



# *Iris domestica* (iso)flavone 7- and 3'-O-Glycosyltransferases Can Be Induced by CuCl<sub>2</sub>

Xiang Zhang<sup>1,2†</sup>, Yan Zhu<sup>1,2†</sup>, Jun Ye<sup>1,2</sup>, Ziyu Ye<sup>1,2</sup>, Ruirui Zhu<sup>1,2</sup>, Guoyong Xie<sup>1,2</sup>, Yucheng Zhao<sup>1,2\*</sup> and Minjian Qin<sup>1,2\*</sup>

<sup>1</sup> Department of Resources Science of Traditional Chinese Medicines, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing, China, <sup>2</sup> Key Laboratory of Modern Traditional Chinese Medicines (Ministry of Education), China Pharmaceutical University, Nanjing, China

## OPEN ACCESS

### Edited by:

María Pilar López-Gresa,  
Universitat Politècnica de  
València, Spain

### Reviewed by:

Alexander Wilson,  
Northern Michigan University,  
United States  
Bing-kai Hou,  
Shandong University, China

### \*Correspondence:

Minjian Qin  
minjianqin@163.com  
Yucheng Zhao  
zhaoyucheng1986@126.com

†These authors have contributed  
equally to this work

### Specialty section:

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

Received: 23 November 2020

Accepted: 18 January 2021

Published: 09 February 2021

### Citation:

Zhang X, Zhu Y, Ye J, Ye Z, Zhu R,  
Xie G, Zhao Y and Qin M (2021) *Iris*  
*domestica* (iso)flavone 7- and  
3'-O-Glycosyltransferases Can Be  
Induced by CuCl<sub>2</sub>.  
*Front. Plant Sci.* 12:632557.  
doi: 10.3389/fpls.2021.632557

In many plants, isoflavones are the main secondary metabolites that have various pharmacological activities, but the low water solubility of aglycones limits their usage. The O-glycosylation of (iso)flavones is a promising way to overcome this barrier. O-glycosyltransferases (UGTs) are key enzymes in the biosynthesis of (iso)flavonoid O-glycosides in plants. However, limited investigations on isoflavonoid O-UGTs have been reported, and they mainly focused on legumes. *Iris domestica* (L.) Goldblatt et Mabblerley is a non-legume plant rich in various isoflavonoid glycosides. However, there are no reports regarding its glycosylation mechanism, despite the *I. domestica* transcriptome previously being annotated as having non-active isoflavone 7-O-UGTs. Our previous experiments indicated that isoflavonoid glycosides were induced by CuCl<sub>2</sub> in *I. domestica* calli; therefore, we hypothesized that isoflavone O-UGTs may be induced by Cu<sup>2+</sup>. Thus, a comparative transcriptome analysis was performed using *I. domestica* seedlings treated with CuCl<sub>2</sub>, and eight new active BcUGTs were obtained. Biochemical analyses showed that most of the active BcUGTs had broad substrate spectra; however, substrates lacking 5-OH were rarely catalyzed. Real-time quantitative PCR results further indicated that the transcriptional levels of BcUGTs were remarkably induced by Cu<sup>2+</sup>. Our study increases the understanding of UGTs and isoflavone biosynthesis in non-legume plants.

**Keywords:** *Iris domestica* (*Belamcanda chinensis*) copper chloride, induce, (iso)flavones, transcriptome, UGT

## INTRODUCTION

Secondary metabolites derived from plants play very important roles in organisms' physiological activities, as well as human health. Isoflavonoid glycosides are a class of significant secondary metabolites mainly found in legume plants, which are associated with the interactions between legumes and both symbiotic and pathogenic microorganisms (Krämer et al., 1984; Graham et al., 1990; Udomsuk et al., 2010; Li et al., 2014; Clúa et al., 2018; Karre et al., 2019). However, the low water solubility of isoflavonoid aglycones is a barrier to the clinical applications. Glycosylation is an effective way to increase water solubility and bioavailability. For example, the apparent solubility of puerarin 7-O-glycoside is 18-fold that of puerarin (Jiang et al., 2008), and genistein 7-O-glycoside showed greater oral bioavailability than genistein (Kwon et al., 2007). O-glycosylation is the main glycosylation form, and O- UDP- sugar glycosyltransferases (O-UGTs) are responsible for the

biosynthesis of various O-glycosides in plants. However, only several genes with isoflavone O-glycosylation ability have been reported, and moreover, they are mainly investigated in leguminous plants. For instance, eight *Medicago* UGTs have been identified as being active against various isoflavones and flavones using a functional genomics approach (Modolo et al., 2007). In soybean (*Glycine max*), GmUGT4 is highly specific for isoflavones, while GmUGT1 and GmUGT7 have broad substrate spectra (Funaki et al., 2015). *Pueraria lobate* UGT1 plays a role in isoflavone 7-O-glycosylation (Li et al., 2014), while PIUGT2 is a bifunctional isoflavone O-UGT responsible for both the 4'-O- and 4',7-O-glycosylation of isoflavonoids (Wang et al., 2016). In non-legumes, isoflavone O-UGTs have only been reported in pomegranate and rice. UGT84A23, UGT84A24, and PgUGT95B2 cloned from pomegranate exhibit 7-O-glycosylation activity of genistein *in vitro* (Ono et al., 2016; Wilson et al., 2019). In the monocotyledonous plant rice, UGT709A4 and UGT706D1 could use isoflavones as substrates *in vitro*, and produce isoflavonoid 7-O-glycosides (Ko et al., 2008).

*Iris domestica* (L.) Goldblatt et Mabberley [syn. *Belamcanda chinensis* (L.) DC] forms a traditional Chinese medicine and isoflavones are the main active ingredients. Previous *in vitro* investigations showed that the isoflavones in rhizomes have a number of biological activities, including anti-mutagenic, anti-inflammatory, anti-angiogenic and anti-tumor activities, as well as inhibiting lipid peroxidation and scavenging free radicals (Seidlová-Wuttke et al., 2004; Zhang et al., 2005; Kang et al., 2009; Wozniak et al., 2010; Wang et al., 2013; Xie et al., 2014). Among the active isoflavones, O-glycosylated compounds, such as tectoridin and iridin, are the main ingredients (Lee et al., 2011; Chen et al., 2014). This implies that some isoflavone O-UGT genes may be abundant in *I. domestica*. In addition, it is unusual that *I. domestica* accumulates isoflavonoid glycosides as a non-legume. Thus, it is important for us to explore the mechanisms of glycosylation in *I. domestica*. However, *I. domestica* is a monocotyledonous plant; therefore, it is very difficult to obtain the active BcUGTs using homology cloning strategies owing to their low homology levels. Transcriptomes have aided in the discovery of isoflavone-related-UGTs in *I. domestica*. However, after analyzing the transcriptomes of six *I. domestica* organs, we did not identify any active (iso)flavone genes (Tian et al., 2018).

Previously, we found that the isoflavonoid glycoside content in *I. domestica* calli was significantly induced by Cu<sup>2+</sup> stress. Specifically, when the isoflavones contents of calli were monitored for 49 days, tectoridin and iridin increased up to 1.14 mg/g and 0.95 mg/g, respectively, at 42 days after a 1.2-mM CuCl<sub>2</sub> treatment. However, no tectoridin or iridin was detected in the control group (Zhu et al., 2020). The same dynamic content variation was observed in the adventitious roots of *Iris germanica* after a CuCl<sub>2</sub> treatment (Tomoyoshi et al., 2005). Thus, CuCl<sub>2</sub> appears to promote the biosynthesis of isoflavone glycosides by inducing related gene expression. Consequently, in this study, we chose CuCl<sub>2</sub> as the elicitor to induce the

expression of (iso)flavone 7-O-UGT genes in *I. domestica*. Using RNA sequencing (RNA-seq) and differentially expressed gene analyses, several BcUGT complementary DNAs (cDNAs), which belonged to the UGT706L family, were obtained. Biochemical analyses showed that eight active BcUGTs shared similar coding DNA sequences, but they exhibited different catalytic activities against different substrates. For most substrates, 7-O-glycoside compounds were the main products. Interestingly, a new compound was catalyzed by BcUGT4 and BcUGT5, and catalysis by the BcUGTs was 5-OH dependent. In addition, real-time quantitative PCR was performed to predict the functions of BcUGTs in *I. domestica* growth, especially its mechanism in CuCl<sub>2</sub>-induced gene expression. Our study elucidated the mechanisms of isoflavonoid glycoside biosynthesis in a non-legume plant through Cu<sup>2+</sup> induction. Furthermore, it offers the potential to engineer new compounds using synthetic biological strategies.

## MATERIALS AND METHODS

### Plant Materials and Stress Treatment

Newly harvested seeds of *I. domestica* were collected from the Botanic Garden of China Pharmaceutical University, Nanjing, China (31°54'36.36 N, 118°54'38.42 E). The seeds were identified by Prof. Minjian Qin (China Pharmaceutical University, Nanjing, China). After being sown in nutrient soil mixed with perlite, the seeds were placed in an artificial climate box at an ambient temperature of 25°C under a 14-h/10-h photoperiod. After 45 days of culturing, all the seedlings were transferred into liquid 1/2 Murashige and Skoog medium (MS, Murashige and Skoog, 1962) for 48 h. Then, 1.5 mM CuCl<sub>2</sub> was added to the medium, while nothing was added to the control group medium. The concentration of CuCl<sub>2</sub> was established in accordance with to the results of our previous copper stress-induced experiment (Zhu et al., 2020). Roots were collected after a 24-h CuCl<sub>2</sub> treatment for RNA-seq and gene expression analyses. Another induction experiment was carried out to rule out the effect of the chlorine ion in the induction, 48-day-old seedlings were transferred into liquid 1/2 MS for 48 h. Then 1.5 mM MgCl<sub>2</sub>, 1.5 mM CaCl<sub>2</sub>, 1.5 mM ZnCl<sub>2</sub>, and 1.5 mM CuCl<sub>2</sub> was added to the medium respectively, while nothing was added to the control group medium. Roots were collected after a 24-h treatment for gene expression analyses. All the samples were collected in three independent biological replicates. Tissues from different organs (flower, fruit, leaf, rhizome, and root) were collected separately in three biological replicates from three independent plants, these samples were frozen immediately in liquid nitrogen after washing and maintained at -80°C until use.

### RNA Extraction and cDNA Library Preparation

Total RNAs were extracted using an RNA simple total RNA kit (TIANGEN, Beijing, China), and the roots were collected from the seedlings treated with CuCl<sub>2</sub> and the control group. RNA degradation and contamination were examined using 1% agarose gels. RNA purity was checked using a NanoPhotometer<sup>®</sup> spectrophotometer (Implen, Carlsbad, CA, USA). The RNA

**Abbreviations:** cDNA, complementary DNA; CuCl<sub>2</sub>, copper chloride; UDPG, uridine diphosphate glucose; UGT, UDP-sugar glycosyltransferase.

concentration was measured using a Qubit<sup>®</sup> 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Sequencing libraries were generated using the NEB Next Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Seattle, WA, USA).

## Illumina Sequencing, Data Filtering, Transcriptome Assembly, and Gene Functional Annotation

Six sequencing libraries of *I. domestica* were subjected to the Illumina Novaseq 6000 platform (Illumina, San Diego, CA, USA) to generate paired-end reads. After removing low-quality reads and reads containing adapters or poly-N, we obtained clean data that were then assembled into contigs using the Trinity assembler (Grabherr et al., 2011). A gene functional annotation was performed in seven public databases: NCBI non-redundant protein sequences, NCBI non-redundant nucleotide sequences, Pfam, euKaryotic Ortholog Groups/Clusters of Orthologous Groups of proteins, Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes, and Gene Ontology.

## Screening of Candidate BcUGTs

Clean data were mapped back onto the assembled transcriptome. The read count for each gene was obtained from the mapping results using the RNA-seq by Expectation–Maximization (RSEM) software package (Li and Dewey, 2011). The expression level of each transcript in each sample was described in terms of fragments per kilobase of exon model per million mapped reads, which is a calculation of normalized read counts. A differential expression analysis of two samples was performed using the DESeq R package (Anders and Huber, 2010). Genes having an adjusted *p*-value < 0.005 and |log<sub>2</sub> Foldchange| > 1 were considered as differentially expressed. To search for genes involved in the biosynthesis of isoflavone 7-O-glycosides in *I. domestica*, we established a local database using the transcriptome data, and then performed a local BLAST algorithm-based search to identify genes that shared similar conserved sequences with isoflavone 7-O-UGT genes of legume plants. To further determine the evolutionary relationships among these genes, the BcUGTs proteins were aligned with sequences of other plant (iso)flavone O-UGT genes having known functions using MEGA7 (<https://www.megasoftware.net/>). A phylogenetic tree was constructed using the neighbor-joining method, with a bootstrap of 1,000. The evolutionary distances were computed using the Poisson correction (Kumar et al., 2016).

## Functional Analysis of BcUGTs Candidates *in vitro*

The open reading frames of BcUGTs were cloned individually into the prokaryotic expression vector pCold-TF, which helps the proteins fold correctly owing to the expression of the coupled trigger factor (chaperone). Consequently, the efficient expression of soluble proteins that are difficult to deal with

in other systems may be achieved. The primers used for the cloning are listed in **Supplementary Table 1**. All the BcUGT-pCold-TF constructs were transferred into *Escherichia coli* after sequencing. The recombinant strains were then inoculated into 100 mL Luria Broth liquid culture medium at 37°C and 220 rpm in an orbital shaker incubator. When the OD<sub>600</sub> of the strains reached 0.6, 0.5 mM isopropyl 1-β-D thiogalactoside (IPTG) was added into the culture to induce the expression of recombinant proteins. This was followed by further incubation at 15°C for 24 h. The recombinant *E. coli* cells were harvested by centrifugation (10,000 g for 10 min at 4°C), resuspended with 50 mM phosphate-buffered saline (PBS; pH 8.0) and disrupted using a sonicator. The recombinant proteins were purified using Ni-NTA columns (Sangon, Shanghai, China). The purified proteins were subjected to SDS-PAGE in accordance with the method of Laemmli (1970), and the protein concentrations were determined using the Bradford method (Bradford, 1976).

## Enzyme Assays

The catalytic UGT reactions were performed in 200 μL. The reaction components included 2 mM uridine diphosphate glucose (UDPG; Sigma-Aldrich, Shanghai, China), 1 mM dithiothreitol, 50 mM PBS, 10 μg of the purified recombinant protein, and 100 μM (iso)flavonoid acceptors. The reactions were incubated for 1 h at 30°C and then 400 μL ethyl acetate was added to stop the reactions. Ethyl acetate was vaporized to dryness in the vacuum drying oven, and finally, the samples were dissolved in 100 μL methanol for HPLC analysis. Firstly, we calculated the amount of residual substrate, then the ratio of residual substrate in total initial substrate was obtained, and the ratio is the substrate conversion rate. For kinetic studies of BcUGTs, typical assays containing 2 mM UDPG, 1 mM dithiothreitol, 50 mM PBS, 2 μg purified recombinant protein, and varying concentrations of substrates were performed. The reactions were incubated for 10 min at 30°C, and the solutions subjected to HPLC were prepared as above. Kinetic parameters were determined by hyperbolic regression analyses using the Prism program (GraphPad, <https://www.graphpad.com/>). The reaction products were analyzed using an Agilent 1100 HPLC system on SB C18 columns (5 μm; 250 × 4.6 mm; Agilent Technologies) in which solvent A was 0.1% formic acid in water and solvent B was 100% acetonitrile. The substrates and products were eluted using a linear gradient of 30% B to 60% B in 15 min at a flow rate of 1 mL/min. The chromatograms were obtained with detections at 270 nm (for isoflavones), 280 nm (for flavanones), 340 nm (for flavones), and 360 nm (for flavonols).

## Homology Modeling and Auto-Docking Analysis

We chose BcUGT4 to generate a protein model using SWISS-MODEL (<https://swissmodel.expasy.org/>, Biasini et al., 2014), the templates were searched using BLAST (Camacho et al., 2009) and HHblits (Steinegger et al., 2019). We used GMQE (Global Model Quality Estimation) to estimate the template quality, GMQE score is expressed as a number between 0 and 1, higher numbers indicate higher reliability. The quality of protein model was evaluated *via* GMQE and QMEAN (Benkert et al., 2008).

Docking of naringenin and liquiritigenin to BcUGT4 model was performed using AutoDock Tools (Trott and Olson, 2010, version 1.5.6). We used the default parameters in AutoDock Tools to prepare the ligands and macromolecule, then ran AutoDock to acquire the docking structures. Finally, we used PyMOL to present the docking results (<https://pymol.org/2/>).

## Transcriptional Analysis

The RNAs of seedling roots treated with MgCl<sub>2</sub>, CaCl<sub>2</sub>, ZnCl<sub>2</sub>, and CuCl<sub>2</sub> were isolated as above. Then, different organs of *I. domestica* and different parts of rhizomes were extracted to detect the expression levels of BcUGTs, and samples were all collected in three independent biological replicates. The quality and the quantity of total RNAs were determined using a NanoDrop Spectrophotometer (Eppendorf, Germany) and 1% agarose gels. The first-strand cDNA for real-time PCR was synthesized using HiScript II Q RT SuperMix for qPCR (Vazyme, Nanjing, China). The PCR reaction was carried out in a total volume of 20 μL containing ChamQ SYBR qPCR Master Mix (Vazyme) on an ABI StepOne plus Real-time PCR system (Applied Biosystem, USA), with cycling parameters 95°C for 30 s and 40 cycles of 95°C for 5 s and 60°C for 30 s. The relative quantification method ( $\Delta\Delta^{-CT}$ ) was used to calculate the expression level of each unigene (Livak and Schmittgen, 2001). The specificity of primers was tested using a melting curve program. The slope of the relative standard curve was calculated to ensure the amplification efficiency. The primers used for qPCR are listed in **Supplementary Table 2**, and each gene has three biological replications and three technical replications.

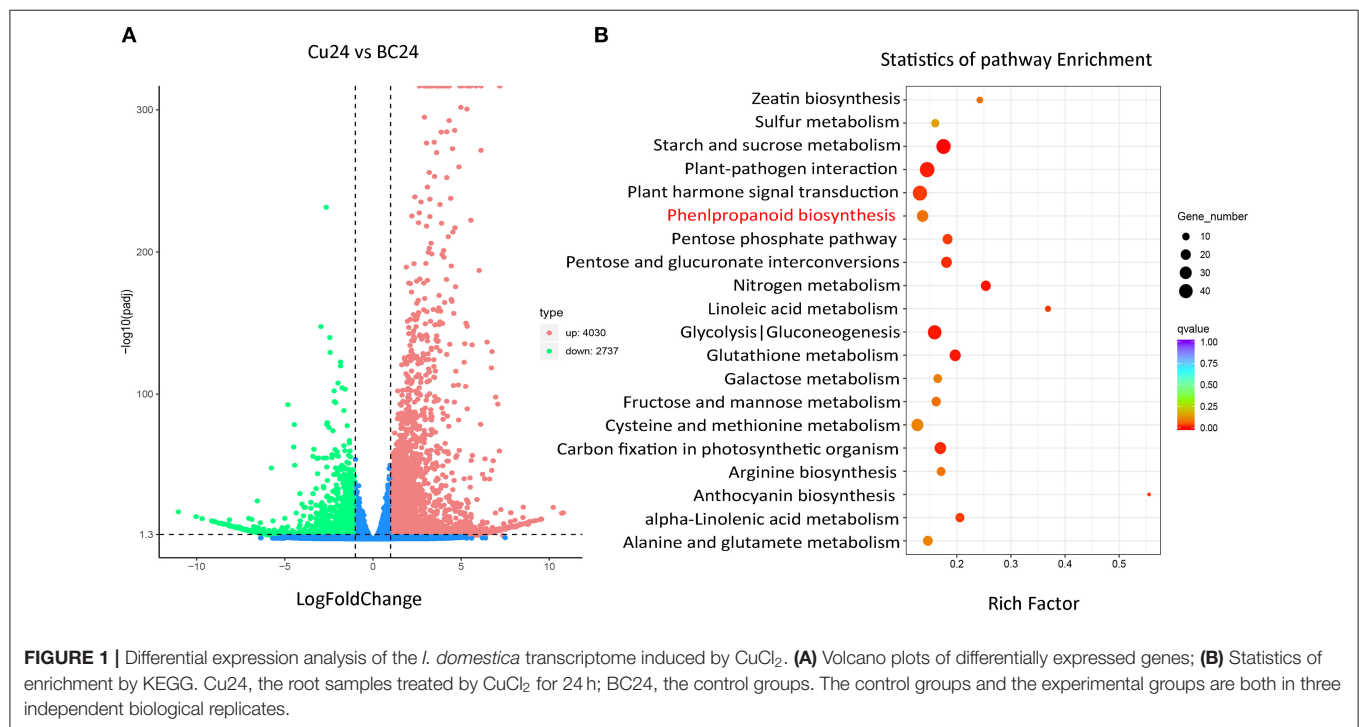
## RESULTS

### De novo Sequencing and Functional Annotation of Unigenes

Our previous study showed that calli of *I. domestica* treated with CuCl<sub>2</sub> accumulated 2.09 mg/g of tectoridin and iridin, while no isoflavonoid glycosides were detected in the control group. Thus, 45-day-old *I. domestica* seedling roots treated with CuCl<sub>2</sub> were used for Illumina sequencing after the RNA quality met specific criteria (**Supplementary Figure 1**). The adapter, unknown and low-quality reads were removed to obtain clean reads. In total, 25.03 and 25.57 Gb clean bases from the control and treated groups, respectively, were obtained and used for transcript assembly (**Supplementary Table 3**). A total of 73,132 unigenes were annotated in the seven referenced public databases, with 24.6% annotated unigenes being homologous to *Asparagus officinalis*, followed by *Elaeis guineensis* and *Phoenix dactylifera*. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes annotations assigned the unigenes to different metabolic pathways. This was the basis for the preliminarily functional predictions for these unigenes (**Supplementary Figures 2, 3**). These results provided a good foundation for the screening of candidate BcUGTs.

### Selection of Candidate BcUGTs Induced by CuCl<sub>2</sub>

Genes involved in the same biosynthetic route are temporally and spatially induced by effective treatments, which indicates that the gene functions may be characterized by a differential

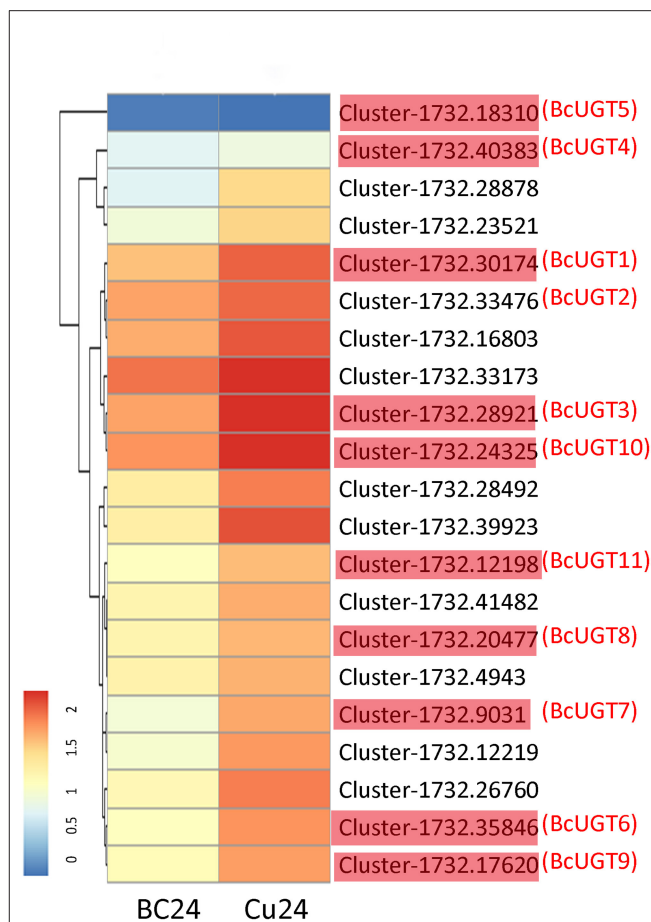


expression analysis. Using the standard criteria of an adjusted  $p$ -value  $< 0.005$  and  $|\log_2 \text{Foldchange}| > 1$ , we obtained 6,767 differentially expressed genes. Among them, 4,030 were up-regulated in *I. domestica* roots treated with CuCl<sub>2</sub>, and many unigenes were enriched in phenylpropanoid biosynthesis (Figure 1). Then, 19 candidate unigenes were selected from the up-regulated unigenes annotated as O-UGTs. To fully screen for possible unigenes involved in isoflavonoid glycoside biosynthesis, two putative UGTs (Cluster-1732.40383 and Cluster-1732.18310) were identified using a local BLAST algorithm-based search, and they shared similar conserved sequences with GmUGTs. The expression levels of these 21 candidate genes were up-regulated after CuCl<sub>2</sub> treatments, except Cluster-1732.18310, and the expression levels of most of the selected genes increased remarkably, having  $\log_2 \text{Foldchanges} > 2$  (Figure 2). The sequences of the 21 genes were used as the query in a BLAST algorithm-based search of National Center for Biotechnology Information (NCBI) databases to predict their gene families and possible functions. We finally selected 11 BcUGTs putatively identified as (iso)flavone-O-UGTs to determine whether they were capable of transferring UDPG to (iso)flavone aglycones (Figure 2).

## Functional Characterization of Active BcUGTs *in vitro*

For *in vitro* biochemical assays, the open reading frame sequences of the selected 11 BcUGTs were all cloned independently into the prokaryotic expression vector pCold-TF and expressed in *E. coli* cells. The recombinant proteins purified with Ni-NTA columns showed strong and single protein bands at 105 kDa by SDS-PAGE (Supplementary Figure 4). The deduced molecular weights of the BcUGTs were ~53 kDa, with the fusion protein having an additional soluble label and His-tag of ~52 kDa. The protein bands identified by SDS-PAGE indicate that these soluble proteins may be used to determine the enzymes' activity levels and kinetic properties. The functional BcUGT assays were performed against the different (iso)flavone substrates: naringenin, apigenin, kaempferol, irigenin and tectorigenin. BcUGT1–8 showed activities toward (iso)flavones, while Cluster-1732.24325, Cluster-1732.17620 and Cluster-1732.12198 did not transfer a glycosyl group onto (iso)flavones, although they were annotated as (iso)flavone UGTs. BcUGT1–8, except BcUGT8, shared very high deduced amino acid sequence similarities, ranging from 81.7 to 97.3% (Supplementary Table 4), all of them were submitted to the UGT Nomenclature Committee and designated as UGT706L family genes.

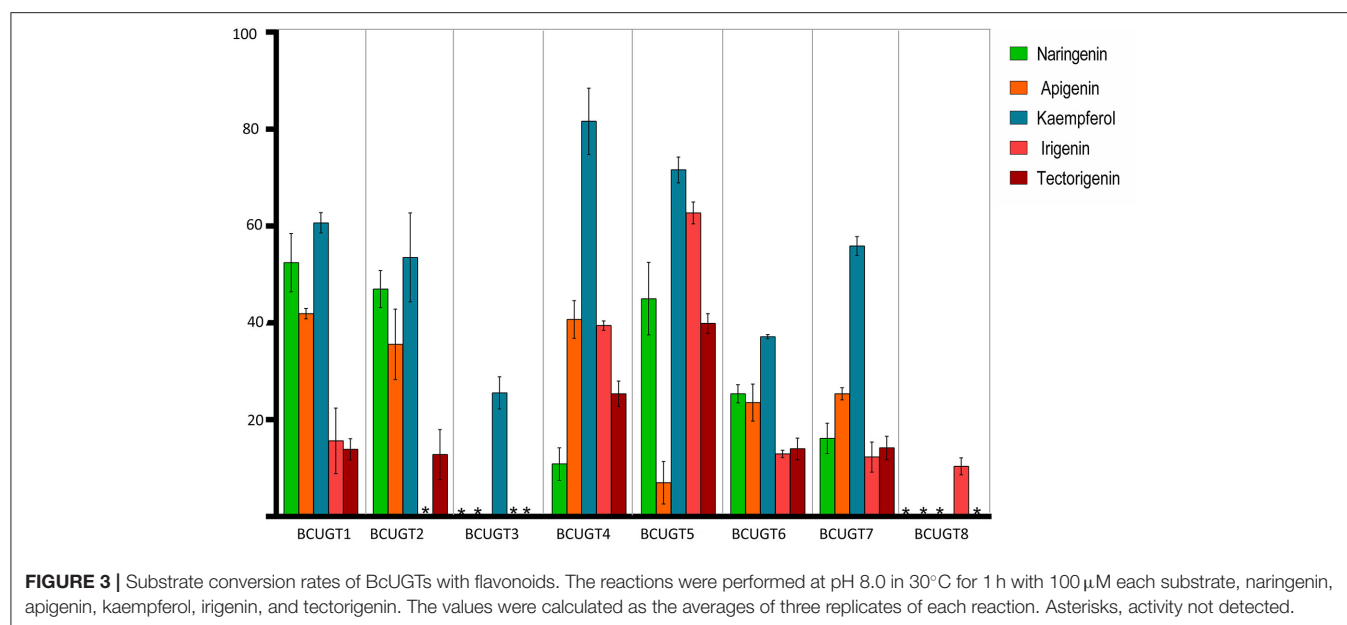
The substrate conversion rates were determined using the flavones: naringenin, liquiritigenin, apigenin, baicalein, luteolin, quercetin, kaempferol, irigenin, tectorigenin, genistein, daidzein, and 6,7,4'-trihydroxyisoflavone (Supplementary Figure 5). BcUGT1, –2 and –4–7 displayed broad activities toward various (iso)flavone substrates, with flavonols being the most preferred (Supplementary Figure 6). However, for different types of substrates, BcUGTs exhibited different catalytic efficiencies. As shown in Figure 3, BcUGT1 exhibited similar catalytic activities toward different sugar acceptors, as did BcUGT2, while BcUGT4



**FIGURE 2** | Analysis of differentially expressed unigenes from *I. domestica*. Hierarchical clustering and corresponding heatmaps of the differentially expressed (iso)flavone glycosyltransferases across all pairwise library comparisons. The clusters in red (BcUGT1–11) were selected to verify the activity against flavones *in vitro*. Cu24, the root samples treated by CuCl<sub>2</sub> for 24 h; BC24, the control groups. The control groups and the experimental groups are both in three independent biological replicates.

was analogous to BcUGT5. Among these substrates, BcUGT3 only accepted luteolin and kaempferol as sugar acceptors, while BcUGT8 only exhibited relatively low activities toward luteolin and irigenin. The other flavonoids were all inert substrates for these enzymes (Table 1). The conversion rates of BcUGT4 and BcUGT5 for irigenin and tectorigenin were greater than other BcUGTs, by 2–5 times (Table 2).

None of the BcUGTs displayed any activity when daidzein and 6,7,4'-trihydroxyisoflavone were used as substrates. These isoflavones are rich in soybean, and their chemical structures lack 5-OH (Table 1). BcUGT1 and BcUGT5 converted 2.5-fold and 5.6-fold more naringenin, respectively, to their 7-O-glycosides compared with liquiritigenin. In addition, for daidzein and 6,7,4'-trihydroxy-isoflavone, none of the BcUGTs showed any activity. Thus, the existence of 5-hydroxy of (iso)flavone aglycones was very important for catalysis by the BcUGTs. In this study, BcUGT4 was chosen to investigate



**FIGURE 3** | Substrate conversion rates of BcUGTs with flavonoids. The reactions were performed at pH 8.0 in 30°C for 1 h with 100 μM each substrate, naringenin, apigenin, kaempferol, irigenin, and tectorigenin. The values were calculated as the averages of three replicates of each reaction. Asterisks, activity not detected.

**TABLE 1** | Substrate conversion rates of BcUGTs with various flavonoids.

Substrate	Conversion rate %							
	BcUGT1	BcUGT2	BcUGT3	BcUGT4	BcUGT5	BcUGT6	BcUGT7	BcUGT8
<b>Flavanone</b>								
Naringenin	52.4 ± 6.0	47.0 ± 3.8	0.0	47.0 ± 3.8	45.0 ± 7.5	25.3 ± 1.9	16.1 ± 3.1	0.0
Liquiritigenin	20.6 ± 5.0	11.2 ± 3.4	0.0	16.1 ± 3.1	7.8 ± 2.4	16.2 ± 1.1	11.3 ± 1.5	0.0
<b>Flavone</b>								
Apigenin	41.9 ± 1.1	35.6 ± 7.3	0.0	40.7 ± 3.9	7.0 ± 4.4	25.5 ± 3.8	25.4 ± 1.2	0.0
Baicalein	42.5 ± 3.5	41.7 ± 1.5	0.0	9.1 ± 3.9	7.9 ± 6.4	15.9 ± 3.7	25.9 ± 2.1	0.0
Luteolin	29.9 ± 3.5	28.0 ± 4.0	13.7 ± 3.9	44.8 ± 9.2	21.3 ± 8.2	18.6 ± 3.7	35.2 ± 1.7	9.7 ± 2.4
<b>Flavonol</b>								
Quercetin	53.6 ± 7.5	62.5 ± 4.3	0.0	39.2 ± 7.1	26.3 ± 4.2	12.2 ± 0.8	12.4 ± 1.8	0.0
Kaempferol	60.6 ± 2.1	53.5 ± 9.2	25.5 ± 3.3	81.6 ± 6.9	71.6 ± 2.6	37.1 ± 0.5	55.8 ± 1.9	0.0
<b>Isoflavone</b>								
Tectorigenin	13.9 ± 2.2	12.8 ± 5.2	0.0	25.3 ± 2.7	39.9 ± 2.0	14.0 ± 2.2	14.2 ± 2.4	0.0
Iriogenin	15.6 ± 6.8	0.0	0.0	39.4 ± 1.0	62.7 ± 2.3	12.9 ± 0.8	12.3 ± 3.1	10.4 ± 1.8
Genistein	11.1 ± 2.8	16.2 ± 7.6	0.0	39.9 ± 1.0	13.5 ± 1.9	24.8 ± 0.2	17.5 ± 2.3	0.0
Daidzein	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6,7,4'-Trihydroxy-isoflavone	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Reactions of BcUGTs with 100 μM substrate at pH 8.0 and 30°C for 1 h. Data are the means ± SDs of one experiment with three technical replicates.

the critical role of the 5-hydroxy in protein binding with substrates, we identified the top 50 templates according to the coverage, identity, and GMQE. A newly reported TcCGT1 (PDB) crystal structure was used as the template for homologous molecular modeling in SWISS-MODEL, the GMQE score of TcCGT was 0.75. A C-glycosyltransferase found in *Trollius chinensis*, TcCGT1, regio-specifically produces 8-C-glycosylation flavones with broad substrate spectra, and site-directed mutagenesis at just two residues switched C- to O-glycosylation (He et al., 2019). The GMQE of the model was 0.72, and the GMEAN was -2.59, GMEAN

scores > -4 is an indication of models with high quality. Afterward, we modeled the structure of BcUGT4 docking with liquiritigenin and naringenin (Figure 4). The spacious binding pocket explained the broad substrate spectra of BcUGT4, with seven residues interacting with naringenin through hydrogen bonds having 1.8–2.8 Å distances. The residues GLN-394 and TYR-391 interacted with naringenin C-5-OH (2.8 Å, 2.1 Å), but for liquiritigenin, five residues, HIS-369, GLU-393, TRP-372, ASN-373, and GLU-377, were discovered around the ligand and they were also identified in the naringenin docking.

**TABLE 2** | Kinetic parameters of BcUGTs with tectorigenin and irigenin.

Substrate	kat(s <sup>-1</sup> )	km(mM)	kcat/km(M <sup>-1</sup> s <sup>-1</sup> )
<b>BcUGT1</b>			
Tectorigenin	57.29 ± 4.94	0.592 ± 0.049	9.68 × 10 <sup>4</sup>
Irigenin	58.88 ± 3.24	0.210 ± 0.013	2.80 × 10 <sup>5</sup>
<b>BcUGT2</b>			
Tectorigenin	70.10 ± 9.86	0.392 ± 0.007	1.79 × 10 <sup>5</sup>
Irigenin	–	–	–
<b>BcUGT4</b>			
Tectorigenin	155.18 ± 12.26	0.115 ± 0.023	1.35 × 10 <sup>6</sup>
Irigenin	57.23 ± 18.21	0.190 ± 0.007	3.01 × 10 <sup>5</sup>
<b>BcUGT5</b>			
Tectorigenin	116.03 ± 6.98	0.201 ± 0.015	5.77 × 10 <sup>5</sup>
Irigenin	296.52 ± 11.8	0.121 ± 0.009	2.45 × 10 <sup>6</sup>
<b>BcUGT7</b>			
Tectorigenin	15.86 ± 7.53	1.269 ± 0.009	1.98 × 10 <sup>4</sup>
Irigenin	17.73 ± 9.89	0.894 ± 0.011	1.98 × 10 <sup>4</sup>
<b>BcUGT8</b>			
Tectorigenin	–	–	–
Irigenin	11.97 ± 1.31	1.423 ± 0.021	8.41 × 10 <sup>3</sup>

Kinetic parameters were determined at pH 8.0 in 30°C with UDP-glucose, as described in the Materials and methods section. Data are the means ± SDs of one experiment with three technical replicates.

## Structural Analysis of the New Compound

Reactions of BcUGTs with most of the sugar acceptors yielded single products, each of which was identified as its 7-O-glycoside on the basis of their co-chromatography with authentic samples using HPLC. However, as shown in **Figure 5**, when irigenin was used as acceptor substrate, BcUGT4 and BcUGT5 tended to yield double products, one of which (1a) was identified as iridin according to the mass spectrum and co-chromatography with authentic iridin, the molecular formula C<sub>24</sub>H<sub>26</sub>O<sub>13</sub> was deduced from the HR-ESI-MS ion peak at *m/z* 521.1832 [M-H]<sup>-</sup>. The other main product (1b) was confirmed to be irigenin 3'-O glycoside using HR-ESI-MS, <sup>1</sup>H-NMR, <sup>13</sup>C-NMR, HMBC, and NEOSY. This structure has not been found previously in *I. domestica* or any other plant. Its UV (MeOH) λ<sub>max</sub> values were 268 nm, which is the typical absorption peak of isoflavones, and its IR absorptions at 3,424, 1,627, and 1,510 revealed the existence of hydroxyl and carbonyl groups (**Supplementary Figure 7**). The molecular formula C<sub>24</sub>H<sub>26</sub>O<sub>13</sub> was deduced from the HR-ESI-MS ion peak at *m/z* 545.1275 [M+Na]<sup>+</sup> (calcd for C<sub>24</sub>H<sub>26</sub>O<sub>13</sub> as 522.4544) (**Supplementary Figure 8**). In the <sup>1</sup>H NMR spectrum, the most representative signal was the characteristic signal for isoflavone at δ8.43. There were three methoxyl group signals at δ3.81, δ3.76, and δ3.75. The signals at δ13.01 and δ10.93 were assigned to the hydroxy groups at the five and seven positions of isoflavone, and the signal of the end group hydrion of glucose occurred at δ4.91. The signals at δ6.96 (1H, d, 2.8) and δ6.96 (1H, d, 2.8) represented two AX coupling systems, which indicated the existence of incomplete symmetry tri-substituted groups in the B ring of the isoflavone skeleton (**Supplementary Figure 9**). The <sup>13</sup>C NMR spectrum indicated 24 carbon resonances, which

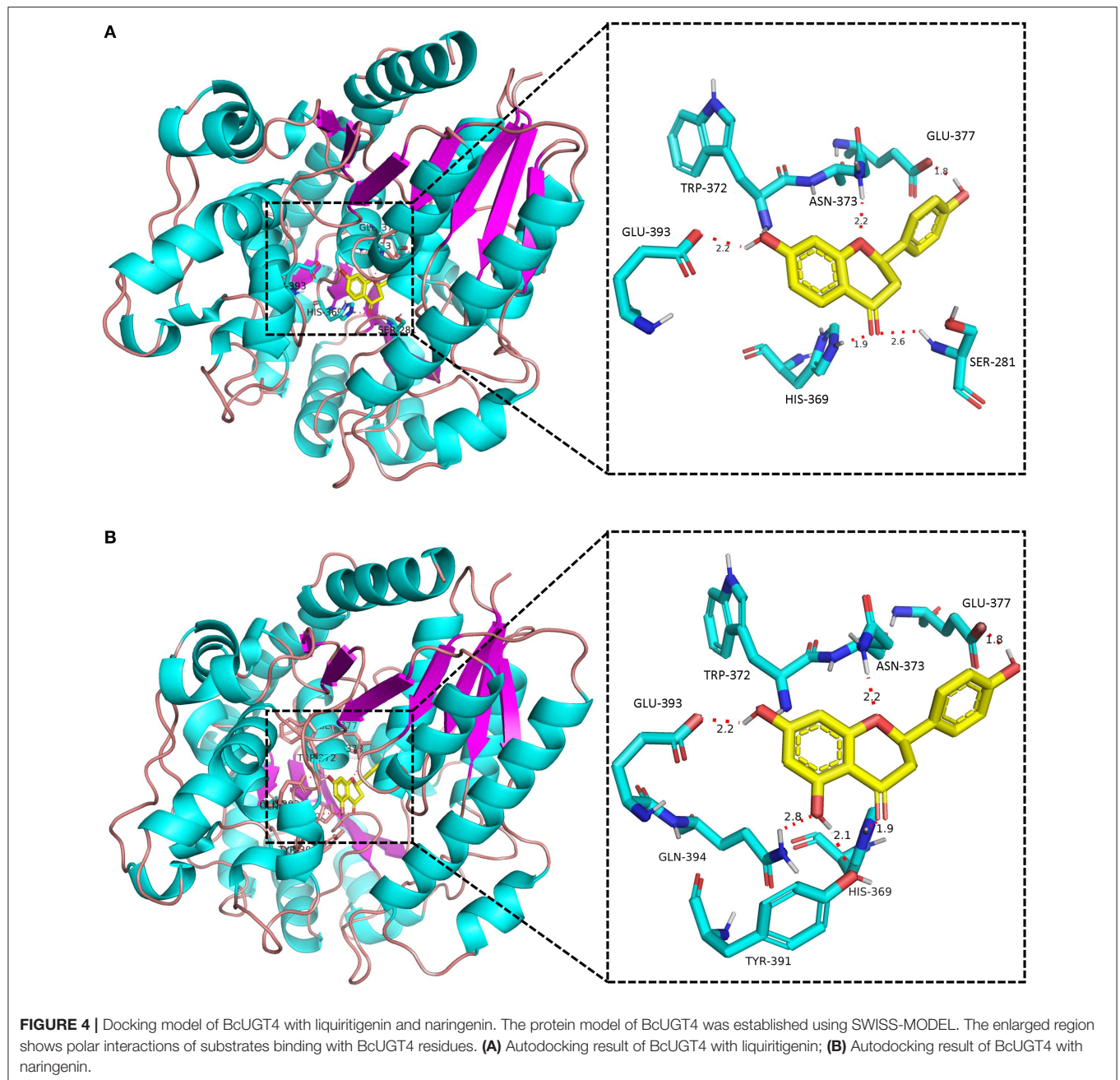
contained 19 carbon resonances of the isoflavone skeleton, 3 methoxyl group signals, and 6 carbon signals of glucose (**Supplementary Figure 10**). In the HMBC experiment, the resonating signal at δH 4.91 correlated with δC 151.2, which indicated that the glucose was transferred to isoflavonoid 3'-hydroxyl position (**Supplementary Figure 11**). <sup>1</sup>H-NMR and <sup>13</sup>C-NMR data for compound 1b and key HMBC correlations were shown in **Supplementary Table 5**. The NEOSY spectrum showed a resonating signal correlation between two hydrogens at a distance of two chemical bonds, and the resonating signal of δH 4.91 displayed a strong correlation with δH 6.96, which implied that 1' H of glucose was close to 2' H in irigenin. The NEOSY result further validated the structure of 3'-O glycoside (**Supplementary Figure 12**).

## Kinetic Analysis of BcUGTs for Isoflavones

Although most of the BcUGTs exhibited broad substrate ranges and preference for the kaempferol aglycone, no data showed the existence of flavone O-glycosides or flavonol O-glycosides in *I. domestica*, only small amounts of flavones and flavonols have been reported, such as hispidulin, apigenin, rhamnazin, kaempferol, and quercetin (Jin et al., 2008). Consequently, we selected irigenin and tectorigenin as the substrates for kinetic parameter determinations (**Table 2**). The UDPG concentration was set at 2 mM, and the substrate concentration ranged from 10 to 400 μM for the assays. The reaction time was strictly controlled at 10 min. The km values of BcUGT4 and BcUGT5 for irigenin and tectorigenin were the lowest among all the active BcUGTs, and BcUGT5 exhibited the highest kcat/km ratio for irigenin, at 2.45 × 10<sup>6</sup>, which was approximately four times the kcat/km ratio for tectorigenin. Using BcUGT4, the kcat/km ratio for tectorigenin was approximately 4.5 times that for irigenin. Thus, although BcUGT4 shared a 94.2% sequence identity with BcUGT5, their substrate specificities were quite different. In addition, the sequence identity between BcUGT1 and BcUGT3 was 97.3%, but their catalytic specificities were totally different. BcUGT3 did not show activity against isoflavones, but BcUGT1 exhibited broad substrate ranges, although its kcat/km ratio for irigenin was just 2.8 × 10<sup>5</sup>. BcUGT8 shared only about 70% sequence identity with other BcUGTs, so irigenin and tectorigenin may not be its natural substrates, and BcUGT8 could almost not glycosylate tectorigenin (**Supplementary Table 4**).

## Phylogenetic Analysis of BcUGTs

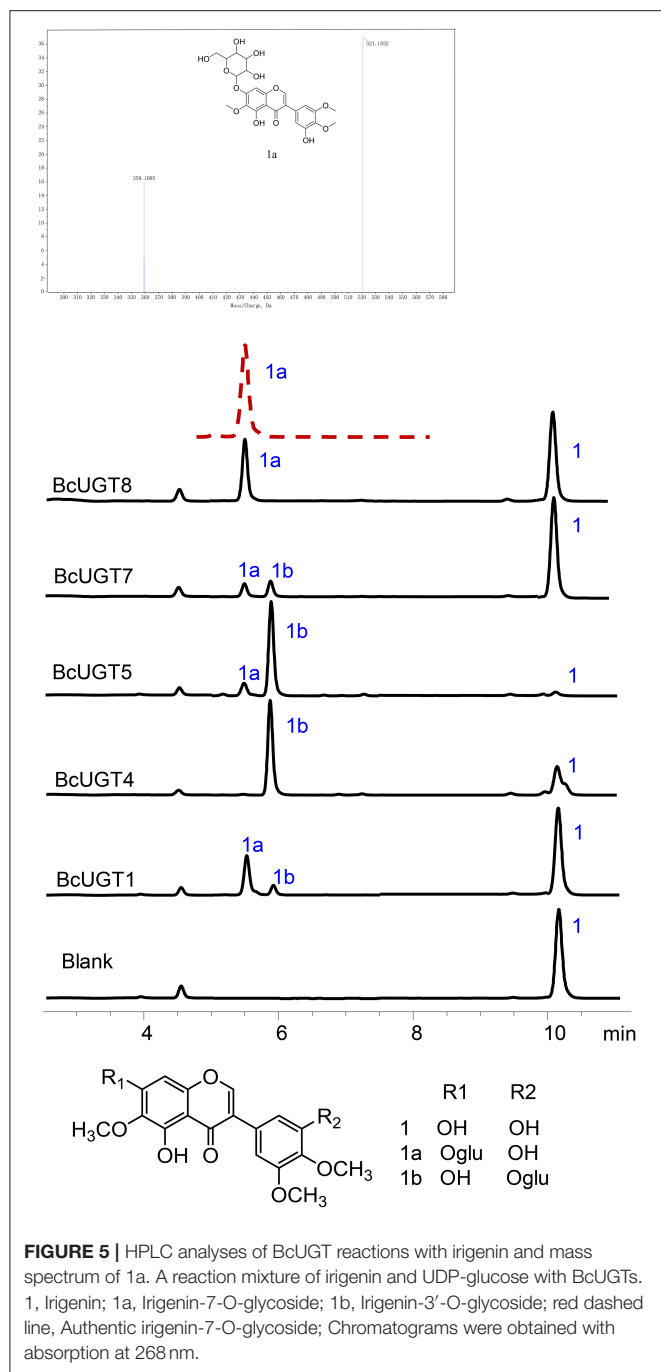
According to the phylogenetic tree, BcUGTs showed a closer relationship with the UDP-UGT706 family (**Figure 6**). UGT706D1 and UGT709A4 were reported to be capable of forming isoflavone 7-O glucosides *in vitro*, while UGT706C2 could use flavones to produce flavone 3-O glucosides (Ko et al., 2008), all BcUGTs showed a closer relationship with UGT706D1 and UGT706C2, and sequence identities were 40–50%. Meanwhile, BcUGTs had 40–50% amino acid sequence similarities with the UGTs of *G. max* and *P. lobata*, which have been confirmed to have activities toward (iso)flavone aglycones. Although all the BcUGTs selected for phylogenetic tree construction had similar sequence similarities with UGT88 family group and UGT706 family group, they clustered into rice



UGT706 clade, perhaps owing to the genetic distance between monocots and eudicots. Genes chosen in the phylogenetic tree were mainly those of *G. max* and *P. lobata*, but they exhibited different substrate specificities and regio-selectivities. PIUGT1 and PIUGT13 showed enzyme activities for isoflavone substrates at the 7-hydroxy group, while other flavonoids were hardly accepted. PIUGT2 was later shown to be the isoflavone 4',7-O-diglucosides' UGT. GmUGT4 was specific for isoflavones at the 7-hydroxy group, but other GmUGTs showed broad glycosyl-acceptor specificities (Li et al., 2014; Funaki et al., 2015; Wang et al., 2016). The PgUGT84A23 and PgUGT84A24

played important roles in formation of hydrolyzable tannins (HTs) in pomegranate, but they also exhibited genistein 7-O-glycosylation activity *in vitro* (Ono et al., 2016), and another PgUGT95B2 showed genistein 7-O-glycosylation activity *in vitro* as well (Wilson et al., 2019). These reports suggested BcUGTs' ability to glycosylate different flavones or isoflavones at their 7-hydroxy positions, but *in vitro* experiments revealed that their specificities and regio-selectivities for these sugar acceptors differed. A further division of the clusters among the 10 unigenes from *I. domestica* revealed that Cluster-1732.24325, Cluster-1732.17620, and BcUGT8 did not fall into the same





subcluster with the other BcUGTs, and Cluster-1732.24325 and Cluster-1732.17620 had no activities toward flavones. Additionally, BcUGT8 only displayed a weak catalytic efficiency toward irigenin. BcUGT1 and BcUGT2 clustered into the same branch, and BcUGT6 and BcUGT7 displayed a close genetic relationship. Cluster-1732.12198 had a relatively higher homology to PIUGT43, which was reported to be a daidzein C-glycosyltransferase cloned from *P. lobata* (Wang et al., 2017). However, to date, there have been no isoflavone C-UGTs detected

in *I. domestica*, and Cluster-1732.12198 did not transfer glucose to any flavones.

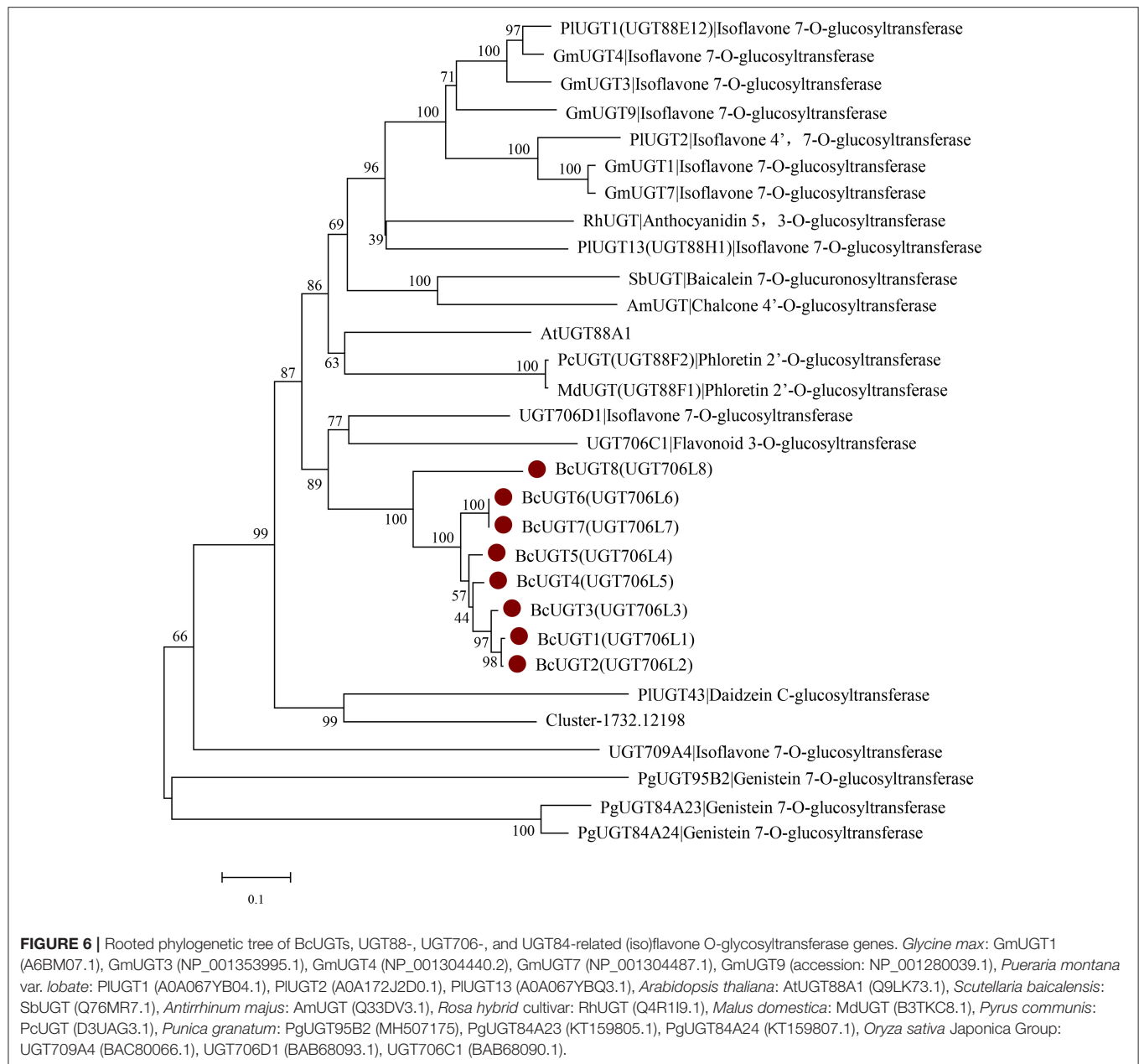
## Expression Level Analyses

The transcription levels of the active BcUGTs (BcUGT1–8) in *I. domestica* seedling roots were detected using quantitative real-time RT-PCR with primers having good specificity and amplification efficiency levels (**Supplementary Figure 13**). After treating with CuCl<sub>2</sub>, the transcript abundances in seedling roots increased to different degrees, especially for BcUGT1 and BcUGT3, in addition, the expression levels of BcUGT1 and BcUGT3 were significantly higher than those of other BcUGTs (**Figure 7A**). The results showed good correlations with the FPKM (fragments per kilobase of exon model per million mapped reads) of each BcUGT calculated from the RNA-seq data. We found that the expression levels of all the active BcUGTs in seedlings not receiving CuCl<sub>2</sub> treatments were very low, and some of them, such as BcUGT5 and BcUGT8, were almost undetectable. Thus, CuCl<sub>2</sub> plays a crucial role in inducing the expression of the eight active BcUGTs. The expression levels of BcUGTs in seedlings receiving MgCl<sub>2</sub>, ZnCl<sub>2</sub> and CaCl<sub>2</sub> treatments were very low as well (**Supplementary Figure 14**). Moreover, ZnCl<sub>2</sub> and MgCl<sub>2</sub> suppressed the expression of all BcUGTs compared with the control group, CaCl<sub>2</sub> showed almost no effect in inducing the expression of the eight BcUGTs. These data indicated that Cu<sup>2+</sup> plays a crucial role in inducing the expression of the BcUGTs, and we could rule out the effect of the chlorine ion in the induction.

The transcript abundances of all the BcUGTs in the underground parts were significantly greater than in the aerial parts, which correlated well with the isoflavone glycosides' distribution pattern (**Figure 7B**). The rhizome accumulated the most isoflavones, but the BcUGTs were more highly expressed in roots than rhizomes (**Figure 8**). To further explore the relationship of BcUGT expression levels between roots and rhizomes, we divided the triennial rhizomes into five parts, lateral, bulb, vertical, sprout primordium, and cork layer (Chen et al., 2014). The vertical part is the first and sprout primordium is the last during rhizome growth, and the lateral and bulb are biennial parts in rhizomes. Among the five rhizome parts, the sprout primordium and vertical part displayed higher BcUGT expression levels than other parts. However, it was obvious that the lateral and bulb parts had the most roots, while the other parts, especially the sprout primordium, had almost no roots. In the cork layer, previous experiments revealed that isoflavone aglycones were the most abundant and almost no isoflavone glycosides were detected (Chen et al., 2014). This may be explained by the low BcUGT expression levels.

## DISCUSSION

To date, the biosynthesis of isoflavonoids, including isoflavone 7-O-, 4'-O-, and C-UGTs, has been mainly studied in legumes, with the only exceptions being studies from pomegranate and rice. Thus, little is known regarding the genes involved in isoflavonoid biosynthesis in other plants containing abundant isoflavones.

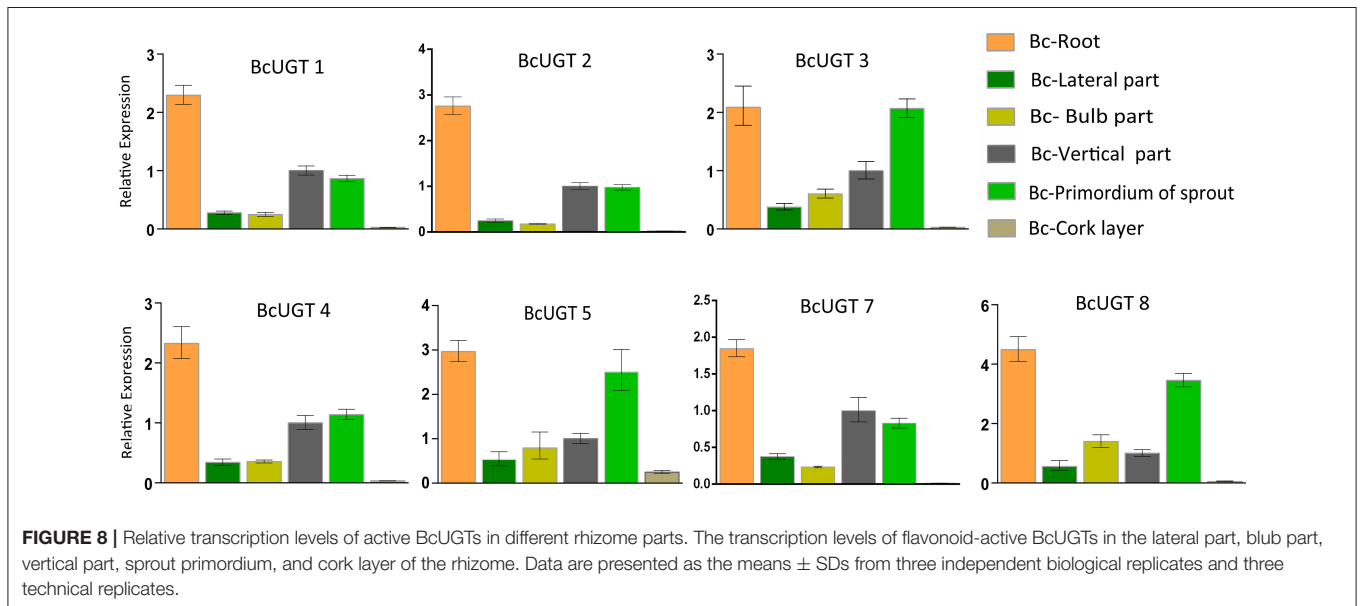
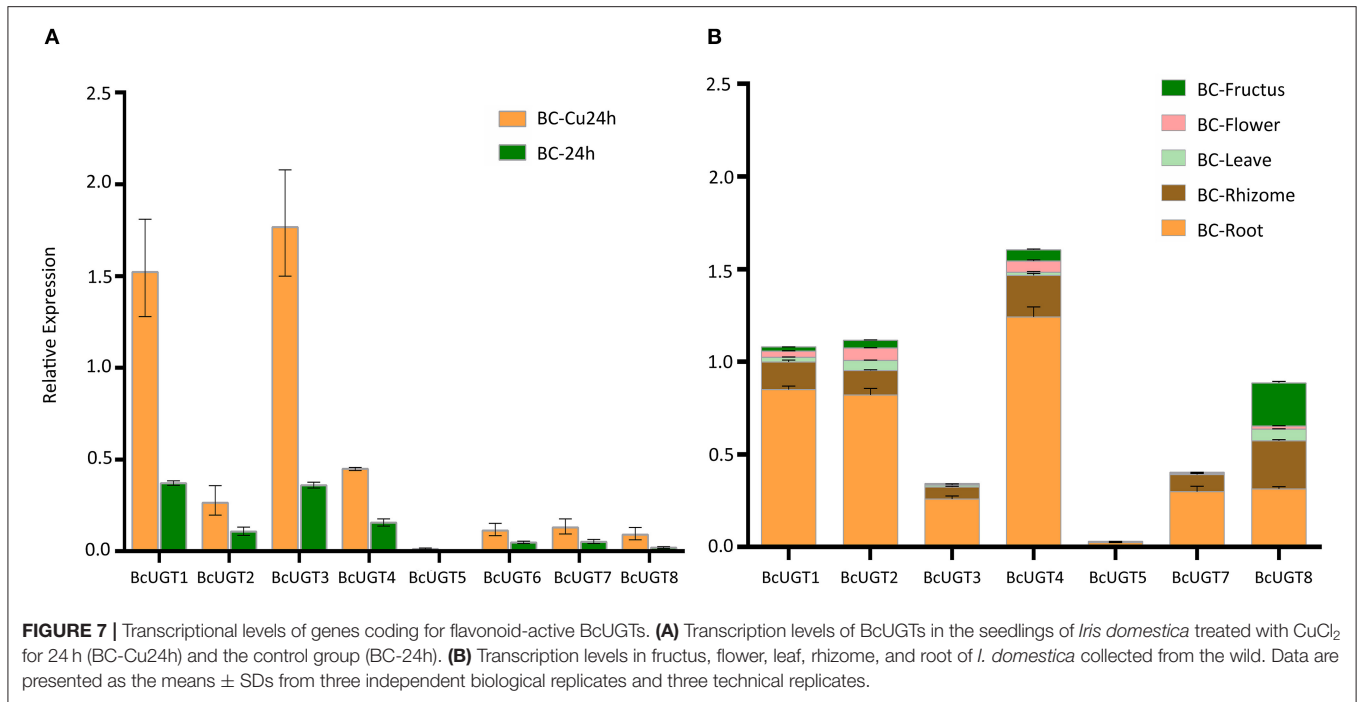


*Iris domestica* accumulates many isoflavonoid glycosides, such as tectoridin, iridin, iristectorin A, and iristectorin B, in its rhizomes. Using copper stress, we obtained eight unigenes that were active toward (iso)flavones. A sequence alignment showed that the active BcUGTs shared a very high sequence homology (Supplementary Figure 15).

## Cu<sup>2+</sup> Mechanism for Up-Regulating BcUGT Expression

Plants are continuously exposed to abiotic stresses during their growth in the wild, including various heavy metal stresses (Piscopo et al., 2016). Consequently, they have evolved a series of skills to deal with the rapid changes in the environment.

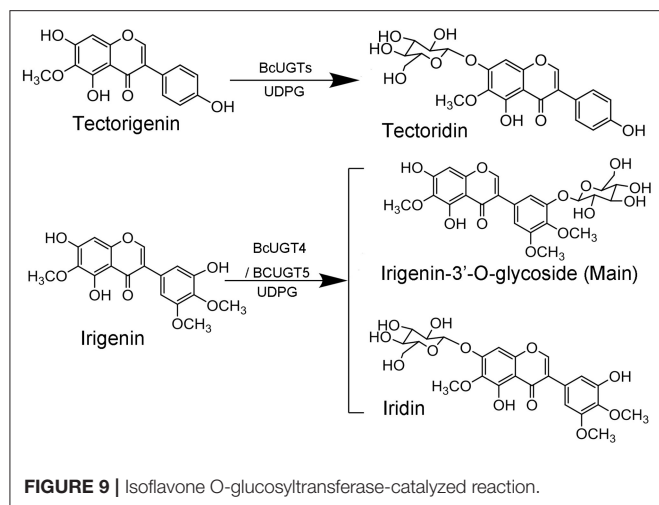
Elements such as Fe, Cu, Zn, Mo, and Ni are considered to be essential heavy metals (HMs) (Hall, 2002; Arif et al., 2016), but excessive HMs may be toxic for plants, resulting in large-scale molecular, biochemical and physiological responses (Hayat et al., 2012; Hussain et al., 2013; Clemens and Ma, 2016). The perception of stress stimuli is the initial plant response when exposed to toxic HMs. Then, the amplified signal is transduced and transmitted into cells to bring about cascade reactions of molecular, biochemical and physiological activities (Wani et al., 2018). Through transcription analyses of *I. domestica* roots, many P450 genes related to the biosynthesis and catabolism of methyl jasmonate and allene oxide (Maksymiec et al., 2005), including CYP94 (cluster1732.31596 and cluster1732.31816) and CYP74



(cluster 1732.20141), were found to be up-regulated after CuCl<sub>2</sub> treatments (Utsunomiya et al., 2000; Koo et al., 2011).

At the biochemical level, plants need more secondary metabolites to cope with excess HMs, resulting in the high expression levels of many unigenes involved in the biosynthesis of secondary metabolites. CuCl<sub>2</sub> was first applied to induce isoflavonoids in *P. lobata* stems, and the daidzein and genistein contents in the CuCl<sub>2</sub>-treated stem increased 5–10-fold compared with the control group (Hakamatsuka et al., 1991). The cotyledons of *Lupinus albus* have been

treated with a fungal elicitor and CuCl<sub>2</sub>, which induced the isoflavone 2'-hydroxygenistein. When challenged with CuCl<sub>2</sub>, isoflavone glycosidic conjugates are largely degraded to isoflavone aglycones, which indicates the important roles of isoflavonoid glycosides in plants coping with Cu<sup>2+</sup> toxicity (Wojtaszek and Stobiecki, 1997). Additionally, the dynamic change in isoflavonoid glycosides contents in *I. domestica* calli suggested that CuCl<sub>2</sub> promotes the biosynthesis of isoflavone glycosides during the first stage of CuCl<sub>2</sub> stress. Afterward, when the Cu<sup>2+</sup> concentration in cells reaches a high level, the



isoflavone glycoside is hydrolyzed to its aglycone. Thus, we speculated that isoflavone glycosides play very important roles in plant defenses against Cu<sup>2+</sup> stress. Through the differential expression analysis, several up-regulated unigenes annotated as (iso)flavone 7-O-UGTs, which belonged to the UGT706 family, were cloned. Then, these unigenes were expressed in a recombinant *E. coli* expression system, and their functions were verified using *in vitro* catalytic experiments. These results implied that the high CuCl<sub>2</sub> concentration enhanced the expression of (iso)flavone O-UGTs in *I. domestica*.

### Substrate Preference and Regio-Specificity of BcUGTs

Many flavone UGTs show more regio-specificity than substrate specificity (Kim et al., 2006; Li et al., 2014; Funaki et al., 2015; Wang et al., 2016). In soybean, GmUGT1 and GmUGT7 efficiently catalyze regio-specific glucosyl transfers to isoflavones at their 7-hydroxy positions, and they accept naringenin, apigenin, quercetin, and kaempferol as substrates (Funaki et al., 2015). In an important crop, *Camellia sinensis*, CsUGT73A20 was identified as a flavonoid UGT. CsUGT73A20 transfers UDP-glucose to produce 3-O- and 7-O-glycosides, depending mainly on the reaction's *in vitro* pH value. Kaempferol is the most preferred substrate, but naringenin and apigenin are also accepted (Zhao et al., 2017). Recently, three new (iso)flavonoid glycosyltransferases (PIUGT4, -15 and -57) were identified in *P. lobate*. PIUGT15 and PIUGT57 exhibit much higher isoflavone specificity than other 7-O-UGTs found in *P. lobate* (Wang X. et al., 2019).

Kaempferol is the most preferred acceptor among the 12 substrates tested for all the BcUGTs. The major product yielded by BcUGTs was (iso)flavone 7-O-glycoside, which indicates that the BcUGTs preferred to transfer the sugar moiety to the hydroxy of ring A, but no (iso)flavone 4'-O-glycosides or 4',7-O-diglycosides were obtained. Thus, most of the BcUGTs had no activity toward the hydroxy of ring B. Tectorigenin is enriched in the rhizomes and roots of *I. domestica*, and all the BcUGTs, except BcUGT3 and BcUGT8, converted tectorigenin to tectoridin,

which exists in large amounts in *I. domestica*. Iridenin is another main isoflavone aglycone in the rhizomes and roots of *I. domestica*, but we unexpectedly found that the main product of BcUGT4 and BcUGT5 using iridenin as the substrate was not iridenin (Figure 9). We identified its structure as iridenin 3'-O-glycoside. This is a new compound that has never been identified in *I. domestica* or other plants. However, other BcUGTs hardly glycosylated iridenin at the 3'-OH position. BcUGT4 yielded two different products using luteolin as the substrate, but the main product was luteolin-7-O-glycoside. Thus, we speculated that the binding conformation of the substrate determines the glycosylation position for BcUGT4 and BcUGT5.

Interestingly, substrates lacking a 5-hydroxy were hardly glycosylated by BcUGTs. Recently, a promiscuous flavonoid-3-O-UGT (sb3GT1) was identified in *Scutellaria baicalensis*, and it accepts many sugar donors to glycosylate 17 flavonols. Through homologous molecular modeling and docking, 18 residues within a 5-Å distance of the binding pocket were selected as candidates, and residue G15 was identified as the critical residue in the sugar donor promiscuity of Sb3GT1 by comparisons with the docking results of *Vitis vinifera* GT1 (Wang Z. L. et al., 2019). In this study, the docking results of BcUGT4 with liquiritigenin and naringenin indicated that the residues GLN-394 and TYR-391 could form hydrogen bond with 5-hydroxy, which may stabilize the substrate in the binding pocket, further study could confirm the role of the 5-hydroxy and the coordinating amino acids in substrate binding.

### Predicting the Function of BcUGTs

The different expression levels of BcUGTs in the transcriptome of *I. domestica* seedlings indicate that BcUGTs may play important roles in responses to abiotic stresses. In plants, glycosylation is a key modification that allows the formation of a myriad of secondary metabolites, which play many important roles in plant responses to challenges faced during growth and development (Bowles et al., 2005; Gachon et al., 2005; Campos et al., 2019). PIUGT4, -15 and -57 are up-regulated in *P. lobate* roots after methyl jasmonate treatments (Wang X. et al., 2019), and PIUGT1 may be related to the defense response against the MeJA treatment. In this paper, all the BcUGTs were up-regulated after CuCl<sub>2</sub> stress, especially BcUGT1 and BcUGT2, which have highly homologous nucleotide sequences and broad substrate ranges. The unigenes encoding the core metabolic pathway usually form small families consisting of unique genes. For example, isoflavone synthases belong to the CYP93C family, which only consists of a few members in leguminous plants (Jung et al., 2000). This is in sharp contrast to the multiple gene families of UGTs, which tend to contain hundreds of members. In addition, the broad substrate ranges may ensure that plants respond rapidly to both abiotic and biotic stresses in the environment.

For wild *I. domestica*, we found that the transcription levels of most BcUGTs in the vertical parts (triennial rhizome) and sprout primordium were greater than in the lateral and bulb parts (biennial rhizome). Additionally, we found that lateral and bulb parts had more roots than the other parts; consequently, we speculated that this is the reason for the differences in BcUGT expression among these rhizome parts. Roots form the main

organ that delivers water and nutrients from the soil to the aerial plant parts; therefore, we hypothesized that many isoflavone glycosides synthesized in roots by BcUGTs are delivered to the lateral and bulb parts, while the sprout primordium and vertical part have no isoflavone glycoside source. Consequently, they have to increase the BcUGT expression levels to defend against the stress. Recently, isoflavone membrane transformation in *Trifolium pratense* was shown to be related to the presence of ABC proteins and vesicular transport (Kubes et al., 2019). By analyzing global transcriptomes of six organs from *I. domestica*, many ABC transporter family unigenes were obtained specifically from the roots.

## DATA AVAILABILITY STATEMENT

The sequencing project has been deposited at the SRA database under the accession number PRJNA596865. The SRA accession numbers are SRR10743030 and SRR107430. The accession numbers for the eight BcUGTs are as follows: BcUGT1 (MN894539), BcUGT2 (MN894540), BcUGT3 (MN894541), BcUGT4 (MN894542), BcUGT5 (MN894543), BcUGT6 (MN894544), BcUGT7 (MN894545), and BcUGT8 (MN894546).

## REFERENCES

- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi: 10.1186/gb-2010-11-10-r106
- Arif, N., Yadav, V., Singh, S., Singh, S., Ahmad, P., Mishra, R., et al. (2016). Influence of high and low levels of plant-beneficial heavy metal ions on plant growth and development. *Front. Env. Sci. Switz.* 4:69. doi: 10.3389/fenvs.2016.00069
- Benkert, P., Tosatto, S. C. E., and Schomburg, D. (2008). QMEAN: a comprehensive scoring function for model quality assessment. *Proteins* 71, 261–277. doi: 10.1002/prot.21715
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 42, W252–W258. doi: 10.1093/nar/gku340
- Bowles, D., Isayenkova, J., Lim, E. K., and Poppenberger, B. (2005). Glycosyltransferases: managers of small molecules. *Curr. Opin. Plant Biol.* 8, 254–263. doi: 10.1016/j.pbi.2005.03.007
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:1. doi: 10.1186/1471-2105-10-421
- Campos, L., López-Gresa, M. P., Fuertes, D., Bellés, J. M., Rodrigo, I., and Lisón, P. (2019). Tomato glycosyltransferase Twi1 plays a role in flavonoid glycosylation and defence against virus. *BMC Plant Biol.* 19:450. doi: 10.1186/s12870-019-2063-9
- Chen, Y. J., Liang, Z. T., Zhu, Y., Xie, G. Y., Tian, M., Zhao, Z. Z., et al. (2014). Tissue-specific metabolites profiling and quantitative analyses of flavonoids in the rhizome of *Belamcanda chinensis* by combining laser-microdissection with UHPLC-Q/TOF-MS and UHPLC-Q/Q-MS. *Talanta* 130, 585–597. doi: 10.1016/j.talanta.2014.07.004
- Clemens, S., and Ma, J. F. (2016). Toxic heavy metal and metalloid accumulation in crop plants and foods. *Annu. Rev. of Plant Biol.* 67, 489–512. doi: 10.1146/annurev-arplant-043015-112301

## AUTHOR CONTRIBUTIONS

XZ, YaZ, GX, and MQ designed research. XZ, JY, and ZY performed research and all authors analyzed the data. XZ, YaZ, YuZ, and MQ wrote the paper. All authors read and approved the final manuscript.

## FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 81872958, 81373918, and 81703637).

## ACKNOWLEDGMENTS

We thank International Science Editing (<https://www.international-scienceediting.com/>) for editing this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.632557/full#supplementary-material>

- Clúa, J., Roda, C., Zanetti, M. E., and Blanco, F. A. (2018). Compatibility between legumes and rhizobia for the establishment of a successful nitrogen-fixing symbiosis. *Genes* 9:125. doi: 10.3390/genes9030125
- Funaki, A., Waki, T., Noguchi, A., Kawai, Y., Yamashita, S., Takahashi, S., et al. (2015). Identification of a highly specific isoflavone 7-O-glucosyltransferase in the soybean (*Glycine max* (L.) Merr.). *Plant Cell Physiol.* 56, 1512–1520. doi: 10.1093/pcp/pcv072
- Gachon, C. M., Langlois-Meurinne, M., and Saindrenan, P. (2005). Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends Plant Sci.* 10, 542–549. doi: 10.1016/j.tplants.2005.09.007
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Graham, T. L., Kim, J. E., and Graham, M. Y. (1990). Role of constitutive isoflavone conjugates in the accumulation of glyceollin in soybean infected with *Phytophthora-Megasperma*. *Mol. Plant Microbe Interact.* 3, 157–166. doi: 10.1094/MPMI-3-157
- Hakamatsuka, T., Ebizuka, Y., and Sankawa, U. (1991). Induced isoflavonoids from copper chloride-treated stems of *Pueraria lobata*. *Phytochemistry* 30, 1481–1482. doi: 10.1016/0031-9422(91)84191-T
- Hall, J. L. (2002). Cellular mechanisms for heavy metal detoxification and tolerance. *J. Exp. Bot.* 53, 1–11. doi: 10.1093/jexbot/53.366.1
- Hayat, S., Khaliq, G., Irfan, M., Wani, A. S., and Tripathi, B. N. (2012). Physiological changes induced by chromium stress in plants: an overview. *Protoplasma* 249, 599–611. doi: 10.1007/s00709-011-0331-0
- He, J. B., Zhao, P., Hu, Z. M., Liu, S., Kuang, Y., Zhang, M., et al. (2019). Molecular and structural characterization of a promiscuous C-glycosyltransferase from *Trollius chinensis*. *Angew Chem. Int. Edit.* 58, 11513–11520. doi: 10.1002/anie.201905505
- Hussain, M. B., Ali, S., Azam, A., Hina, S., Farooq, M. A., Ali, B., et al. (2013). Morphological, physiological and biochemical responses of plants to nickel stress: a review. *Afr. J. Agric. Res.* 8, 1596–1602. doi: 10.5897/AJAR12.407
- Jiang, J. R., Yuan, S., Ding, J. F., Zhu, S. C., Xu, H. D., Chen, T., et al. (2008). Conversion of puerarin into its 7-O-glycoside derivatives by *Microbacterium oxydans* (CGMCC 1788) to improve its water solubility and pharmacokinetic properties. *Appl. Microbiol. Biot.* 81, 647–657. doi: 10.1007/s00253-008-1683-z

- Jin, L., Chen, H. S., Jin, Y. S., Liang, S., Xiang, Z. B., and Lu, J. (2008). Chemical constituents from *Belamcanda chinensis*. *J. Asian Nat. Prod. Res.* 10, 89–94. doi: 10.1080/10286020701273619
- Jung, W., Yu, O., Lau, S. M. C., O'Keefe, D. P., Odell, J., Fader, G., et al. (2000). Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nat. Biotechnol.* 18, 208–212. doi: 10.1038/72671
- Kang, K., Lee, S. B., Jung, S. H., Cha, K. H., Park, W. D., Sohn, Y. C., et al. (2009). Tectoridin, a poor ligand of estrogen receptor  $\alpha$ , exerts its estrogenic effects via an ERK-dependent pathway. *Mol. Cells* 27, 351–357. doi: 10.1007/s10059-009-0045-8
- Karre, S., Kumar, A., Yogendra, K., Kage, U., Kushalappa, A., and Charron, J. B. (2019). HvWRKY23 regulates flavonoid glycoside and hydroxycinnamic acid amide biosynthetic genes in barley to combat Fusarium head blight. *Plant Mol. Biol.* 100, 591–605. doi: 10.1007/s11103-019-00882-2
- Kim, J. H., Kim, B. G., Park, Y., Ko, J. H., Lim, C. E., Lim, J., et al. (2006). Characterization of flavonoid 7-O-glucosyltransferase from *Arabidopsis thaliana*. *Biosci. Biotech. Biochem.* 70, 1471–1477. doi: 10.1271/bbb.60006
- Ko, J. H., Kim, B. G., Kim, J. H., Kim, H., Lim, C. E., Lim, J., et al. (2008). Four glucosyltransferases from rice: cDNA cloning, expression, and characterization. *J. Plant Physiol.* 165, 435–444. doi: 10.1016/j.jplph.2007.01.006
- Koo, A. J. K., Cooke, T. F., and Howe, G. A. (2011). Cytochrome P450 CYP94B3 mediates catabolism and inactivation of the plant hormone jasmonoyl-L-isoleucine. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9298–9303. doi: 10.1073/pnas.1103542108
- Krämer, R. P., Hindorf, H., Jha, H. C., Kallage, J., and Zilliken, F. (1984). Antifungal activity of soybean and chickpea isoflavones and their reduced derivatives. *Phytochemistry* 23, 2203–2205. doi: 10.1016/S0031-9422(00)80520-8
- Kubes, J., Skalicky, M., Tumova, L., Martin, J., Hejnak, V., and Martinkova, J. (2019). Vanadium elicitation of *Trifolium pratense* L. cell culture and possible pathways of produced isoflavones transport across the plasma membrane. *Plant Cell Rep.* 38, 657–671. doi: 10.1007/s00299-019-02397-y
- Kumar, S., Stecher, G., and Tamura, K. (2016). Mega7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Kwon, S. H., Kang, M. J., Huh, J. S., Ha, K. W., Lee, J. R., Lee, S. K., et al. (2007). Comparison of oral bioavailability of genistein and genistin in rats. *Int. J. Pharmacol.* 337, 148–154. doi: 10.1016/j.ijpharm.2006.12.046
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- Lee, Y. S., Kim, S. H., Kim, J. K., Lee, S., Jung, S. H., and Lim, S. S. (2011). Preparative isolation and purification of seven isoflavones from *Belamcanda chinensis*. *Phytochem. Anal.* 22, 468–473. doi: 10.1002/pca.1306
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323
- Li, J., Li, Z. B., Li, C. F., Gou, J. B., and Zhang, Y. S. (2014). Molecular cloning and characterization of an isoflavone 7-O-glucosyltransferase from *Pueraria lobata*. *Plant Cell Rep.* 33, 1173–1185. doi: 10.1007/s00299-014-1606-7
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-ΔΔC(T))</sup> method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Maksymiec, W., Wianowska, D., Dawidowicz, A. L., Radkiewicz, S., Mardarowicz, M., and Krupa, Z. (2005). The level of jasmonic acid in *Arabidopsis thaliana* and *Phaseolus coccineus* plants under heavy metal stress. *J. Plant Physiol.* 162, 1338–1346. doi: 10.1016/j.jplph.2005.01.013
- Modolo, L. V., Blount, J. W., Achnine, L., Naoumkina, M. A., Wang, X., and Dixon, R. (2007). A functional genomics approach to (iso)flavonoid glycosylation in the model legume *Medicago truncatula*. *Plant Mol. Biol.* 64, 499–518. doi: 10.1007/s11103-007-9167-6
- Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco cultures. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x
- Ono, N. N., Qin, X. Q., Wilson, A. E., Li, G., and Tian, L. (2016). Two UGT84 family glycosyltransferases catalyze a critical reaction of hydrolyzable Tannin Biosynthesis in pomegranate (*Punica granatum*). *PLoS ONE* 11, 1–25. doi: 10.1371/journal.pone.0156319
- Piscopo, M., Ricciardiello, M., Palumbo, G., and Troisi, J. (2016). Selectivity of metal bioaccumulation and its relationship with glutathione S-transferase levels in gonadal and gill tissues of *Mytilus galloprovincialis* exposed to Ni (II), Cu (II) and Cd (II). *Rend. Lincei Sci. Fis.* 27, 737–748. doi: 10.1007/s12210-016-0564-0
- Seidlová-Wuttke, D., Hesse, O., Jarry, H., Rimoldi, G., Thelen, P., Christoffel, V., et al. (2004). *Belamcanda chinensis* and the thereof purified tectorigenin have selective estrogen receptor modulator activities. *Phytomedicine* 11, 392–403. doi: 10.1016/j.phymed.2004.01.003
- Steinegger, M., Meier, M., Mirdita, M., Vöhringer, H., Haunsberger, S. J., and Söding, J. (2019). HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* 20:473. doi: 10.1186/s12859-019-3019-7
- Tian, M., Zhang, X., Zhu, Y., Xie, G. Y., and Qin, M. J. (2018). Global transcriptome analyses reveal differentially expressed genes of six organs and putative genes involved in (Iso)flavonoid biosynthesis in *Belamcanda chinensis*. *Front. Plant Sci.* 9:1160. doi: 10.3389/fpls.2018.01160
- Tomoyoshi, A., Masayuki, I., Toshio, A., and Shin-ichi, A. (2005). Isoflavonoid production by adventitious-root cultures of *Iris germanica*. *Plant Biotechnol.* 22, 207–215. doi: 10.5511/plantbiotechnology.22.207
- Trott, O., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461. doi: 10.1002/jcc.21334
- Udomsuk, L., Jarukamjorn, K., Tanaka, H., and Putalum, W. (2010). Improved isoflavonoid production in *Pueraria candollei* hairy root cultures using elicitation. *Biotechnol. Lett.* 33, 369–374. doi: 10.1007/s10529-010-0417-3
- Utsunomiya, Y., Jarukamjorn, K., Tanaka, H., and Putalum, W. (2000). Purification and inactivation by substrate of an allene oxide synthase (CYP74) from corn (*Zea mays* L.) seeds. *Phytochemistry* 53, 319–323. doi: 10.1016/S0031-9422(99)00534-8
- Wang, Q., Cheng, X. L., Zhang, D. Y., Gao, X. J., Zhou, L., Qin, X. Y., et al. (2013). Tectorigenin attenuates palmitate-induced endothelial insulin resistance via targeting ROS-associated inflammation and IRS-1 pathway. *PLoS ONE* 8:e66417. doi: 10.1371/journal.pone.0066417
- Wang, X., Fan, R. Y., Li, J., Li, C. F., and Zhang, Y. S. (2016). Molecular cloning and functional characterization of a novel (Iso)flavone 4',7-O-diglucoside glycosyltransferase from *Pueraria lobata*. *Front. Plant Sci.* 7:387. doi: 10.3389/fpls.2016.00387
- Wang, X., Li, C. F., Zhou, C., Li, J., and Zhang, Y. S. (2017). Molecular characterization of the C-glucosylation for puerarin biosynthesis in *Pueraria lobata*. *Plant J.* 90, 535–546. doi: 10.1111/tpj.13510
- Wang, X., Li, C. F., Zhou, Z. L., and Zhang, Y. S. (2019). Identification of three (Iso)flavonoid glucosyltransferases from *Pueraria lobata*. *Front. Plant Sci.* 10:28. doi: 10.3389/fpls.2019.00028
- Wang, Z. L., Wang, S., Xu, Z., Li, M. W., Chen, K., Zhang, Y. Q., et al. (2019). Highly promiscuous flavonoid 3-O-glucosyltransferase from *Scutellaria baicalensis*. *Org. Lett.* 21, 2241–2245. doi: 10.1021/acs.orglett.9b00524
- Wani, W., Masoodi, K. Z., Zaid, A., Wani, S. H., Shah, F., Meena, V. S., et al. (2018). Engineering plants for heavy metal stress tolerance. *Rend. Lincei Sci. Fis.* 29, 709–723. doi: 10.1007/s12210-018-0702-y
- Wilson, A. E., Wu, S., and Tian, L. (2019). PgUGT95B2 preferentially metabolizes flavones/flavonols and has evolved independently from flavone/flavonol UGTs identified in *Arabidopsis thaliana*. *Phytochemistry* 157, 184–193. doi: 10.1016/j.phytochem.2018.10.025
- Wojtaszek, P., and Stobiecki, M. (1997). Differential secretion and accumulation of isoflavonoids in *Lupinus albus* in response to fungal elicitor and CuCl<sub>2</sub>. *Plant Physiol. Biochem.* 35, 129–135.
- Wozniak, D., Janda, B., Kapusta, I., Oleszek, W., and Matkowski, A. (2010). Antimutagenic and anti-oxidant activities of isoflavonoids from *Belamcanda chinensis* (L.) DC. *Mutat. Res. Gen. Toxicol. Environ. Mut.* 696, 148–153. doi: 10.1016/j.mrgentox.2010.01.004
- Xie, G. Y., Zhu, Y., Shu, P., Qin, X. Y., Wu, G., Wang, Q., et al. (2014). Phenolic metabolite profiles and antioxidants assay of three Iridaceae medicinal plants for traditional Chinese medicine “She-gan” by on-line HPLC-DAD coupled with chemiluminescence (CL) and ESI-Q-TOF-MS/MS. *J. Pharmaceut. Biomed.* 98, 40–51. doi: 10.1016/j.jpba.2014.05.008
- Zhang, C. Z., Wang, S. X., Zhang, Y., Chen, J. P., and Liang, X. M. (2005). *In vitro* estrogenic activities of Chinese medicinal plants traditionally used for

- the management of menopausal symptoms. *J. Ethnopharmacol.* 98, 295–300. doi: 10.1016/j.jep.2005.01.033
- Zhao, X. Q., Wang, P. Q., Li, M. Z., Wang, Y. R., Jiang, X. L., Cui, L. L., et al. (2017). Functional characterization of a new tea (*Camellia sinensis*) flavonoid glycosyltransferase. *J. Agric. Food Chem.* 65, 2074–2083. doi: 10.1021/acs.jafc.6b05619
- Zhu, Y., Chen, Y., Zhang, X., Xie, G. Y., and Qin, M. J. (2020). Copper stress-induced changes in biomass accumulation, antioxidant activity and flavonoid contents in *Belamcanda chinensis* calli. *Plant Cell Tiss. Org.* 142, 299–311. doi: 10.1007/s11240-020-01863-w

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

Copyright © 2021 Zhang, Zhu, Ye, Ye, Zhu, Xie, Zhao and Qin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.