earlier ¹⁴C-feeding studies, into puerarin (17.1%) (Table 2). Taken together, the labeling studies clearly support the formation of puerarin by direct C-glycosylation of daidzein by UGT71T5. However, the higher incorporation of label from both L-Phe and isoliquiritigenin into puerarin than into daidzin suggests the co-involvement of an alternative pathway that by-passes daidzein, as further discussed below.

2.8 | Silencing of 2-HIS blocks formation of puerarin and leads to accumulation of chalcone glycosides in kudzu hairy roots

If C-glycosylation in the biosynthesis of puerarin can occur at or before the chalcone/flavanone stage, down-regulation of 2-HIS would be predicted to cause accumulation of C-glycosylated chalcone/flavanone intermediates. To test this, we generated *P. m. lobata* hairy root lines (Figure 8a,b) in which the *P. m. lobata* 2-HIS gene identified above (Supplemental Figure S3) was targeted for RNAi-mediated down-regulation. The level of 2-HIS transcripts in the 2-HIS RNAi lines was severely reduced as compared to the GUS control (Figure 8c). 2-HIS down-regulation led to a large reduction in puerarin and daidzin levels, associated with a significant increase in the amount of isoliquiritigenin (Figure 8d), resulting in the roots possessing a distinct yellow coloration (Figure 8a,b). Similar results were obtained on examination of two other independent 2-HIS knock-down lines (Supplemental Figure S10).

Flavonoid extracts from the 2-HIS RNAi hairy roots were analyzed by LC-MS, scanning at m/z 417, the [M-H] of glycosylated liquiritigenin/isoliquiritigenin. Ten compounds with m/z = 417 were

present in 2-HIS RNAi hairy roots that were not present in the same quantity in the control roots (Figure 8e). To determine if any of these compounds were C-glycosylated, the extracts were incubated with β-glucosidase (Figure 8f). Five of the ten compounds remained and were therefore potentially C-glycosylated liquiritigenin or isoliquiritigenin. Although these isomers had the same molecular mass, they possessed different UV spectra (Figure 8g). The peaks at retention time 9.4 and 11 min had UV spectra similar to that of liquiritigenin, whereas the peaks at 15 and 15.2 min had UV spectra similar to that of isoliquiritigenin (Figure 8g). The peaks at 15 and 15.2 min are also consistent with the peaks produced in the in vitro enzyme reactions of UGT71T5 with isoliquiritigenin (Figure 5c). The peak at 14.7 min had a UV spectrum that suggested it contains a mixture of isoliquiritigenin and liquiritigenin conjugates (Figure 8g).

The potential chalcone/flavanone C-glycosides were further analyzed by MS/MS. The fragmentation of O-glycosylated compounds usually occurs at the linkage of the sugar whereas in the fragmentation of C-glycosides the sugar molecule is broken at two positions resulting in masses recorded as [M-H-90] and [M-H-120] (Geng et al., 2016; Prasain et al., 2003). This diagnostic pattern is seen by comparison of the MS/MS spectra of puerarin and daidzin (Supplemental Figure S11). Most of the five potential C-glycosylated compounds showed fragment ions at m/z 297, or [M-H-120], and m/z 327, or [M-H-90]. The signal for the peak corresponding to [M-H-120] was higher than that of the peak for [M-H-90] (Supplemental Figure S12). These findings confirm the presence of C-glycosides. For compounds that disappeared following β-glucosidase digestion, ions were found at [M-H] or m/z = 417, and [M-H-162] or m/z = 255, indicating that the fragmentation took place on the oxygen linkage between the aglycone and the glucose molecule (Supplemental Figure S13). Based on these results, peaks at 15 min and 15.2 min are identified as isoliquiritigenin C-glycosides and peaks 9.4 and 11 min as liquiritigenin C-glycosides.

To further investigate the origin of the C-glycosides, we compared their levels in extracts from 2-HIS RNAi, UGT71T5 RNAi and GUS control hairy roots by LC-MS analysis using a highly sensitive TripleTOF system. The approximate retention times of the peaks on

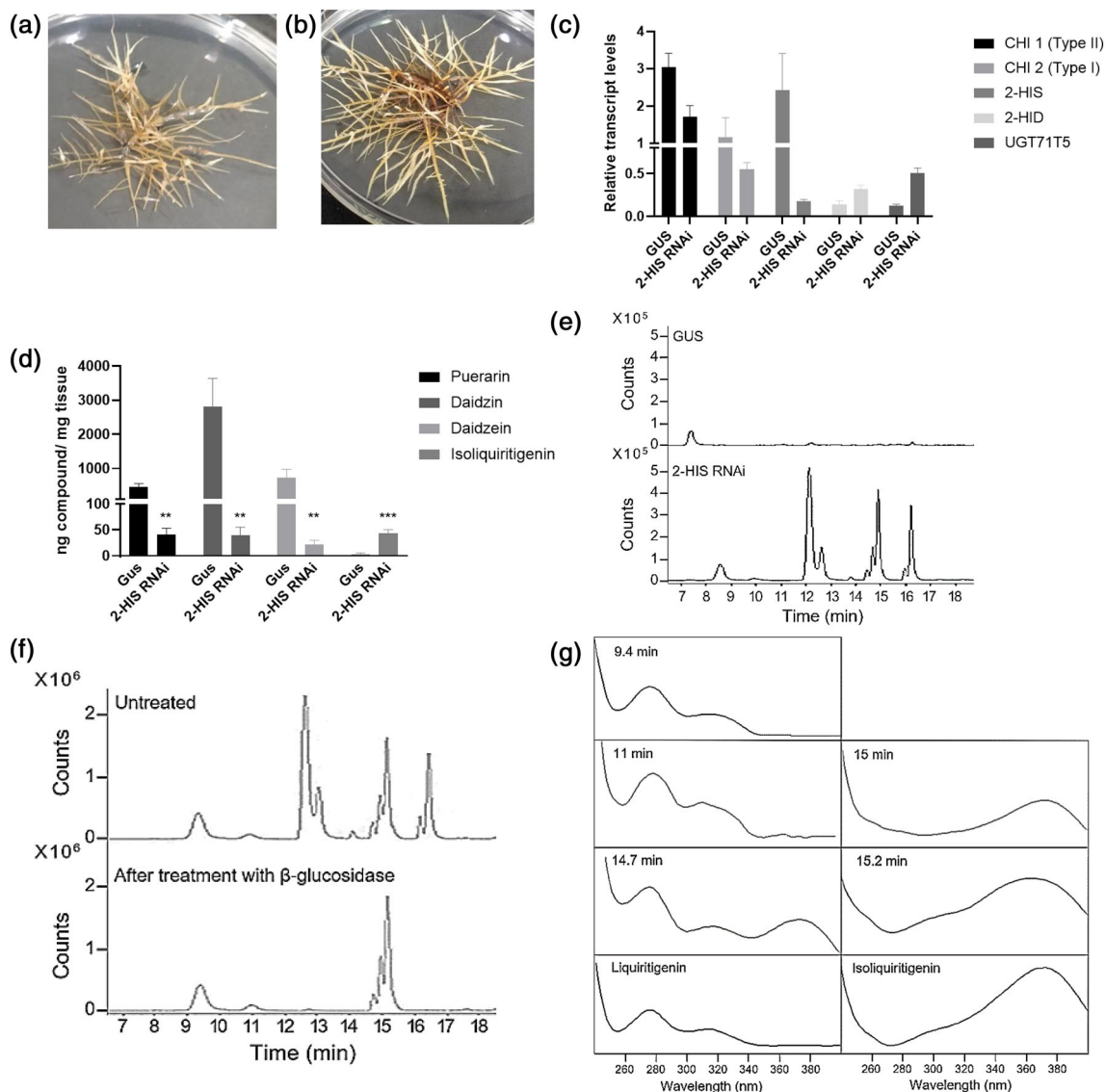


FIGURE 8 Consequences of down-regulation of 2-HIS in hairy roots of *P. m. lobata*. (a–b), phenotypic comparison of kudzu hairy roots growing on solid medium. (a), GUS control roots and (b), 2-HIS RNAi roots showing yellow coloration. (c), relative transcript levels of 2-HIS and other genes in the flavonoid pathway in 2-HIS RNAi and GUS control lines. CHI 1, chalcone isomerase 1; CHI 2, chalcone isomerase 2; 2-HIS, 2-hydroxyisoflavanone synthase; 2-HID, 2-hydroxyisoflavanone dehydratase. Data show means and standard deviations for at least three technical replicates. (d), levels of flavonoid compounds in 2-HIS RNAi lines as compared to the GUS control. Data are means from 3 biological replicates. ** = significant at $p < .01$ and *** = significant at $p \leq .0005$, unpaired t-test. (e), EICs for (iso)liquiritigenin glycosides at m/z 417 in extracts from GUS control and 2-HIS RNAi hairy roots. (f), EICs at m/z 417 after treatment of flavonoid extracts from 2-HIS RNAi hairy roots with β -glucosidase. (g), UV spectra of potential chalcone/flavanone glycosides at m/z 417 in extracts from 2-HIS RNAi hairy roots. Spectra correspond to compounds in panel F at the following retention times: 1, 9.4 min; 2, 11 min; 3, 14.7 min; 5, 15 min; 6, 15.2 min. 4 and 7 are liquiritigenin and isoliquiritigenin standards, respectively. Data for additional independent hairy root lines are given in supplemental Figure S12.

the TripleTOF system that corresponded to the suspected C-glycosylated isoliquiritigenin and C-glycosylated liquiritigenin were 6.2, 7.1, 12.7, 13.0 and 13.2 min (corresponding retention times in Figure 8f are 9.4, 11, 14.7, 15, and 15.2, respectively) (Figure 9a–c). Comparison of the EICs at m/z 417 showed a reduction in the amounts of several of the predicted C-glycosylated flavonoids in the UGT71T5 RNAi line (Figure 9b), with the relative proportions of these peaks in the UGT71T5 RNAi line being very similar, but slightly reduced compared to those in the control line (Figure 9c), and the overall peak intensities

being at least an order of magnitude lower than in the 2-HIS RNAi line. Together, these data suggest that UGT71T5 is involved in the biosynthesis of the C-glycosylated chalcones/flavanones.

Finally, to determine whether any of the potentially C-glycosylated flavanones occurring in 2-HIS RNAi hairy roots was a substrate for 2-HIS, the total flavonoid extract from 2-HIS RNAi hairy roots was incubated with an excess of recombinant 2-HIS for 16 h, and the products examined by LC–MS. 2,7,4'-Trihydroxyisoflavanone ($m/z = 271$) was formed from the activity of 2-HIS with liquiritigenin

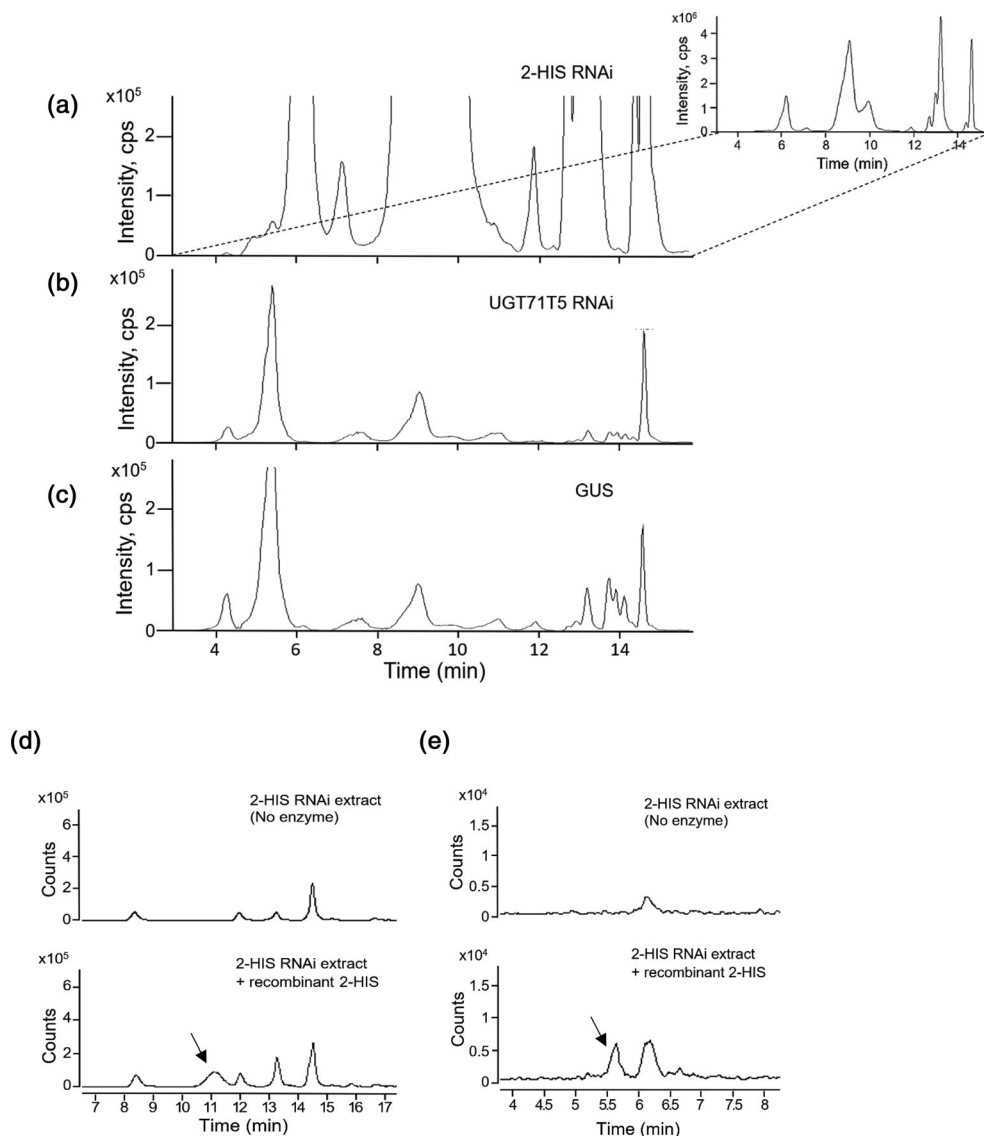


FIGURE 9 UGT71T5 is involved in C-glycosylation of early puerarin precursors in *P. m. lobata* hairy roots. (a–c), identification of chalcone/flavanone C-glycosides using a triple-TOF high resolution mass spectrometer. Extracted ion chromatograms at m/z 417 of flavonoid extracts from (a), 2-HIS RNAi; (b), UGT71T5 RNAi; (c), GUS control kudzu hairy roots. (d), extracted ion chromatogram of extract from 2-HIS RNAi hairy roots before and after incubation with recombinant 2-HIS at m/z 271 for 2-hydroxyisoflavanone. (e), extracted ion chromatogram of extract from 2-HIS RNAi hairy roots before and after incubation with recombinant 2-HIS at m/z 433 for 2-hydroxyisoflavanone C-glycoside.

present in the extracts, and a product with m/z 433 corresponding to a potential C-glycosylated 2,7,4'-trihydroxyisoflavanone was detected by LC-MS, although at significantly lower levels than 2,7,4'-trihydroxyisoflavanone itself (arrows in Figure 9d,e).

2.9 | Differential effects of 2-HID Down-regulation on Isoflavone biosynthesis in kudzu hairy roots

The requirement for 2-HID for isoflavone biosynthesis appears to differ in different plant species (Shimamura et al., 2007; Subramanian et al., 2005). To investigate whether 2-HID is critically required for C-glycosyl isoflavone biosynthesis in kudzu and determine whether a 2-HID knockdown might result in accumulation of upstream precursors, we generated 2-HID RNAi lines in kudzu hairy roots. The 2-HID transcript levels were reduced more than half in the 2-HID knockdown lines compared to the control lines (Figure 10a). The knockdown of 2-HID did not dramatically reduce the amount of puerarin or

daidzin, but there was a small but significant reduction in daidzein (Figure 10b).

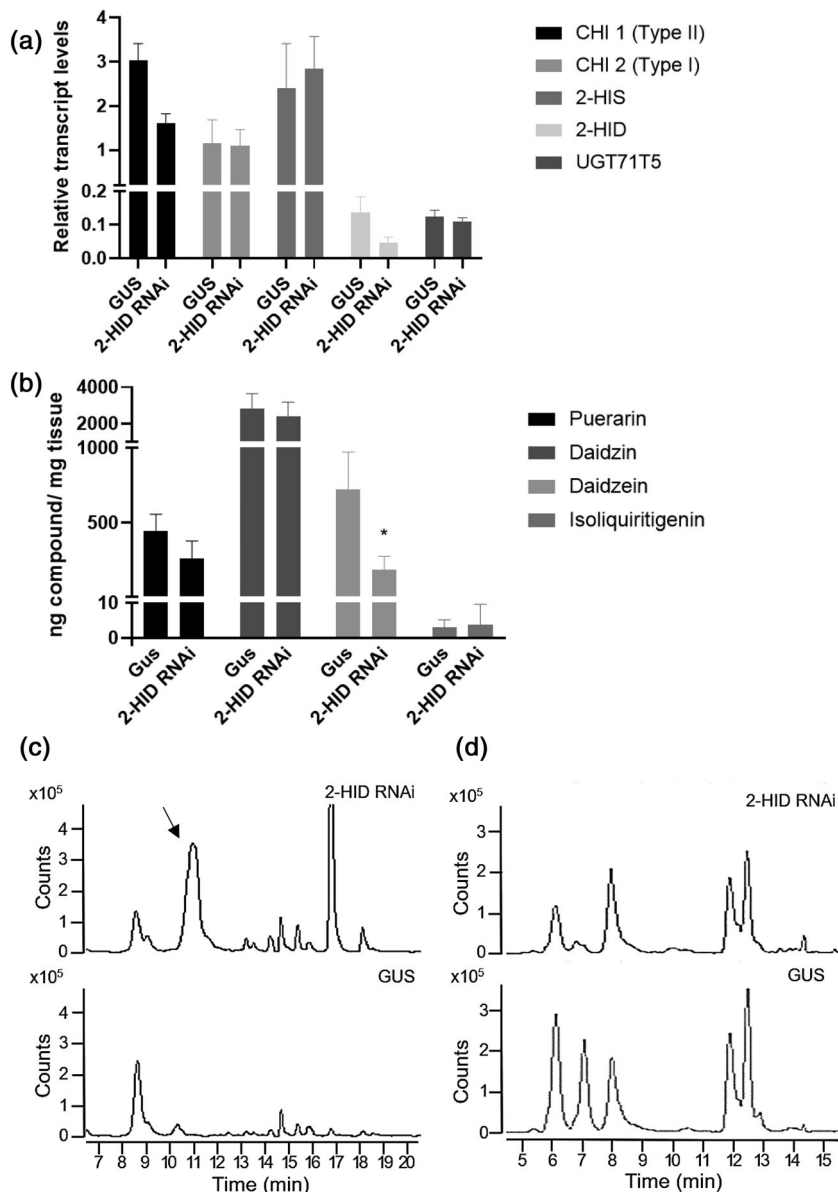
The extracts of the 2-HID RNAi and control lines were then analyzed by LC-MS at m/z 271, the [M-H] of 2,7,4'-trihydroxyisoflavanone, and at m/z 433, the [M-H] of glycosylated 2,7,4'-trihydroxyisoflavanone. There was a large increase in the amount of 2,7,4'-trihydroxyisoflavanone (peak at RT 10.5 min) in the 2-HID RNAi lines (Figure 10c), but no corresponding increase in the level of a compound that could represent a glycoside of 2,7,4'-trihydroxyisoflavanone (Figure 10d).

3 | DISCUSSION

3.1 | Experimental Systems for the Elucidation of Puerarin biosynthesis

The lack of good genetic tools has been an impediment to solving the biosynthetic pathway to puerarin. Previous work, which clearly

FIGURE 10 Down-regulation of 2-HID reduces daidzein levels in kudzu hairy roots. (a), relative transcript levels of 2-HID and other enzymes of isoflavone biosynthesis in 2-HID RNAi lines. CHI 1, chalcone isomerase 1; CHI 2, chalcone isomerase 2; 2-HIS, 2-hydroxyisoflavanone synthase; 2-HID, 2-hydroxyisoflavanone dehydratase. Data show means and standard deviations for at least three technical replicates. (b), levels of flavonoid compounds in 2-HID RNAi lines as compared to a GUS control. Data are means from three biological replicates. * = significant at $p < .05$, unpaired t-test. (c), EIC of hairy root extracts from 2-HID RNAi (top) and GUS control (bottom) at m/z 271 (for 2-hydroxyisoflavanone). (d), EIC of hairy root extracts from 2-HID RNAi (top) and GUS control (bottom) at m/z 433 (for a C-glycoside of 2-hydroxy-isoflavanone).



indicated that kudzu contains an enzyme that can directly convert daidzein to puerarin (Wang et al., 2017), did not provide conclusive proof as to whether this enzyme is totally or partially responsible for C-glycosylation in puerarin biosynthesis. Many plant UGTs act redundantly (Adiji et al., 2021) or show promiscuous activities in vitro that do not necessarily reflect their in vivo functions (Modolo et al., 2007). Although the demonstrated formation of puerarin when UGT71T5 was expressed in soybean indicates that this enzyme can contribute to puerarin biosynthesis (Wang et al., 2017), loss of function analysis is required to ascribe an in vivo role in kudzu.

In the present work, we utilized two approaches to address the in vivo function of UGT71T5. First, a transcriptomic differential expression approach was used to identify candidate UGTs, based on the availability of two kudzu “lines” that did and did not produce puerarin, respectively (He et al., 2011). Although both these “lines” were thought previously to be true kudzu, the commercial material

was recently shown to be *P. phaseoloides*, and the wild-collected material confirmed to be *P. m. lobata* (Adolfo et al., 2022). UGT71T5 was preferentially expressed in *P. m. lobata* compared to *P. phaseoloides* and was the only UGT that we could show had activity as a C-GT. Additional UGTs were identified that, at least in vitro, were capable of the 7-O-glycosylation of daidzein. Enzymes with this activity have previously been described from *G. max*, *Andrographis paniculata* and *P. m. lobata* (Funaki et al., 2015; Li et al., 2014, 2019; Wang et al., 2016). However, the *P. m. lobata* isoflavone 7-O-UGTs identified here did not include the previously reported 7-O-UGT from this species (Li et al., 2014).

For loss of function analysis, we developed a hairy root system for *P. m. lobata*. In a previous study with *P. m. lobata* hairy roots the isoflavone content was similar to that of the natural roots of the plant (Hakamatsuka et al., 1994). Hairy root transformation systems have been used to understand gene function, regulation of secondary

metabolite biosynthetic pathways, and the association of secondary metabolism with resistance to fungal pathogens in legumes (Lozovaya et al., 2007; Shimamura et al., 2007; Subramanian et al., 2005; Zhang et al., 2009). Over-expression of soybean 2-HID in *Lotus japonicus* hairy roots led to accumulation of daidzein and genistein, whereas roots over-expressing licorice 2-HIS did not accumulate isoflavones, suggesting that 2-HID is the rate limiting step in isoflavone production in *L. japonicus* (Shimamura et al., 2007); however, RNAi-mediated down-regulation of soybean 2-HIS in cotyledons reduced isoflavone levels by 60–94% and enhanced susceptibility to *Phytophthora sojae* (Subramanian et al., 2005).

Down-regulation of UGT71T5 by RNA interference in *P. m. lobata* hairy roots led to a strong decrease in puerarin levels, confirming the conclusion of Wang et al. (2017) that this enzyme is indeed involved in puerarin biosynthesis in kudzu. However, this approach does not address the step in the pathway at which the C-glycosidic bond is introduced.

3.2 | Precursor labeling studies indicate channeling of intermediates in puerarin biosynthesis

Although C-glycosylation can clearly occur at the level of daidzein based on in vitro activity of UGT71T5, such activity was not supported by previous radiolabeled precursor feeding studies, in which ^{14}C -daidzein was reported to not be a precursor of puerarin, which was preferentially labeled from ^{14}C -chalcone (Inoue & Fujita, 1977). Our use of ^{13}C - and ^2H -labeled substrates in the present work, with products analyzed by high-resolution HPLC and MS methods, allows for a more definitive determination of label incorporation into products than use of radiolabeled precursors. All precursors were effectively taken up into the roots. Our results confirmed that daidzein was indeed a precursor of puerarin, its approximately 17% incorporation being more than twice the incorporation recorded into daidzein 7-O-glucoside. However, ^{13}C -isoliquiritigenin was incorporated into puerarin to a similar extent (13%), but only weakly into daidzin (3%). Remarkably, however, L-Phe was much more efficiently incorporated than the other precursors into both puerarin and daidzin, with higher incorporation into puerarin (62%) than into its potential substrate daidzein.

Metabolic channeling has been reported at multiple stages of isoflavone biosynthesis from L-Phe, starting with the interaction of PAL and C4H (an ER-membrane bound enzyme) at the ER. This channeling was first suggested based on localization data from microsomal preparations and subcellular fractionation, as well as from metabolite labeling studies (Wagner & Hrazdina, 1984). In *Arabidopsis*, CHI and CHS were proposed to interact based on several assays including yeast two-hybrid and immunoprecipitation, whereas in rice they did not interact based on yeast two-hybrid assays (Burbulis & Winkel-Shirley, 1999; Shih et al., 2008). In soybean, CHS, CHI, and CHR (required for synthesis of 5-deoxyflavonoids) were all found to interact with 2-HIS, but not with each other (Mameda et al., 2018). The evidence of channeling in the flavonoid/isoflavonoid pathway in other

plants, coupled with the higher incorporation of L-Phe into puerarin than its more immediate precursors isoliquiritigenin and daidzein, as well as the higher incorporation of the labeled compounds into puerarin than daidzin, all support the idea that channeling or some other form of metabolic compartmentation of the isoflavone pathway is present in kudzu.

3.3 | Evidence for C-glycosylation before the formation of daidzein

If daidzein were the only substrate for C-glycosylation, the incorporation of any downstream precursor into puerarin should be less than or equal to that into daidzein. As this is not the case, our data suggest that an additional point for introduction of the C-glycosyl group occurs, most likely at the level of chalcone. Thus, down-regulation of 2-HIS results in accumulation of a mixture of C-glycosyl chalcones and flavanones, the latter presumably arising from the former by the action of CHI or spontaneous isomerization. These compounds are present in small amounts in control hairy roots and their levels are decreased on down-regulation of UGT71T5. Based on our preliminary analysis, UGT71T5 has a higher affinity for daidzein than for isoliquiritigenin in vitro. However, this does not necessarily have direct implications for functions of UGT71T5 in vivo. For example, a previous report from *M. truncatula* identified a UGT with in vitro preference for isoflavones (Modolo et al., 2007), but later in vivo analysis revealed the UGT's role in anthocyanin glycosylation (Peel et al., 2009).

Previous studies showed that crude protein extracts from kudzu could convert 2,7,4'-trihydroxyisoflavanone to an intermediate that was converted to puerarin by acid treatment, although this compound was not identified (He et al., 2011). Unfortunately, 2,7,4'-trihydroxyisoflavanone is unstable and spontaneously dehydrated to daidzein in the present studies, making it not possible to determine whether the puerarin generated in vitro on incubation of 2,7,4'-trihydroxyisoflavanone with UDP-Glc and UGT71T5 is a result of glycosylation of the trihydroxyisoflavanone or daidzein. When using CHS and UGT71T5 together a cluster of products appeared that was not present in the CHS assay alone. These products are likely from the glycosylation of intermediates in the formation of naringenin chalcone. It is not clear whether these might be true metabolic intermediates, or artifacts of the in vitro assay. When UGT71T5 was added following the CHS reaction, the small peak produced was most likely glycosylated naringenin.

3.4 | Parallel pathways for the biosynthesis of Puerarin

If the C-glycosylation reaction in puerarin biosynthesis occurs at the chalcone level, as suggested by the early labeling studies and the results discussed above, then three reactions, the chalcone to flavanone isomerization catalyzed by CHI, the formation of 2-hydroxyisoflavanone catalyzed by 2-HIS, and the dehydration of

2-hydroxyflavanone catalyzed by 2-HID, should all be able to utilize a C-glycosylated substrate. However, both the CHI and 2-HID reactions can occur spontaneously (Kochs & Grisebach, 1986; Liu et al., 2002), so it is possible that the only enzyme-catalyzed reaction after CHS required in a pathway for puerarin biosynthesis via glycosylation pre-daidzein is that catalyzed by 2-HIS. 2-HIS is encoded by a single gene in both *P. phaseoloides* and *P. m. lobata*, so a specific form of the enzyme with altered specificity for the glycoside cannot exist unless it is encoded by a phylogenetically distinct P450.

Although further biochemical studies are required to further address 2-HIS substrate specificity and kinetics, the accumulation of chalcone and flavanone C-glycosides in 2-HIS RNAi lines, and the demonstration of activity, albeit weak, for formation of 2-hydroxyisoflavanone glucoside on co-incubation of 2-HIS with a mixture of chalcone/flavanone C-glycosides, suggests that flavanone C-glycosides may be natural substrates for 2-HIS. Our model for the biosynthesis of puerarin (Figure 11) proposes parallel routes in which UGT71T5 is present in either free or channel-associated forms. We propose that free UGT71T5 acts on daidzein that is produced via a metabolon that includes all the enzymes of the pathway from L-phenylalanine (Figure 11b), consistent with the very high % labeling of puerarin from this initial precursor. Additionally, channel associated UGT71T5 situates itself in the metabolon at the point of chalcone formation resulting in C-glycosylation of isoliquiritigenin, which is then

converted to puerarin via CHI, 2-HIS and 2-HID, the first and last reactions also being able to proceed non-enzymatically (Figure 11a). Our results did not directly support the model in which the 2-hydroxyisoflavanone product of 2-HIS is the precursor for C-glycosylation (Figure 1, route 2). This situation reflects the recent report that direct 4'-O-methylation of daidzein is responsible for synthesis of formononetin in *Pueraria* (Li et al., 2016), in contrast to a previous report demonstrating that the 4-O-methyl group of formononetin was introduced at the level of 2-hydroxyisoflavanone in *Glycyrrhiza echinata* (Akashi et al., 2000). Clearly, the (iso)flavonoid metabolon requires further study to understand how pools of intermediates may be sequestered and directed into specific downstream intermediates in different species or under different conditions.

4 | MATERIALS AND METHODS

4.1 | Plant materials and growth conditions

Seeds of wild kudzu plants were collected approximately 3 miles southwest of the Samuel Roberts Noble Foundation, Ardmore, Oklahoma, GPS co-ordinates 34.159020, -97.108043 as described (He et al., 2011). Wild kudzu seeds were also collected off Copeland Road under Batman the Ride at Six Flags Over Texas in Arlington, GPS

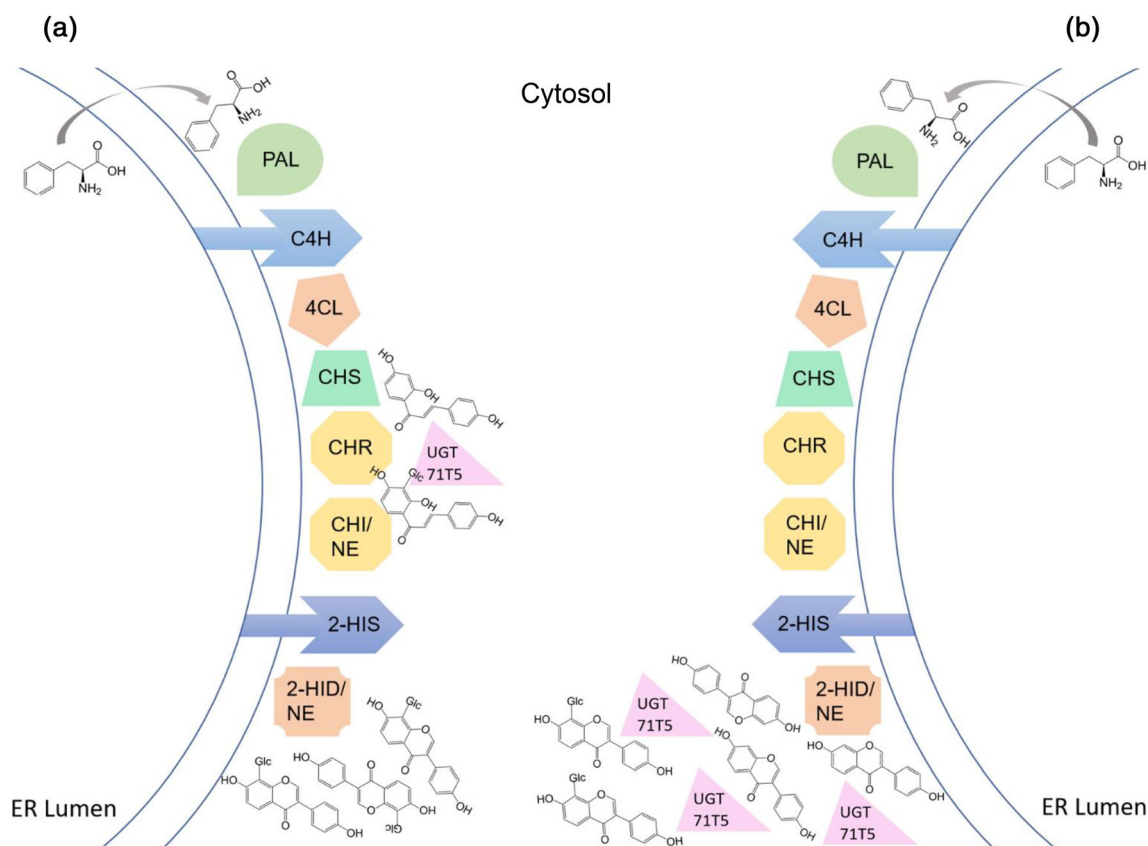


FIGURE 11 A model for the biosynthesis of puerarin involving UGT71T5 in both free and channel-associated forms. (a), The enzymatic synthesis of puerarin through channel-associated UGT71T5 proceeding via C-glycosylation of isoliquiritigenin. (b), The enzymatic synthesis of puerarin following the formation of daidzein with free UGT71T5. Enzyme names as in legend to Figure 1. NE, non-enzymatic.

co-ordinates 32.758692.-97.067350. Seeds of commercial “kudzu” plants were purchased from Kudzu Kingdom Division of Suntop Inc., PO Box 98, Kodak, TN 37764, USA. Seeds were scarified in sulfuric acid for 8 min (*M. truncatula*), 20 min (*P. phaseoloides*), or 45 min (*P. m. lobata*). They were then rinsed with copious amounts of water three times, dried and sterilized in 20% bleach for 5 min. The seeds were allowed to dry before being plated on water agar. The plates were placed in the dark at 4 °C for 3 (*M. truncatula*) or 5 d (kudzu accessions). The plates were then moved to a 24°C light chamber and monitored for germination. Kudzu accessions were maintained in the greenhouse in soil and propagated by cuttings. Morphological features of the two kudzu lines have been described (Adolfo et al., 2022; He et al., 2011).

4.2 | Chemicals

Media were purchased from Phytotechnology Laboratories (Lenexa, KS). ²H₆-daidzein was purchased from Santa Cruz Biotechnologies (Dallas, TX), ¹³C₉-L-Phe was from Sigma-Aldrich (St. Louis, MO), and ¹³C₆-4-hydroxybenzaldehyde was from Cambridge Isotope Laboratories Inc (Tewksbury, MA). Daidzin, genistein, and genistin were from Cayman Chemical Company (Ann Arbor, MI), and esculin, liquiritin, and polydatin were from MedChemExpress (Monmouth Junction, NJ). All other standards were purchased from Indofine Chemical Company (Hillsborough, NJ). Antibiotics were purchased from GoldBio (St. Louis, MO). HPLC solvents were from FisherSci (Waltham, MA) and LC-MS solvents were from Honeywell (Charlotte, NC). Other chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise indicated.

¹³C-Labeled isoliquiritigenin was synthesized by condensation of ¹³C₆-4-hydroxybenzaldehyde and 2',4'-dihydroxyacetophenone using a published protocol (Geissmann and Clinton, 1947). 2,7,4'-Trihydroxyisoflavanone was synthesized from liquiritigenin using recombinant 2-HIS. The 2-HIS used was modified by removing the membrane spanning domain to make the protein more soluble (Chang et al., 2018). LB broth + 50 mg/l carbenicillin + 1 mM thiamine + .5 mM δ-aminolevulinic acid (TCI America) was inoculated with *E. coli* strain BL21 (DE3) harboring 2-HIS in pCWori+ (Chang et al., 2018) and incubated at 37°C until the OD₆₀₀ reached .5. The culture was moved to 30°C and grown to an OD₆₀₀ = .8–1.0. One mM IPTG was added and the culture incubated in a shaker at 28°C for 48 h. The culture was centrifuged and placed at –80°C overnight.

The cell pellet was thawed on ice and resuspended in lysis buffer (50 mM sodium phosphate pH 7.2, 1 M sodium chloride, 15 mM imidazole, and 5 mM β-mercaptoethanol). The resuspended cells were sonicated for 30 min and pelleted. The supernatant was removed and incubated with Pierce HisPur Ni-NTA resin for 2 h at 4 °C. The beads were washed three times with lysis buffer (minus the β-mercaptoethanol) and eluted three times with 50 mM sodium phosphate pH 7.2, 300 mM sodium chloride, and 250 mM imidazole. The eluates were pooled and concentrated on an Amicon Ultra-4 centrifugal filter. The concentrated eluted protein was aliquoted and stored at –80°C.

Following purification, an excess of 2-HIS protein was incubated in a 200 μl reaction with .4 M sucrose, .5 mM glutathione, 1 mM NADPH, 100 mM potassium-phosphate, 25 μl P450 oxidoreductase Supersomes (Corning) and 80 μM liquiritigenin for 16 h on an end-over-end rotator at 16°C. Two vols of ethyl acetate were added to the reaction, centrifuged for 10 min at 12,000 x g, and the upper layer transferred to a new tube. The ethyl acetate extraction was repeated and the contents of the tube dried under air/nitrogen on an Organomation evaporator and resuspended in methanol to be used for LC-MS analysis and as a substrate.

4.3 | RNA extraction, cDNA library construction and Illumina sequencing

The RNA extraction, cDNA library construction and Illumina sequencing were done as described in Adolfo et al. (2022).

4.4 | Short read De novo assembly and annotation of transcriptomes

The transcriptome assembly and annotations used for analysis of CHS, CHR, CHI, 2-HIS, and 2-HID were done as described in Adolfo et al. (2022).

4.5 | Amino acid sequence alignments

For phylogenetic analysis of CHI and 2-HID, multiple protein sequences were aligned using the MUSCLE algorithm in Geneious Prime 2022.0.1 (Edgar, 2004). Neighbor-joining (NJ) phylogenetic trees were generated by the Geneious Prime 2022.0.1, using 1,000 replicates of bootstrap analysis. The multiple protein alignment for IFS/2-HIS was performed using MUSCLE through the web interface provided by EMBL-EBI (Madeira et al., 2022).

4.6 | De novo transcriptome assembly and analysis for glycosyltransferases

Raw RNA-seq data were analyzed by Fast QC (Andrews, 2010) to ensure the data were good quality. The low-quality sequences were removed using the Fast X toolkit (Hannon, 2010). Adapter sequences were removed using Trimmomatic (Bolger et al., 2014). Fast QC was then used to validate adapter removal and read trimming. A de novo transcriptome was built using Trinity (Grabherr et al., 2011) and kmer lengths of 25 and 31.

For digital expression, raw counts were generated using Salmon (Patro et al., 2017) and imported into DeSeq2 (Love et al., 2014) for normalization and differential expression analysis with a joint transcriptome of the *P. phaseoloides* and *P. m. lobata* data. The individual transcriptomes were then compared to the joint transcriptome to



determine expression profiles. For identification of glycosyltransferases, Transdecoder (Haas & Papanicolaou, 2018) was used to identify possible protein sequences of at least 100 amino acids. The sequences were then analyzed using Interproscan (Jones et al., 2014) to identify characteristics found among GO terms, Pfam, PROSITE, and other categories for motifs. IDs from PROSITE and Pfam were identified as matching glycosyltransferases. The sequences with those IDs were analyzed for digital expression levels between *P.m. lobata* and *P. phaseoloides*.

4.7 | Phylogenetic analysis of glycosyltransferases

Phylogenetic trees were made in Geneious Prime 2022.0.1 from a protein alignment using an ATGC PhymL plugin 2.2.4 and an LG substitution model (Guindon et al., 2010) with 1,000 bootstrap replicates. The protein alignment was done in Geneious Prime 2022.0. 1 with MUSCLE (Edgar, 2004).

4.8 | RNA isolation, DNase treatment and cDNA synthesis

Tissue was collected and placed in 2 ml Eppendorf tubes with a single ball bearing. The tubes were placed in liquid nitrogen and the tissue was ground using a Retsch Mixer Mill 400 at 30 Hz for 15 s. The samples were then checked for degree of grinding and placed in liquid nitrogen. If the tissue was not thoroughly ground, it was run on the Retsch Mill again. RNA was isolated using PureLink™ Plant RNA reagent from Invitrogen™ following the manufacturer's instructions for small scale isolation except for the addition of repeating steps 4–6. The concentration of RNA was estimated using a NanoDrop™ 2000. Following RNA isolation, the samples were treated with TURBO DNA-free™ following the manufacturer's instructions in a 25 µl reaction volume for DNase digestion of contaminating DNA in the RNA samples. cDNA synthesis was performed using the SuperScript™ III First-strand synthesis system following the manufacturer's instructions for oligo (dT)₂₀ primers.

4.9 | qPCR

qPCR was performed on a QuantStudio 6 using PowerUp™ SYBR™ Green Master Mix in a total reaction volume of 10 µl. qPCR primers were used at a final concentration of 500 nM. Primers were designed with the Integrated DNA technologies (IDT) qPCR primer design tool. The PSPG motif was avoided in primer design.

4.10 | PCR

Amplification of target regions was performed following the manufacturer's instructions for NEB's Q5 Hot-start polymerase including

extension time. The annealing temperature was calculated using NEB's Tm calculator. The amplified products were visualized on a 1% agarose gel pre-stained with SYBR Safe DNA gel stain on a transilluminator.

Colony PCR was performed on individual colonies. The colonies were picked and resuspended in 10 µl ddH₂O. Two µL of resuspended colony was used for colony PCR using GoTaq Green Master Mix from Promega. Vector specific primers such as M13 F and M13 R were used if directionality did not need to be confirmed. A vector-specific primer coupled with a gene-specific primer were used to determine directionality. The manufacturer's instructions for the GoTaq Green master mix was used for thermocycler conditions and PCR recipe. Primer melting temperatures were calculated using the T_m for oligo calculator from Promega. The PCR reactions were visualized on a 1% agarose gel pre-stained with SYBR Safe DNA gel stain on a transilluminator.

4.11 | Generation of constructs for genetic transformation

cDNA from kudzu was amplified with primers designed for either full length amplification or amplification of a ~ 300 bp section of the gene for RNAi. For glycosyltransferase candidates, the RNAi section was designed outside the PSPG motif to avoid common sequences among the candidates. The amplification products were cleaned following the DNA clean up protocol from the Promega Wizard SV gel and PCR clean-up kit.

The cleaned amplifications were cloned into either pENTR/D-TOPO or pCR8/GW/TOPO TA entry vectors following the manufacturer's protocols. Cloning into pENTR/D-TOPO required a CACC to be included in the forward primer used for amplification. Cloning into pCR8/GW/TOPO TA required the addition of A's which were added after PCR clean up. The A-tail was added using BioReady rTaq DNA polymerase from Bulldog Bio.

The cloning reactions were transformed into chemically competent OneShot TOP10 cells following the manufacturer's instructions. pENTR/D-TOPO transformations were plated on LB medium (Bertani, 1951) + kanamycin while pCR8/GW/TOPO transformations were plated on LB (Bertani, 1951) + spectinomycin. Colonies were checked for the insert using colony PCR.

Colonies with the appropriate construct were grown overnight in 3 ml of broth following the manufacturer's instructions for plasmid preps using either the Qiagen QIAprep Spin Miniprep kit or the Bio-ner AccuPrep Plasmid Miniprep DNA Extraction kit with suggested improvements (<https://bitesizebio.com/13502/how-to-get-better-plasmid-midprep-yields/>). Purified plasmids were sequenced by Eurofins Genomics or Genscript.

Entry vectors with the verified gene inserted were cloned into destination vectors using LR clonase II. The reaction followed the manufacturer's protocol with the exception of being left overnight at room temperature before transformation with chemically competent OneShot TOP10 cells. Genes that were used for expression of



recombinant proteins were cloned into pDEST15 vectors with a GST-tag. Genes that were expressed heterologously in *M. truncatula* were cloned into the Ghent expression vector, pK7WG2. Genes that were knocked down in kudzu were cloned into the Ghent expression vector pK7GWIWG2(I).

To recombine the pCR8/GW/TOPO plasmid with any of the Ghent expression vectors the pCR8/GW/TOPO plasmid was first cut with *pvuI* from NEB. The restriction digest was carried out following the manufacturer's protocol except that the incubation was for 3 h. The plasmid was then run on a .8% agarose gel pre-stained with SYBR™ Safe DNA gel stain and visualized on a transilluminator. The band on the gel was cut out and cleaned using the Promega wizard SV PCR and gel clean-up kit following the manufacturer's instructions.

The transformations with the cloning reactions and competent cells were done following the manufacturer's protocol. The transformations were plated on either LB medium + spectinomycin for the reactions with the Ghent expression vectors or with LB + carbenicillin for the reactions with pDEST15. Colonies from pDEST15, pK7GWIWG2(I), or pK7WG2 were confirmed using colony PCR and positive colonies were grown for plasmid preparation as above.

Ghent expression plasmids were electroporated into *A. rhizogenes* strain A4TC24 following the protocol below. pDEST15 plasmids were transformed into rosetta2 pLysS following the manufacturer's instructions.

4.12 | Transformation of Agrobacterium

Electrocompetent Agrobacterium cells were made using *A. rhizogenes* A4TC24 using the following procedure. Glycerol stocks of Agrobacterium were streaked on LB medium + 25 mg/l rifampicin and 100 mg/l streptomycin. After 2 d of growth at 28°C several colonies were used to inoculate 50 µl of LB broth + 25 mg/l rifampicin and 100 mg/l streptomycin overnight. Two hundred mL of fresh LB broth + 25 mg/l rifampicin and 100 mg/l streptomycin was inoculated with 25 ml of the overnight Agrobacterium culture. The 225 ml culture was grown to an OD₆₀₀ of 1.0. The culture was poured into a cold centrifuge bottle and chilled on ice for 20 min. The cells were collected by centrifugation at 4,000 x g for 15 min at 4 °C, the medium removed, and the cells washed by resuspending them in 250 ml of cold sterile ddH₂O and centrifuging at 4,000 x g for 15 min at 4 °C. The water was removed and an additional 250 ml of sterile ddH₂O used to resuspend the cells. The cell washing was repeated for a total of five washes. After the last wash, the cells were collected by centrifugation at 4,000 x g for 15 min at 4 °C and the liquid removed. The cells were resuspended in 25 ml of cold 10% (v/v) glycerol, transferred to a 50 ml conical tube, and collected by centrifugation at 3,000 x g for 10 min at 4 °C. The liquid was again removed, and the cells resuspended in 1 ml of cold 10% glycerol. Fifty µL of the resuspended cells were aliquoted into tubes, which were placed in liquid nitrogen and then stored at -80°C.

To transform the electrocompetent cells, the cells were completely thawed on ice. Two µL of purified plasmid was pipetted

into the competent cells without mixing. The tubes were gently flicked to mix. The tubes were placed back on ice for 10 min. The entire contents of the tubes were transferred to an ice-cold 2 mm electroporation cuvette. The cuvette was electroporated on an electroporator at 2400 V. Immediately following electroporation, 250 µl of super optimal broth (SOC) was added to the cuvette and the cuvette placed on ice. The contents of the cuvette were gently poured into a microfuge tube and placed in a 28°C incubator to shake gently for 2–4 h to recover. Fifty µL of recovered electroporated Agrobacterium cells were plated on LB medium + 25 mg/l rifampicin and 100 mg/l streptomycin. After 2 d, colony PCR was performed to check for the presence of the vector.

4.13 | Recombinant protein expression and purification

E. coli cells harboring recombinant UGTs were grown at 37°C in LB medium + 50 mg/l carbenicillin to OD = .8 and induced with 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside). The cultures were placed at 18–20°C for 20 h, then centrifuged and the pellets placed at -20°C overnight. The bacterial pellets were thawed on ice and resuspended in PBS pH 7.4 and freeze/thawed three times. The cells were sonicated for 15 min on ice and centrifuged. The soluble protein sonicate was transferred to GE Sepharose 4B glutathione agarose beads and incubated on an end-over-end rotator for 2 h at 4 °C. Following incubation the beads were washed three times with PBS. The protein was eluted three times with 50 mM glutathione and 50 mM Tris pH 8. Eluted proteins were pooled and concentrated on an Amicon Ultra-4 centrifugal filter. The concentrated eluted protein was aliquoted and stored at -80°C.

4.14 | Protein quantification

Five µL of concentrated eluted protein was run in a 20 µl reaction volume on an SDS-PAGE gel. Additionally, BSA was loaded on the gel at final concentrations of 5, 2.5, 1.25, .625, .3125, .15265 and .0625 µg protein. The gel was stained with GelCode Blue Safe Protein Stain for 1 h and destained for at least 1 h. The gel image was analyzed on imageJ using the known BSA concentrations to determine the recombinant protein concentration following the procedure on openwetware (https://openwetware.org/wiki/Protein_Quantification_Using_ImageJ).

4.15 | Glycosyltransferase enzyme assay

Ten µg of recombinant UGT was incubated with 200 µM substrate, 2.5 mM UDP-glucose, 50 mM Tris pH 7.5 for 3 h at 30°C. Two vols ethyl acetate were added to the reaction, centrifuged for 10 min at 12,000 xg, and the top layer transferred to a new tube. The ethyl acetate extraction was performed once more. The supernatants were



pooled and dried under air/nitrogen on an Organomation evaporator and resuspended in methanol.

For β -glucosidase hydrolysis, rather than resuspend in methanol, the samples were resuspended in 400 μ l sodium phosphate buffer (pH 6). The supernatant was divided into two tubes, one as a control without glucosidase digestion, the other for treatment with 6 mg β -glucosidase for 12 h at 37°C. Following treatment the samples were extracted twice with ethyl acetate, resuspended in 25 μ l methanol, and run on the LC-MS as described below.

4.16 | Generation of *M. truncatula* hairy roots

M. truncatula A17 seeds were scarified and sterilized as above. One-2 d after transfer of plates to the 24°C long day chamber, a lawn of *Agrobacterium* was plated on LB with appropriate antibiotics. After approximately 3–4 d in the light chamber, the root tips were cut off the seedlings at the hypocotyl and root junction and the freshly cut hypocotyl end was wiped across a lawn of *A. rhizogenes* strain A4TC24 with appropriate construct plated on LB medium + antibiotic(s) appropriate for the bacteria and plasmid. The seedlings were then plated on modified Fahreus medium (Fahreus, 1957), and the plates sealed with parafilm (slit to allow for gas exchange) and placed in the dark at 24°C at a raised angle to promote root growth for 3 d. The plates were then moved to a 24°C long day chamber. Approximately 2 weeks later, roots were cut and plated on B5 medium (Gamborg et al., 1968) plus 200 μ g/l Timentin™ to cure *Agrobacterium* growth and further antibiotics specific for plasmid selection.

4.17 | Generation of kudzu hairy roots

A. rhizogenes was grown to OD₆₀₀ = 1.2 and centrifuged before resuspending at room temperature in double the volume of MS medium (Murashige & Skoog, 1962) supplemented with 30% sucrose, 5 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D), 1 mg/l 6-benzylaminopurine (BAP), and 100 μ M acetosyringone. Leaves were sterilized with .8% (v/v) bleach for 10 min and then soaked in 2% PPM (Plant Preservative Mixture) for no more than 4 h. The sterilized leaf disks were added to the resuspended *Agrobacterium* and a vacuum was pulled for 30 min to facilitate *Agrobacterium* infiltration. The cultures were then placed on a gentle shaker at room temperature for 28°C.

The leaf disks were plated on MS medium supplemented with 30% sucrose, 5 mg/l 2,4-D, 1 mg/l BAP, and 100 μ M acetosyringone solidified with agar before being placed in a 24°C dark chamber. After 3 d the leaf disks were transferred to MS medium supplemented with 30% sucrose, 5 mg/l 2,4-D, 1 mg/l BAP, and 200 μ M Timentin™ solidified with Phytigel™. The plates were placed in a 24°C dark chamber. After several weeks calli began to form, and then roots. The transformed roots were moved to MS medium supplemented with 30% sucrose, 25–50 mg/l kanamycin, 200 μ M Timentin™, and IBA solidified with Phytigel™.

4.18 | Labeling experiments

One week old seedlings were placed on a float disk allowing the roots to be submerged in liquid ½ B5 medium. The plants were placed in 40 ml medium that contained either no compound, 100 μ M L-phenylalanine (labeled or unlabeled), 100 μ M isoliquiritigenin (labeled or unlabeled), or 45 μ M daidzein (labeled or unlabeled). The plants were placed in a 24°C long day chamber for 3 weeks. Following incubation, roots were patted dry using filter paper, cut off from the rest of the plant, and placed in 2 ml Eppendorf tubes with a single ball bearing. Flavonoid extraction from the roots was as described below.

4.19 | Flavonoid extraction

Kudzu roots were dried on a lyophilizer and ground to a powder in liquid nitrogen. About 10–20 mg of powder was weighed in Eppendorf tubes and extracted with 80% methanol. The extracts were centrifuged at 12,000 x g for 20 min, dried under a stream of air and resuspended in water. Ethyl acetate extraction was performed twice before resuspension in 150 μ l methanol.

4.20 | LC-MS/MS analysis of labeled phenolic compounds

Detection and quantification of ¹²C- and ¹³C-isoflavones and ²H- and ¹H- isoflavones was conducted as previously described (Cocuron et al., 2019). Isoflavones were detected and quantified using an Agilent 1,290 Infinity II liquid chromatography system coupled to a hybrid Triple Quadrupole 6,500 + triple quadrupole from ABSciex. The extracts were kept at 10 °C in an auto-sampler. The metabolites were resolved at 30°C using a reverse phase C18 Symmetry column (4.6 x 75 mm; 3.5 μ m) associated with a Symmetry C18 pre-column (3.9 x 20 mm; 5 μ m) from Waters. The liquid chromatography gradient was made of .1% (v/v) acetic acid in acetonitrile (A) and .1% (v/v) acetic acid in water (B). The total LC-MS/MS run was 15 min with a flow rate of 800 μ l/min. The following gradient was applied to separate the phenolic compounds: B = 0–1 minute 85%, 1–7 minutes 42%, 7–9 minutes 20%, 9–9.1 minutes 15%, 9.1–12 min 15%, 12–12.1 minutes 85%, 12.1–15 minutes 85%. The injection needle was rinsed with 50% aqueous methanol. Five μ l of extracts or external standard mixture was injected onto the column.

The detection of isoflavones was accomplished using an AB Sciex hybrid Triple Quadrupole/Ion trap mass spectrometer QTRAP 6500+. Electrospray ionization with polarity switch was utilized to acquire mass spectra of the metabolites. In this study, negative ionization was applied to formononetin, genistein, genistin, isoliquiritigenin, and liquiritigenin, whereas positive ionization was used for daidzein, daidzin, ononin, and puerarin. The settling time between each polarity was 15 msec. ¹²C- and ¹³C- metabolites and ²H- and ¹H- metabolites were simultaneously detected as precursor ion/product ion pair using

multiple reaction monitoring (MRM) (see Supplemental Table S4 for ^{12}C - and ^{13}C - metabolite parameters). The source parameters such as curtain gas, temperature, nebulizer gas (GS1), heating gas (GS2), and collision activated dissociation (CAD) were kept constant during MRM survey scan (Supplemental Table S5). The dwell time in the mass spectrometer was set to 10 msec. Analyst 1.7 software from AB Sciex was used to acquire and process the LC-MS/MS data. Isoflavones were identified and quantified using a mixture of known external standards run at the same time as the biological extracts.

4.21 | Determination of label incorporation

The percent of labeled compound incorporated into flavonoids was calculated by adding the peak areas for the labeled and unlabeled compounds. The peak area of the labeled compound was then divided by total amount of the compound (labeled and unlabeled). Finally, the number was multiplied by 100.

4.22 | HPLC for analysis of Isoflavones from roots of mature kudzu

Twenty μL samples were injected on an Agilent 1,220 Infinity II with a C18 reverse phase column. The 50 min run used the solvents .1% (v/v) formic acid (A) and acetonitrile (B) with a gradient of: 0–5 min, 95% A; 5–10 min, 85% A; 10–25 min, 77% A; 25–30 min, 67% A; 30–35 min, 60% A; 35–40 min, 0% A; 40–45 min, 0% A; 45–50 min, 95% A with a flow rate of 1 ml/min. The absorption was measured 254 nm.

4.23 | LC-MS analysis for products of enzyme reactions

Products from enzyme reactions in methanol (10 μL) were injected on an Agilent 1,290 Infinity II with a Xterra C18 reverse phase column (2.1 x 250 mm; 5 μm) set to 45°C. The 35 min run used the solvents .1% (v/v) formic acid in water (A) and .1% (v/v) formic acid in acetonitrile (B) with the following gradient: 1 min, 95% A; 2 min, 87% A; 8 min, 87% A; 14 min, 70% A; 19 min, 60% A; 24 min, 50% A; 30 min, 30% A; 31 min, 5% A; 32 min, 5% A; 33 min, 95% A; 35 min, 95% A with a flow rate of .45 ml/min. The MS was acquired in negative mode from ESI + Agilent Jet Stream in MS2 scan.

To separate kaempferol glucosides, 10 μL samples were injected on an Agilent 1,290 Infinity II with a reverse phase C18 Symmetry column (4.6 x 75 mm; 3.5 μm) set to 30°C. The 15 min run used the solvents .1% (v/v) acetic acid in acetonitrile (A) and .1% (v/v) acetic acid in water (B) with the following gradient: 1 min, 98% B; 7 min, 42% B; 9 min, 20% B; 11 min, 10% B; 13 min, 10% B; 13.1 min, 98% B; 15 min 98% B with a flow rate of .8 ml/min. The MS was acquired in negative mode from ESI + Agilent Jet Stream in MS2 scan.

4.24 | High resolution MS analysis

Further analysis of metabolites was carried out using an Exion ultra high performance liquid chromatography system coupled with a high resolution mass spectrometer TripleTOF6600 + from AB Sciex. The compounds were separated using a C18 Xterra MS (250 x 2.1 mm, 5 μm) column from Waters. The temperatures of the column compartment and the autosampler were kept at 45°C and 15°C, respectively. The analytes were eluted using a gradient of .1% formic acid in acetonitrile (Solvent A) and .1% formic acid in water (Solvent B) under a flow rate of .45 ml/min. The following gradient was applied: 0–1.0 min, 5.0% A; 1.0–2.0 min, 5.0–13.0% A; 2.0–8.0 min, 13.0% A; 8.0–14.0 min, 13.0–30.0% A; 14.0–19.0 min, 30.0–40.0% A; 19.0–24.0, 40.0–50.0% A; 24.0–24.1 min, 50.0–95.0% A; 24.1–27.0 min, 95.0% A; 27.0–27.1 min, 95.0–5.0% A; 27.1–32.0 min; 5% A.

The mass spectrometer was set to scan metabolites from m/z 200–700 amu in negative mode with an ion spray voltage of 4,500 V. The accumulation time was 100 msec. Parameters such as declustering potential and collision energy were 50 V and 10 V, respectively. MS/MS spectra were acquired over m/z 30–700 amu with an accumulation time of 25 msec. Parameters such as declustering potential, collision energy, collision energy spread were set to 50 V, 35 V and 15 V, respectively.

The curtain gas (nitrogen), nebulizing and heating gas were fixed at 30 psi, 60 psi and 60 psi, respectively. The temperature of the source was 500°C. An APCI negative calibration solution was delivered by a calibrant delivery system every 5 samples in order to correct for any mass drift that may occur during the run. MS spectra were acquired using Analyst TF 1.8.1 software. The analysis of the data was performed using Sciex OS software.

4.25 | Compound quantification

Compound amounts were calculated by first taking the peak area of the compound's standard and dividing it by the amount of standard run on the machine to give the response factor. To calculate unknown amounts of compounds from samples, the peak area of the compound from the sample was divided by the response factor calculated from the standard.

4.26 | Statistical analysis and significance tests

To determine if there was a significant difference between levels of compounds in different plant materials, an unpaired two-tailed *t* test was performed in GraphPad Prism 8.

4.27 | Accession numbers

Sequence data from this article can be found in the NCBI Sequence Read Archive (SRA) repository, NCBI SRA accession No. SRX768865.

UGT and isoflavone pathway sequences can be found in GenBank (OL860977- OL860984 and ON986485-ON986500 respectively).

ACKNOWLEDGMENTS

This work was supported by the University of North Texas. We thank Christophe Cocuron and the BioAnalytical Facility at the University of North Texas for assistance with mass spectrometry, Dr Xiaoqiang Wang for providing recombinant 2-HIS, Dr Ji Hyung Jun for providing images of GUS control *M. truncatula* hairy roots, Dr Michael Court and the UGT nomenclature committee at Washington State University for providing official UGT nomenclatures for the kudzu UGTs, Dr Awinash Bhatkar of the Texas Department of Agriculture for providing kudzu transportation permits, David McSweeney and Yuhong Tang of the Noble Research Institute for providing *M. truncatula* A17 seeds and for RNA sequencing preparation, respectively, Sue Miller of the USDA for providing *A. rhizogenes* strain A4TC24, and Enrique Carranco for assistance with programming.

CONFLICT OF INTEREST

The Authors did not report any conflict of interest.

ORCID

Laci M. Adolfo  <https://orcid.org/0000-0001-9199-3818>

David Burks  <https://orcid.org/0000-0003-0098-265X>

Xiaolan Rao  <https://orcid.org/0000-0001-7939-8362>

Anislay Alvarez-Hernandez  <https://orcid.org/0000-0002-9962-3456>

Richard A. Dixon  <https://orcid.org/0000-0001-8393-9408>

REFERENCES

- Adiji, O. A., Docampo-Palacios, M. L., Alvarez-Hernandez, A., Pasinetti, G. M., Wang, X., & Dixon, R. A. (2021). UGT84F9 is the major flavonoid UDP-glucuronosyltransferase in *Medicago truncatula*. *Plant Physiology*, *185*, 1617–1637. <https://doi.org/10.1093/plphys/kiab016>
- Adolfo, L. A., Rao, X., & Dixon, R. A. (2022). Identification of *Pueraria* spp. through DNA barcoding and comparative transcriptomics. *BMC Plant Biology*, *22*, 10. <https://doi.org/10.1186/s12870-021-03383-x>
- Akashi, T., Aoki, T., & Ayabe, S.-I. (2005). Molecular and biochemical characterization of 2-hydroxyisoflavanone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis. *Plant Physiology*, *137*, 882–891. <https://doi.org/10.1104/pp.104.056747>
- Akashi, T., Sawada, Y., Aoki, T., & Ayabe, S. (2000). New scheme of the biosynthesis of formononetin involving 2,7,4'-trihydroxyisoflavanone but not daidzein as the methyl acceptor. *Bioscience, Biotechnology, and Biochemistry*, *64*(10), 2276–2279. <https://doi.org/10.1271/bbb.64.2276>
- Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Boisson-Dernier, A., Chabaud, M., Garcia, F., Bécard, G., Rosenberg, C., & Barker, D. G. (2001). *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Molecular Plant-Microbe Interactions*, *14*, 695–700. <https://doi.org/10.1094/MPMI.2001.14.6.695>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Boonsongcheep, P., Korsangruang, S., Soonthornchareonnon, N., Chintapakorn, Y., Saralamp, P., & Prathanturug, S. (2010). Growth and isoflavonoid accumulation of *Pueraria candollei* var. *candollei* and *P. candollei* var. *mirifica* cell suspension cultures. *Plant Cell, Tissue and Organ Culture*, *101*, 119–126. <https://doi.org/10.1007/s11240-010-9668-x>
- Brazier-Hicks, M., Evans, K. M., Gershater, M. C., Puschmann, H., Steel, P. G., & Edwards, R. (2009). The C-glycosylation of flavonoids in cereals. *The Journal of Biological Chemistry*, *284*, 17926–17934. <https://doi.org/10.1074/jbc.M109.009258>
- Britton, K. O., Orr, D., & Sun, J. H. (2002). Biological control of kudzu, *Pueraria montana* var. *lobata*. Biological control of invasive plants in the eastern United States. USDA Forest Service Publication FHTET-2002-04: 3
- Burbulis, I. E., & Winkel-Shirley, B. (1999). Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 12929–12934. <https://doi.org/10.1073/pnas.96.22.12929>
- Chang, Z., Wang, X., Wei, R., Liu, Z., Shan, H., Fan, G., & Hu, H. (2018). Functional expression and purification of CYP93C20 a plant membrane-associated cytochrome P450 from *Medicago truncatula*. *Protein Expression and Purification*, *150*, 44–52. <https://doi.org/10.1016/j.pep.2018.04.017>
- Cocuron, J. C., Casas, M. I., Yang, F., Grotewold, E., & Alonso, A. P. (2019). Beyond the wall: High-throughput quantification of plant soluble and cell-wall bound phenolics by liquid chromatography tandem mass spectrometry. *Journal of Chromatography. A*, *1589*, 93–104. <https://doi.org/10.1016/j.chroma.2018.12.059>
- Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 113. <https://doi.org/10.1186/1471-2105-5-113>
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L. J., Salazar, G. A., Smart, A., & Sonnhammer, E. L. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research*, *47*, D427–D432. <https://doi.org/10.1093/nar/gky995>
- EPPO. (2007). Data sheets on quarantine pests: *Pueraria lobata*. *European and Mediterranean Plant Protection Organization Bulletin*, *37*, 230–235.
- Farag, M. A., Huhman, D. V., Lei, Z., & Sumner, L. W. (2007). Metabolic profiling and systematic identification of flavonoids and isoflavonoids in roots and cell suspension cultures of *Medicago truncatula* using HPLC-UV-ESI-MS and GC-MS. *Phytochemistry*, *68*, 342–354. <https://doi.org/10.1016/j.phytochem.2006.10.023>
- 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 Physiology*, *56*, 1512–1520. <https://doi.org/10.1093/pcp/pcv072>
- Gamborg, O. L., Miller, R. A., & Ojima, K. (1968). Nutrient requirements of suspension cultures of soybean root cells. *Experimental Cell Research*, *50*, 151–158. [https://doi.org/10.1016/0014-4827\(68\)90403-5](https://doi.org/10.1016/0014-4827(68)90403-5)
- Gandia-Herrero, F., Lorenz, A., Larson, T., Graham, I. A., Bowles, D., Rylott, E. L., & Bruce, N. (2008). Detoxification of the explosive 2,4,6-trinitrotoluene in *Arabidopsis*: Discovery of bifunctional O- and C-glucosyltransferases. *The Plant Journal*, *56*, 936–974. <https://doi.org/10.1111/j.1365-3113.2008.03653.x>
- Geissmann, T. A., & Clinton, R. O. (1946). Flavanones and related compounds. I. the preparation of polyhydroxy-chalcones and flavanones.

- Journal of the American Chemical Society*, 68, 697–700. <https://doi.org/10.1021/ja01208a051>
- Geng, P., Sun, J., Zhang, M., Li, X., Harnly, J. M., & Chen, P. (2016). Comprehensive characterization of C-glycosyl flavones in wheat (*Triticum aestivum* L.) germ using UPLC-PDA-ESI/HRMSⁿ and mass defect filtering. *Journal of Mass Spectrometry*, 51, 914–930. <https://doi.org/10.1002/jms.3803>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., ... Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29, 644–652. <https://doi.org/10.1038/nbt.1883>
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59, 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Haas, B., & Papanicolaou, A. (2018). [TransDecoder.GitHub.io](https://github.com/TransDecoder)
- Hakamatsuka, T., Ebizuka, Y., & Sankawa, U. (1994). *Pueraria lobata* (kudzu vine): In vitro culture and the production of isoflavonoids. *Medicinal and Aromatic Plants*, VII, 386–400. https://doi.org/10.1007/978-3-662-30369-6_23
- Hannon, G. J. (2010). FASTX-Toolkit. http://hannonlab.cshl.edu/fastx_toolkit
- He, J. B., Zhao, P., Hu, Z. M., Liu, S., Kuang, Y., Zhang, M., Li, B., Yun, C. H., Qiao, X., & Ye, M. (2019). Molecular and structural characterization of a promiscuous C-glycosyltransferase from *Trollius chinensis*. *Angewandte Chemie (International Ed. in English)*, 58, 11513–11520. <https://doi.org/10.1002/anie.201905505>
- He, X., Blount, J. W., Ge, S., Tang, Y., & Dixon, R. A. (2011). A genomic approach to isoflavone biosynthesis in kudzu (*Pueraria lobata*). *Planta*, 233, 843–885. <https://doi.org/10.1007/s00425-010-1344-1>
- Hirade, Y., Kotoku, N., Terasaka, K., Saijo-Hamano, Y., Fukumoto, A., & Mizukami, H. (2015). Identification and functional analysis of 2-hydroxyflavanone C-glycosyltransferase in soybean (*Glycine max*). *FEBS Letters*, 589, 1778–1786. <https://doi.org/10.1016/j.febslet.2015.05.010>
- Hughes, J., & Hughes, M. A. (1994). Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA Sequence*, 5, 41–49. <https://doi.org/10.3109/10425179409039703>
- Inoue, T., & Fujita, M. (1977). Biosynthesis of puerarin in *Pueraria* root. *Chemical & Pharmaceutical Bulletin*, 25, 3226–3231. <https://doi.org/10.1248/cpb.25.3226>
- Jewett, D. K., Jiang, C. J., Britton, K. O., Sun, J. H., & Tang, J. (2003). Characterizing specimens of kudzu and related taxa with RAPD's. *Castanea*, 68, 254–260.
- Jez, J., Bowman, M., Dixon, R. A., & Noel, J. P. (2000). Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nature Structural & Molecular Biology*, 7, 786–791. <https://doi.org/10.1038/79025>
- Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, 30, 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>
- Jung, W., Yu, O., Lau, S.-M. C., O'Keefe, D. P., Odell, J., Fader, G., & McGonigle, B. (2000). Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nature Biotechnology*, 18, 208–212. <https://doi.org/10.1038/72671>
- Kaufman, P. B., Duke, J. A., Briemann, H., Boik, J., & Hoyt, J. E. (1997). A comparative survey of leguminous plants as sources of the isoflavones, genistein and daidzein: Implications for human nutrition and health. *The Journal of Alternative and Complementary Medicine*, 3, 7–12. <https://doi.org/10.1089/acm.1997.3.7>
- Kochs, G., & Grisebach, H. (1986). Enzymic synthesis of isoflavones. *European Journal of Biochemistry*, 155, 311–318. <https://doi.org/10.1111/j.1432-1033.1986.tb09492.x>
- Li, J., Li, C., Gou, J., Wang, X., Fan, R., & Zhang, Y. (2016). An alternative pathway for formononetin biosynthesis in *Pueraria lobata*. *Frontiers in Plant Science*, 7, 861. <https://doi.org/10.3389/fpls.2016.00861>
- Li, J., Li, Z., Li, C., Gou, J., & Zhang, Y. (2014). Molecular cloning and characterization of an isoflavone 7-O-glucosyltransferase from *Pueraria lobata*. *Plant Cell Reports*, 33, 1173–1185. <https://doi.org/10.1007/s00299-014-1606-7>
- Li, J. M., Wang, G., Liu, J. C., Zhou, L., Dong, M. X., Wang, R., Li, X. Y., Li, X. M., Lin, C. R., & Niu, Y. C. (2010). Puerarin attenuates amyloid-beta-induced cognitive impairment through suppression of apoptosis in rat hippocampus in vivo. *European Journal of Pharmacology*, 649, 195–201. <https://doi.org/10.1016/j.ejphar.2010.09.045>
- Li, L., Xue, Z., Chen, L., Che, X., Wang, H., & Wang, X. (2017). Puerarin suppression of A β ₁₋₄₂-induced primary cortical neuron death is largely dependent on ER β . *Brain Research*, 1657, 87–94. <https://doi.org/10.1016/j.brainres.2016.11.023>
- Li, Y., Li, X.-L., Lai, C.-J.-S., Wang, R.-S., Kang, L.-P., Ma, T., Zhao, Z.-H., Wei, G., & Huang, L.-Q. (2019). Functional characterization of three flavonoid glycosyltransferases from *Andrographis paniculata*. *Royal Society Open Science*, 6, 190150. <https://doi.org/10.1098/rsos.190150>
- Liang, J., & Olsen, R. W. (2014). Alcohol use disorders and current pharmacological therapies: The role of GABA(a) receptors. *Acta Pharmacologica Sinica*, 35, 981–993. <https://doi.org/10.1038/aps.2014.50>
- Liu, C. J., Blount, J. W., Steele, C. L., & Dixon, R. A. (2002). Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 14578–14583. <https://doi.org/10.1073/pnas.212522099>
- Liu, H., Zhang, X., Zhong, X., Li, Z., Cai, S., Yang, P., Ou, C., & Chen, M. (2019). Puerarin inhibits vascular calcification of uremic rats. *European Journal of Pharmacology*, 855, 235–243. <https://doi.org/10.1016/j.ejphar.2019.05.023>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lozovaya, V. V., Lygin, A. V., Zernova, O. V., Ulanov, A. V., Li, S., Hartman, G. L., & Widholm, J. M. (2007). Modification of phenolic metabolism in soybean hairy roots through down regulation of chalcone synthase or isoflavone synthase. *Planta*, 225, 665–679. <https://doi.org/10.1007/s00425-006-0368-z>
- Lukas, S. E. (2002). Human studies of kudzu as a treatment for alcohol abuse. In W. M. Keung (Ed.), *Pueraria: The genus Pueraria* (pp. 159–179). CRC Press.
- Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, 50(W1), W276–W279. <https://doi.org/10.1093/nar/gkac240>
- Mameda, R., Waki, T., Kawai, Y., Takahashi, S., & Nakayama, T. (2018). Involvement of chalcone reductase in the soybean isoflavone metabolism: Identification of GmCHR5, which interacts with 2-hydroxyisoflavanone synthase. *The Plant Journal*, 96, 56–74. <https://doi.org/10.1111/tpj.14014>
- Meezan, E., Meezan, E. M., Jones, K., Moore, R., Barnes, S., & Prasain, J. K. (2005). Contrasting effects of puerarin and daidzin on glucose homeostasis in mice. *Journal of Agricultural and Food Chemistry*, 53, 8760–8767. <https://doi.org/10.1021/jf058105e>
- Modolo, L., Blount, J. W., Achnine, L., Naoumkina, M., Wang, X., & Dixon, R. A. (2007). A functional genomics approach to (iso)flavonoid



- glycosylation in the model legume *Medicago truncatula*. *Plant Molecular Biology*, 64, 499–518. <https://doi.org/10.1007/s11103-007-9167-6>
- Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schaffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics*, 24, 1757–1764. <https://doi.org/10.1093/bioinformatics/btn322>
- Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum*, 15, 473–497. <https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>
- Nagatomo, Y., Usui, S., Ito, T., Kato, A., Shimosaka, M., & Taguchi, G. (2014). Purification, molecular cloning and functional characterization of flavonoid C-glycosyltransferases from *Fagopyrum esculentum* M. (buckwheat) cotyledon. *The Plant Journal*, 80, 437–448. <https://doi.org/10.1111/tpj.12645>
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14, 417–419. <https://doi.org/10.1038/nmeth.4197>
- Peel, G. J., Pang, Y., Modolo, L. V., & Dixon, R. A. (2009). The LAP1 MYB transcription factor orchestrates anthocyanidin biosynthesis and glycosylation in *Medicago*. *The Plant Journal*, 59, 136–149. <https://doi.org/10.1111/j.1365-313X.2009.03885.x>
- Prasain, J. K., Jones, K., Brissie, N., Moore, R., Wyss, J. M., & Barnes, S. (2004). Identification of puerarin and its metabolites in rats by liquid chromatography-tandem mass spectrometry. *Journal of Agricultural and Food Chemistry*, 52, 3708–3712. <https://doi.org/10.1021/jf040037t>
- Prasain, J. K., Jones, K., Kirk, M., Wilson, L., Smith-Johnson, M., Weaver, C., & Barnes, S. (2003). Profiling and quantification of isoflavonoids in kudzu dietary supplements by high-performance liquid chromatography and electrospray ionization tandem mass spectrometry. *Journal of Agricultural and Food Chemistry*, 51, 4213–4218. <https://doi.org/10.1021/jf030174a>
- Rong, H., Keukeleire, D., & Cooman, L. (2002). Chemical constituents of *Pueraria* plants: identification and methods of analysis. In W. M. Keung (Ed.), *Pueraria: The genus Pueraria* (pp. 83–96). Taylor and Francis.
- Sasaki, N., Nishizaki, Y., Yamada, E., Tatsuzawa, F., Nakatsuka, T., Takahashi, H., & Nishihara, M. (2015). Identification of the glucosyltransferase that mediates direct flavone C-glycosylation in *Gentiana triflora*. *FEBS Letters*, 589, 182–187. <https://doi.org/10.1016/j.febslet.2014.11.045>
- Shen, X. L., Witt, M. R., Nielsen, M., & Sterner, O. (1996). Inhibition of H-flunitrazepam binding to rat brain membranes in vitro by puerarin and daidzein. *Acta Pharmaceutica Sinica*, 31, 59–62.
- Shih, C. H., Chu, H., Tang, L. K., Sakamoto, W., Maekawa, M., Chu, I. K., Wang, M., & Lo, C. (2008). Functional characterization of key structural genes in rice flavonoid biosynthesis. *Planta*, 228, 1043–1054. <https://doi.org/10.1007/s00425-008-0806-1>
- Shimamura, M., Akashi, T., Sakurai, N., Suzuki, H., Saito, K., Shibata, K., Ayabe, S., & Aoki, T. (2007). 2-Hydroxyisoflavanone dehydratase is a critical determinant of isoflavone productivity in hairy root cultures of *Lotus japonicus*. *Plant & Cell Physiology*, 48, 1652–1657. <https://doi.org/10.1093/pcp/pcm125>
- Sigrist, C. J., de Castro, E., Cerutti, L., Cuch, B. A., Hulo, N., Bridge, A., Bougueleret, L., & Xenarios, I. (2013). New and continuing developments at PROSITE. *Nucleic Acids Research*, 41, D344–D347. <https://doi.org/10.1093/nar/gks1067>
- Steele, C. L., Gijzen, M., Qutob, D., & Dixon, R. A. (1999). Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. *Archives of Biochemistry and Biophysics*, 367, 147–150. <https://doi.org/10.1006/abbi.1999.1238>
- Subramanian, S., Graham, M. Y., Yu, O., & 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 Physiology*, 137, 1345–1353. <https://doi.org/10.1104/pp.104.057257>
- Sun, J. H., Li, Z. C., Jewett, D. K., Britton, K. O., Ye, W. H., & Ge, X. J. (2005). Genetic diversity of *Pueraria lobata* (kudzu) and closely related taxa as revealed by inter-simple sequence repeat analysis. *Weed Research*, 45, 255–260. <https://doi.org/10.1111/j.1365-3180.2005.00462.x>
- Van der Maesen, L. J. G. (2002). *Pueraria*: botanical characteristics. In W. M. Keung (Ed.), *Pueraria: The genus Pueraria* (pp. 1–28). Taylor and Francis.
- Wagner, G. J., & Hrazdina, G. (1984). Endoplasmic reticulum as a site of phenylpropanoid and flavonoid metabolism in *Hippeastrum*. *Plant Physiology*, 74, 901–906. <https://doi.org/10.1104/pp.74.4.901>
- Wang, X., Fan, R., Li, J., Li, C., & Zhang, Y. (2016). Molecular cloning and functional characterization of a novel (iso)flavone 4',7-O-diglucoside glucosyltransferase from *Pueraria lobata*. *Frontiers in Plant Science*, 7, 387. <https://doi.org/10.3389/fpls.2016.00387>
- Wang, X., Li, C., Zhou, C., Li, J., & Zhang, Y. (2017). Molecular characterization of the C-glycosylation for puerarin biosynthesis in *Pueraria lobata*. *The Plant Journal*, 90, 535–546. <https://doi.org/10.1111/tpj.13510>
- Wong, K. H., Li, G. Q., Li, K. M., Razmovski-Naumovski, V., & Chan, K. (2011). Kudzu root: Traditional uses and potential medicinal benefits in diabetes and cardiovascular diseases. *Journal of Ethnopharmacology*, 134, 584–607. <https://doi.org/10.1016/j.jep.2011.02.001>
- Wu, X.-D., Wang, C., Zhang, Z.-Y., Fu, Y., Liu, F.-Y., & Liu, X.-H. (2014). Puerarin attenuates cerebral damage by improving cerebral microcirculation in spontaneously hypertensive rats. *Evidence-Based Complementary and Alternative Medicine*, 2014, 1–7. <https://doi.org/10.1155/2014/408501>
- Zhang, H. C., Liu, J. M., Lu, H. Y., & Gao, S. L. (2009). Enhanced flavonoid production in hairy root cultures of *Glycyrrhiza uralensis* Fisch by combining the over-expression of chalcone isomerase gene with the elicitation treatment. *Plant Cell Reports*, 28, 1205–1213. <https://doi.org/10.1007/s00299-009-0721-3>
- Zhang, H. Y., Liu, Y. H., Wang, H. Q., Xu, J. H., & Hu, H. T. (2008). Puerarin protects PC12 cells against beta-amyloid-induced cell injury. *Cell Biology International*, 32, 1230–1237. <https://doi.org/10.1016/j.cellbi.2008.07.006>
- Zhang, X., Xiong, J., Liu, S., Wang, L., Huang, J., Liu, L., Yang, J., Zhang, G., Guo, K., Zhang, Z., Wu, P., Wang, D., Lin, Z., Xiong, N., & Wang, T. (2014). Puerarin protects dopaminergic neurons in Parkinson's disease models. *Neuroscience*, 280, 88–98. <https://doi.org/10.1016/j.neuroscience.2014.08.052>
- Zhao, C., Chan, H. Y., Yuan, D., Liang, Y., Lau, T. Y., & Chau, F. T. (2011). Rapid simultaneous determination of major isoflavones of *Pueraria lobata* and discriminative analysis of its geographical origins by principal component analysis. *Phytochemical Analysis*, 22, 503–508. <https://doi.org/10.1002/pca.1308>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Adolfo, L. M., Burks, D., Rao, X., Alvarez-Hernandez, A., & Dixon, R. A. (2022). Evaluation of pathways to the C-glycosyl isoflavone puerarin in roots of kudzu (*Pueraria montana lobata*). *Plant Direct*, 6(9), e442. <https://doi.org/10.1002/pld3.442>