



# Genome-wide identification of glycosyltransferases converting phloretin to phloridzin in *Malus* species

Kun Zhou, Lingyu Hu, Pengmin Li, Xiaoqing Gong\*, Fengwang Ma\*

State Key Laboratory of Crop Stress Biology for Arid Areas, College of Horticulture, Northwest A&F University, Yangling, Shaanxi 712100, China

## ARTICLE INFO

### Keywords:

Dihydrochalcone  
Phloridzin  
UGT  
*Malus*  
Expression pattern  
Transgenic plants

## ABSTRACT

Phloridzin (phloretin 2'-O-glucoside) is the most abundant phenolic compound in *Malus* species, accounting for up to 18% of the dry weight in leaves. Glycosylation of phloretin at the 2' position is the last and key step in phloridzin biosynthesis. It is catalyzed by a uridine diphosphate (UDP)-glucose:phloretin 2'-O-glucosyltransferase (P2'GT), which directly determines the concentration of phloridzin. However, this process is poorly understood. We conducted a large-scale investigation of phloridzin accumulations in leaves from 64 *Malus* species and cultivars. To identify the responsible P2'GT, we performed a genome-wide analysis of the expression patterns of UDP-dependent glycosyltransferase genes (*UGTs*). Two candidates were screened preliminarily in *Malus* spp. cv. Adams (North American Begonia). Results from further qRT-PCR analyses of the genotypes showed a divergence in phloridzin production. Our assays of enzyme activity also suggested that MdUGT88F4 and MdUGT88F1 regulate the conversion of phloretin to phloridzin in *Malus* plants. Finally, when they were silenced in 'GL-3' ('Royal Gala'), the concentrations of phloridzin and phloretin (and trilobatin) were significantly reduced and increased, respectively.

## 1. Introduction

Dihydrochalcones (DHCs) are phenylpropanoids that are very similar in structure to chalcones, the intermediates in flavonoid formation (Fig. 1). The DHC phloridzin, a mono-glycoside of phloretin, has a glucose moiety attached at the 2'-OH position (Fig. 1). It is the most abundant phenolic compound in the leaves of apple (*Malus domestica*), accounting for 14–18% of the total dry weight for that tissue [1,2]. This makes apple unique in the plant kingdom because, in addition to members of the *Malus* genus, only a few other plants, i.e., *Camellia japonica* [3], *Fragaria × ananassa* [4], *Rosa canina* [5], and *Lithocarpus litseifolius* [6], can accumulate tiny amounts of phloridzin [7]. Even *Pyrus* sp., closely related to *M. domestica*, can not accumulate phloridzin and its precursor phloretin [8,9]. Within *Malus*, production of phloridzin varies among cultivars [10,11], developmental stages [12], tissue types [10], and external factors such as pathogen attacks [13]. Phloridzin is enriched in apple seeds, bark, and leaves, but only minimally detected in fruit [10,14].

Phloridzin and its derivatives are excellent antioxidants [15–17], and are widely investigated in the field of human health research [18,19]. Evidence from studies with diabetic mice has shown that

phloridzin blocks glucose adsorption in liver and intestinal cells by inhibiting sodium-linked glucose transporters, thereby ameliorating the effects of hyperglycemia [20–22]. However, the benefits of phloridzin in plants are largely unknown, although it is thought to be involved in pathogen resistance. When plants are attacked by a pathogen, phloridzin is hydrolyzed into phloretin by specific cytosolic glucosidases after cellular decompartmentalization. It is then oxidized by peroxidase and/or polyphenol oxidase into o-diphenols and toxic o-quinoids to inactivate microbial proteins [23,24]. Despite these preliminary reports, the role of phloridzin in pathogen defenses remains unclear because such accumulations have not yet been proven to be relevant to the resistance mechanism [9]. Some researchers have proposed that various metabolic pathways of DHCs, occurring in different apple genotypes, determine the potential for resistance to fire blight [25]. Petkovšek et al. have suggested that a consistent level of polyphenols is not necessary in the plant, but that their post-infection accumulation and further conversion are prerequisites for making plants more resistant to such threats [26]. Overexpression of *chalcone 3-hydroxylase* (CH3H) from *Cosmos sulphureus* stimulates the accumulation of 3-hydroxyphlorizin and reduces the susceptibility of *M. domestica* cv. Pinova to fire blight and scab [27]. Phloretin also displays broad-spectrum

**Abbreviations:** DHC, dihydrochalcone; GTs, glycosyltransferases; HPLC, high performance liquid chromatography; PSPG, plant secondary product glycosyltransferase; qRT-PCR, quantitative real-time PCR; RT-PCR, reverse-transcription PCR; UDP, uridine diphosphate; UGTs, UDP-dependent glycosyltransferases

\* Corresponding authors.

E-mail addresses: [gongxq0103@nwsuaf.edu.cn](mailto:gongxq0103@nwsuaf.edu.cn) (X. Gong), [fwm64@sina.com](mailto:fwm64@sina.com), [fwm64@nwsuaf.edu.cn](mailto:fwm64@nwsuaf.edu.cn) (F. Ma).

<http://dx.doi.org/10.1016/j.plantsci.2017.10.003>

Received 13 July 2017; Received in revised form 4 October 2017; Accepted 6 October 2017

Available online 10 October 2017

0168-9452/ © 2017 Elsevier B.V. All rights reserved.

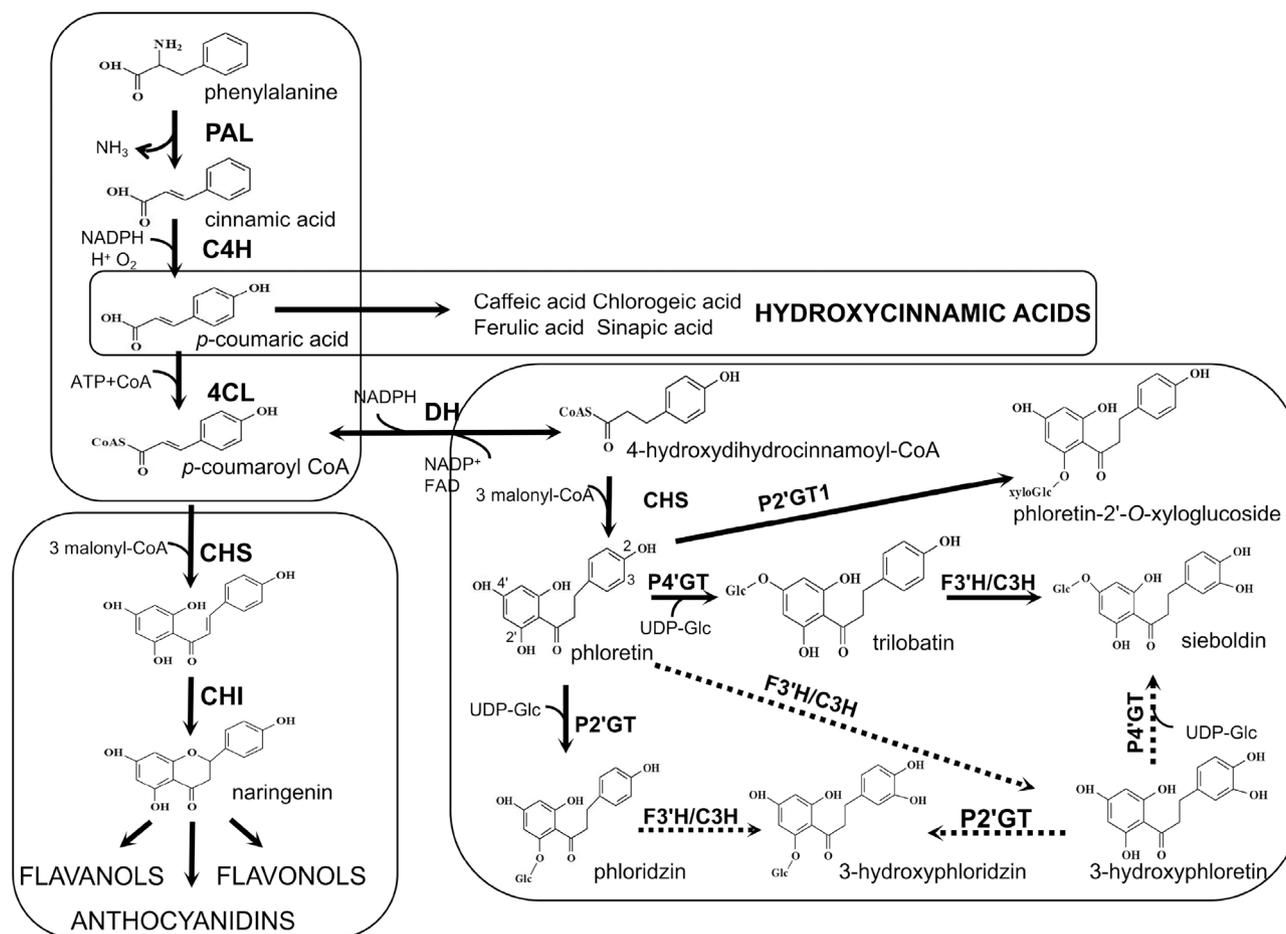


Fig. 1. Biosynthetic pathway of dihydrochalcones. PAL, phenylalanine ammonia lyase;

C4H, cinnamate-4-hydroxylase; 4CL, 4-hydroxycinnamoyl-CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; DH, dehydrogenase; P2'GT, UDP-glucose:phloretin 2'-O-glycosyltransferase; P2'GT1, UDP-xyloglucose:phloretin 2'-O-glycosyltransferase; P4'GT, UDP-glucose:phloretin 4'-O-glycosyltransferase; F3'H, flavonoid 3'-hydroxylase; C3H, chalcone-3-hydroxylase. Dotted arrow represents proposed pathway.

antibacterial activity and much stronger antioxidant activity [16,28], and is considered as an active compound against fungal infections [26].

Glycosylation, mediated by glycosyltransferases (GTs; EC 2.4.x.y), is ubiquitous in all living organisms [29]. According to CAZy (<http://www.cazy.org/GlycosylTransferases.html>), the GTs from various species can be classified into 94 families. The largest, Family 1, comprises diverse enzymes from animals, plants, bacteria, fungi, and viruses. In plants, Family-1 GTs, also called uridine diphosphate (UDP)-dependent glycosyltransferases, or UGTs, utilize UDP-activated sugars as the major donor molecule. One UGT feature is a conserved UGT-defining motif close to the C-terminus. The plant secondary product glycosyltransferase (PSPG) motif has 44 amino acid residues at the UDP-sugar binding site. A wide array of small molecules, including hormones, lipophilic acceptors, and secondary metabolites, can be glycosylated by UGTs, which improves their solubility in water and their chemical stability, reduces their toxicity, and changes their biological activity in plants [30,31]. Phloridzin is synthesized in three steps (Fig. 1), the final one being the glycosylation of phloretin [1,32,33]. Because this glycosylation has rarely been reported, it is essential that we identify the UDP-glucose:phloretin 2'-O-glycosyltransferase (P2'GT) in *Malus* plants if we are to conduct a more thorough investigation of the physiological relevance of phloridzin. To date, only four candidates have been isolated from apple leaves and tested for substrate specificity in vitro [1,32–34].

Apple is one of the most important crop trees because its popularly consumed fruit are available year-round in markets. Because those plants contain high levels of beneficial phloridzin, more research is

being focused on elucidating its mechanism of biosynthesis and its physiological relevance. In this study, we sampled the leaves of 64 *Malus* species and cultivars grown in the same location and measured their amounts of DHCs (i.e., phloretin, phloridzin, trilobatin, and sieboldin). Among them, eight genotypes exhibited a novel mode of phloridzin metabolism. Each of the tested samples featured a different phloridzin profile, which we used to determine whether phloridzin levels are correlated with the expression of UGTs across the entire genome. Protein activity as well as the transgenic plants analysis demonstrated that both MdUGT88F1 and its paralog MdUGT88F4 convert phloretin to phloridzin in apple. These findings will be beneficial to future examinations of the physiological role and mechanism of phloridzin biosynthesis.

## 2. Materials and methods

### 2.1. Plant materials, growing conditions, treatments, and chemicals

This study involved 64 *Malus* species and cultivars (Supplemental Table 1) plus one other cultivar, 'Royal Gala'. For the quantification analysis of DHCs, we collected mature leaves, branches, bark, and open flowers from healthy trees growing at the Horticultural Experimental Station and on the campus of Northwest A&F University, Yangling (34°20'N, 108°24'E), China. For each tissue type, at least five samples were harvested from individual trees at 9:00 AM, then immediately frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

The sources for chemicals were Yuanye (Shanghai, China) for

phloridzin and phloretin; Tauto Biotech (Shanghai, China) for phloretin 4'-O-glucoside (trilobatin); Extrasynthese (Genay, France) for 3-hydroxyphloretin 2'-O-glucoside (sieboldin); Sigma-Aldrich (Shanghai, China) for uridine 5'-diphosphoglucose disodium salt hydrate; and Tedia (Fairfield, OH, USA) for HPLC-grade methanol. All other chemicals were of analytical grade.

## 2.2. Quantification analysis of dihydrochalcones

Quantification of DHCs (including phloretin, phloridzin, trilobatin, and sieboldin) in *Malus* samples was performed as described by Zhang [35], with some modifications. Briefly, each frozen sample was ground to a fine powder in liquid nitrogen. Afterward, 250 mg of the powder was homogenized in 1.0 mL of extraction solution (70% methanol containing 2% formic acid), then blended in a thermomixer (1000 rpm, 30 °C, 30 min), followed by centrifugation (12,000 rpm, 10 min). Each supernatant was transferred to a new tube and, except for the fruit samples, the lower sediment was extracted twice more with 1.0 mL of extraction solution. The collected supernatants were then mixed and diluted 20–100 times with that solution before being passed through a 0.22- $\mu$ m syringe filter. For each sample, a 20- $\mu$ L aliquot was loaded into the Agilent 1200 HPLC system (Agilent Technology, Palo Alto, CA, USA), which was equipped with a diode array detector (DAD) and an Inertsil ODS-3 column (5.0  $\mu$ m particle size, 4.6 mm  $\times$  250 mm; GL Sciences Inc., Tokyo, Japan). The solvent system consisted of 50:50 (v:v) methanol (A) and water with 0.3% trifluoroacetic acid (B). Each sample was eluted for 40 min (flow rate, 1 mL min<sup>-1</sup>; 30 °C). The DAD was set at 280 nm to detect DHCs, and authentic samples for each were used for identification. They were then quantified by comparing the peak area against the standard curve (Table 1). Total DHCs were expressed as the sum of the mean values for the four DHCs mentioned above. All samples were analyzed three times.

## 2.3. Identification and phylogenetic analysis of UGT genes in apple

To identify apple UGTs, we used the PSPG domain, consisting of 44 amino acids, as our query sequence in the BLASTP program, searching against the Genome Database for Rosaceae (<https://www.rosaceae.org/node/1>) with an E-value cutoff of 0.01. The predicted amino acid sequences were searched in the NCBI database (<https://www.ncbi.nlm.nih.gov/>) to obtain accession numbers and descriptions of functions for these UGTs. Multiple alignments of MdUGTs were performed with amino acid sequences, using the Clustal X (version 2.1) program or DNAMAN software, and a phylogenetic analysis was conducted with the MEGA 5.0 program. The UGTs for *Arabidopsis thaliana* (hereafter *Arabidopsis*) and maize (*Zea mays*) were obtained from NCBI and the maize sequence database (<http://www.maizesequence.org/index.html>), respectively (Supplemental Table 2). We used the Neighbor-Joining (NJ) and *p*-distance methods with the pairwise deletion option to deal with gaps in the amino acid sequences, and set the bootstrap value to 1000 replicates. Our phylogenetic tree was visualized in the Interactive Tree Of Life program (iTOL, <http://itol.embl.de/>).

**Table 1**

Calibration curves and regression equations. <sup>1</sup>Y, value of peak area; X, concentration of reference compound ( $\mu$ g mL<sup>-1</sup>).

Compound	Calibration curve <sup>1</sup>	R <sup>2</sup>
Phloretin	Y = 668.044X - 50.240	0.997
Phloridzin	Y = 336.495X + 16.854	0.993
Trilobatin	Y = 386.284X - 11.976	0.995
Sieboldin	Y = 331.190X + 22.401	0.994

## 2.4. Gene isolation and qRT-PCR

Total RNA was extracted from different tissues using a Wolact Plant RNA Isolation Kit (Wolact, Hongkong, China). For reverse-transcription PCR (RT-PCR) and quantitative real-time PCR (qRT-PCR), first-strand cDNA was synthesized using a PrimeScript™ RT reagent Kit with the gDNA Eraser (Perfect Real Time) (Takara, Tokyo, Japan) according to the manufacturer's instructions. Genomic DNA was isolated from leaves with a Wolact Plant Genomic DNA purification Kit (Wolact).

To investigate gene expression, we performed qRT-PCR with a SYBR Premix Ex Taq Kit (Takara) on a StepOne Plus Real-time PCR Detection system (Applied Biosystems, Foster City, CA, USA). Each sample was analyzed in three biological replicates. The level of relative expression for each gene was calculated according to the 2<sup>- $\Delta\Delta$ CT</sup> method [36], and *Malus elongation factor 1 alpha (EF-1 $\alpha$ )* (DQ341381) was used as the reference gene [37]. The qRT-PCR primers are shown in Supplemental Table 3. The heatmap was conducted with HemI software [38].

The open reading frames of MDP0000219282, MDP0000461555, MDP00002888715, MDP0000170162, MDP0000375160, MDP0000361449, and MDP0000318101 were amplified with primers (Supplemental Table 3), using PrimeSTAR HS DNA polymerase (Takara). The PCR products were cloned into PMD 19-T Simple vectors (Takara) for sequencing (Sango, Shanghai, China).

## 2.5. Assays of enzyme activity

The full-length cDNA fragment of *MdUGT1* was PCR-amplified with primers containing BamHI or HindIII restriction sites (Supplemental Table 3) and inserted into a pET-32 a (+) vector (Novagen, Darmstadt, Germany) that was pre-digested by the same restriction enzymes. Other fusion vectors were constructed as described above. Recombinant pET-32 a (+) vectors harboring different genes were independently transformed into *Escherichia coli* strain BL21(DE3). Afterward, 100  $\mu$ L of a positive clone strain that was cultured overnight was inoculated into 5 mL of an LB medium containing 50 mg L<sup>-1</sup> ampicillin and then incubated at 37 °C until the OD600 reached 0.6–1.0. Following the addition of 0.8 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG, Takara), the culture was grown at 20 °C and shaken at 120 rpm for 36 h before a 1-mL aliquot was centrifuged (12,000 rpm, 2 min). The cell pellet was suspended in buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 30 mM NaCl, and 20 mM imidazole), then incubated at 37 °C for 30 min and frozen in liquid nitrogen after the addition of 1 mg L<sup>-1</sup> (final concentration) of lysozyme. The samples were centrifuged (12,000 rpm, 10 min, 4 °C) and the supernatant fractions were used for assaying enzyme activity.

The activity of UGT was monitored using 500- $\mu$ L reaction mixtures that contained 100  $\mu$ L of protein preparation, 1  $\mu$ L of phloretin (50 mM), 1  $\mu$ L of UDP-glucose (50 mM), 250  $\mu$ L of Tris buffer (100 mM; pH 7.5), 1  $\mu$ L of dithiothreitol (DTT, 1 M), and 147  $\mu$ L of deionized water. The assay mixtures were incubated for 30 min at 30 °C and the reactions were terminated by adding 40  $\mu$ L of acetic acid. After centrifugation (12,000 rpm, 10 min, 4 °C), the supernatants were passed through a 0.22- $\mu$ m syringe filter. Finally, 20- $\mu$ L aliquots were directly subjected to HPLC analysis on the Agilent 1200 HPLC system, as described above, with the pET-32 a (+) vector control being run in parallel with the enzyme reaction.

## 2.6. Production of silencing 'GL-3' apple plants and HPLC analysis

A 238-bp fragment of *MdUGT88F1* was PCR-amplified with 282F and 282R (Supplemental Table 3) and cloned into the Gateway transfer vector pDONR222 and then into the binary vector pK7WIWG2D, or directly cloned into pHellsgate 2 to create the two RNAi silencing constructs. All Gateway reactions were performed as recommended by the manufacturer (Invitrogen). Afterward, the binary vectors pHellsgate 2 and pK7WIWG2D were electroporated into *Agrobacterium tumefaciens* strain EHA105.

Leaves of *Malus* genotype ‘GL-3’ (‘Royal Gala’) were used for *A. tumefaciens*-mediated transformation and selected on kanamycin-containing media, according to the method of Dai [39]. Kanamycin-resistant lines were further tested for the presence of the transgene using genomic-PCR. Transgenic plants and ‘GL-3’ controls were grown on a subculture medium [Murashige and Skoog (MS) medium supplemented with 1.33  $\mu\text{M}$  6-benzyladenine and 1.14  $\mu\text{M}$  indoleacetic acid]. The expanding leaves of 35-day-old plants were collected and subjected to qRT-PCR analysis and analysis of dihydrochalcones as described previously.

## 2.7. Statistical analysis

We used SPSS software (version 17.0) for the statistical analysis. All data were subjected to one-way ANOVA.

## 3. Results

### 3.1. Identification and phylogenetic analysis of UGTs in *Malus* genotypes

The complete sequencing of the domesticated apple (*Malus × domestica* Borkh.) genome [40] has greatly facilitated the identification of apple gene families. We screened that genome using the PSPG domain sequence and found 299 candidate UGTs. After removing genes encoding the same proteins and sequences that were too short or long, we selected for further analysis 237 that encode proteins ranging from 208 to 932 amino acids and that have the conserved PSPG domain (Supplemental Table 4). Gene annotation based on the apple genome sequence database and NCBI was used to determine GenBank accession numbers, protein lengths, numbers of exons, chromosome locations, and functional prediction/descriptions (Supplemental Table 4). These included some characterized UGTs involved in glycosylation of phloretin in vitro, i.e., MDP0000219282 (MdPGT1, Accession no. EU246349; UGT88F1, ACZ44840; MdP2’GT, KT444675), MDP0000163017 (UGT71K1, ACZ44835), MDP0000215525 (UGTA15, AAZ80472), and MDP0000617956 (MdPh-4’-OGT, AY786997). All have been characterized as harboring P2’GT activity while MdPh-4’-OGT and UGTA15 also catalyze phloretin into trilobation in vitro [7,9,32–34].

The UGTs in *Linum usitatissimum* cluster into 14 major groups (A-N) based on the protein sequences of typical *Arabidopsis* UGTs in each group [41]. Genome-wide analysis of Family-1 UGTs from 12 land plants, including *M. domestica*, have revealed 16 groups, including 14 that are conserved (A-N), plus new groups called O and P [42,43]. To classify all 237 of our apple UGTs, we constructed an unrooted tree by aligning their full-length amino acid sequences with 19 *Arabidopsis* UGTs (from Groups A-N) [42] and three maize UGTs (from Groups O-Q) (Supplemental Table 2) [31]. This created 16 major groups, A through P (Fig. 2), as well as three members in Group O, including two zeatin O-glucosyltransferase (ZOG)-like genes, MDP0000850255 and MDP0000793268, that are always found in that group [43]. Another seven apple UGTs were clustered with a Group-P maize UGT. We thought it interesting that all of these genes, except for MDP0000299496, belong to UGT85A. Group Q, novel to maize, was not found in apple UGTs.

### 3.2. Investigation of dihydrochalcone accumulations

We utilized the rapid, convenient HPLC system to determine DHC concentrations in *Malus* leaves and detected primarily phloretin, phloridzin, trilobatin, and sieboldin in amounts that varied by genotype. The aglycone phloretin occurred at low concentrations (from an undetermined amount to 2,077.85  $\mu\text{g g}^{-1}$  fresh weight, or FW), therefore contributing very little to the total DHCs. Among the three phloretin glycosides, phloridzin was predominant, accumulating in all species/cultivar samples at concentrations ranging from 3.56 to 161,966.20  $\mu\text{g g}^{-1}$  FW, and

followed a Gaussian distribution (data not shown). Extremely high concentrations of phloridzin ( $> 150,000 \mu\text{g g}^{-1}$  FW) were calculated for ZA5 (*M. honanensis* Rehd.) and ZA14 (*M. kansuensis* (Batal.) Schneid.), followed by ZL9 (*M. domestica* ‘Oekonomierat Echtermeyer’), ZH18 (*M. bhutanica*), ZK23 (*M. hybrida* ‘Irene’), ZB10 (*M. rockii* Rehd.), ZB3 (*M. hupehensis* Rehd.), and ZI13 (unknown genotype), each containing  $> 50,000 \mu\text{g g}^{-1}$  FW. In contrast, very low levels of phloridzin ( $< 100 \mu\text{g g}^{-1}$  FW) were found in ZA0 (*M. spp.* ‘Adams’ (North American Begonia)), ZB15 (*M. sieboldii* (Reg.) Rehd.), ZD1 (*M. hupehensis* Rehd.), ZD7 (*M. toringoides* (Rehd.) Hughes), ZD9 (*M. spectabilis* Borkh), ZE8 (*M. zumi*), ZH16 (*M. sargentii*), and ZK2 (*M. hybrida* ‘Robinson’) (Supplemental Table 1). Those last eight genotypes, along with 28 others, accumulated both trilobatin and sieboldin (Supplemental Table 1). No plant lines contained only one or the other of those two DHCs. Among the 36 showing these simultaneous accumulations, they also displayed a wide range in phloridzin concentrations (3.56–29,498.54  $\mu\text{g g}^{-1}$  FW), but only moderate variations in total DHCs (12,797.51–99,369.24  $\mu\text{g g}^{-1}$  FW) (Supplemental Table 1).

Measurements were also made with samples of other tissue types from ZA0, ZB15, ZD1, ZD7, ZD9, ZE8, ZH16, ZK2, ZA5, and ZA14. Levels of total DHCs were relatively lower in the fruit and relatively higher in the branches (Supplemental Table 6). The same trend was noted for phloridzin, trilobatin, and sieboldin, but no such pattern was found for phloretin. The leaves had relatively higher total DHCs, although the amounts of individual compounds were rather variable among genotypes. For ZA0, ZB15, ZD1, ZD7, ZD9, ZE8, ZH16, and ZK2, phloridzin was replaced by trilobatin and sieboldin. Similar results were found in the branches, which accumulated more phloridzin but less of the other two. The greatest variability in phloridzin concentrations was detected in samples from ‘Adams’. In that genotype, those levels were 10,825.45  $\mu\text{g g}^{-1}$  FW in the bark and 5,267.84  $\mu\text{g g}^{-1}$  FW in the branches, both of which were much higher than the amounts measured in the leaves (83.21  $\mu\text{g g}^{-1}$  FW) and flowers (233.24  $\mu\text{g g}^{-1}$  FW). Total DHCs were most abundant in the leaves (51,819.73  $\mu\text{g g}^{-1}$  FW) rather than in the bark (36,494.0371  $\mu\text{g g}^{-1}$  FW) (Fig. 3A). The corresponding tissues from ‘Royal Gala’ also contained relatively higher levels of phloridzin, ranging from 5,308.76  $\mu\text{g g}^{-1}$  FW in the flowers to 20,651.76  $\mu\text{g g}^{-1}$  FW in the leaves and 23,365.63  $\mu\text{g g}^{-1}$  FW in the bark. Although the distribution of phloridzin differed among tissue types in those two genotypes, it was the predominant DHC in their bark. In ‘Royal Gala’, the pattern of accumulation was similar for total DHCs, being highest in the bark and somewhat lower in the leaves (Fig. 4A).

### 3.3. Identification of UDP-glucose: phloretin 2’-O-glucosyltransferase (P2’GT)

The concentration of phloridzin was extremely low in ‘Adams’ leaves, where trilobatin and sieboldin were found instead (Fig. 3A). In fact, those latter two were the most highly accumulated of the DHCs in all tissues sampled from that cultivar (Fig. 3A). DHC concentration range was also rather narrow when compared with other genotypes, suggesting that low P2’GT activity was the key factor, and not two upstream enzymes, in the biosynthetic pathway. We believe this is why the level of phloridzin was extremely low in ‘Adams’ leaves.

Using samples from ‘Adams’ we selected 37 UGTs from each sub-group to compare their spatial expression patterns in corresponding tissues. The qRT-PCR assays showed that all of these genes were differentially expressed in all tested tissues, but most were preferentially expressed in the leaves (Fig. 5). The exceptions were MDP0000189941, MDP0000163017, MDP0000213638, MDP0000128265, MDP0000850255, MDP0000289413, MDP0000120322, MDP0000209283, and MDP0000154932, which were preferentially expressed in the flowers (and MDP0000286574 and MDP0000134856 being exclusively expressed there) (Fig. 5). Only a few genes were highly expressed in the bark and branches. They included MDP0000293001, MDP0000209283, and MDP0000135820—expressed in only one tissue type — and MDP0000219282 and MDP0000839414, showing high transcript levels in both (Fig. 5). The expression patterns of

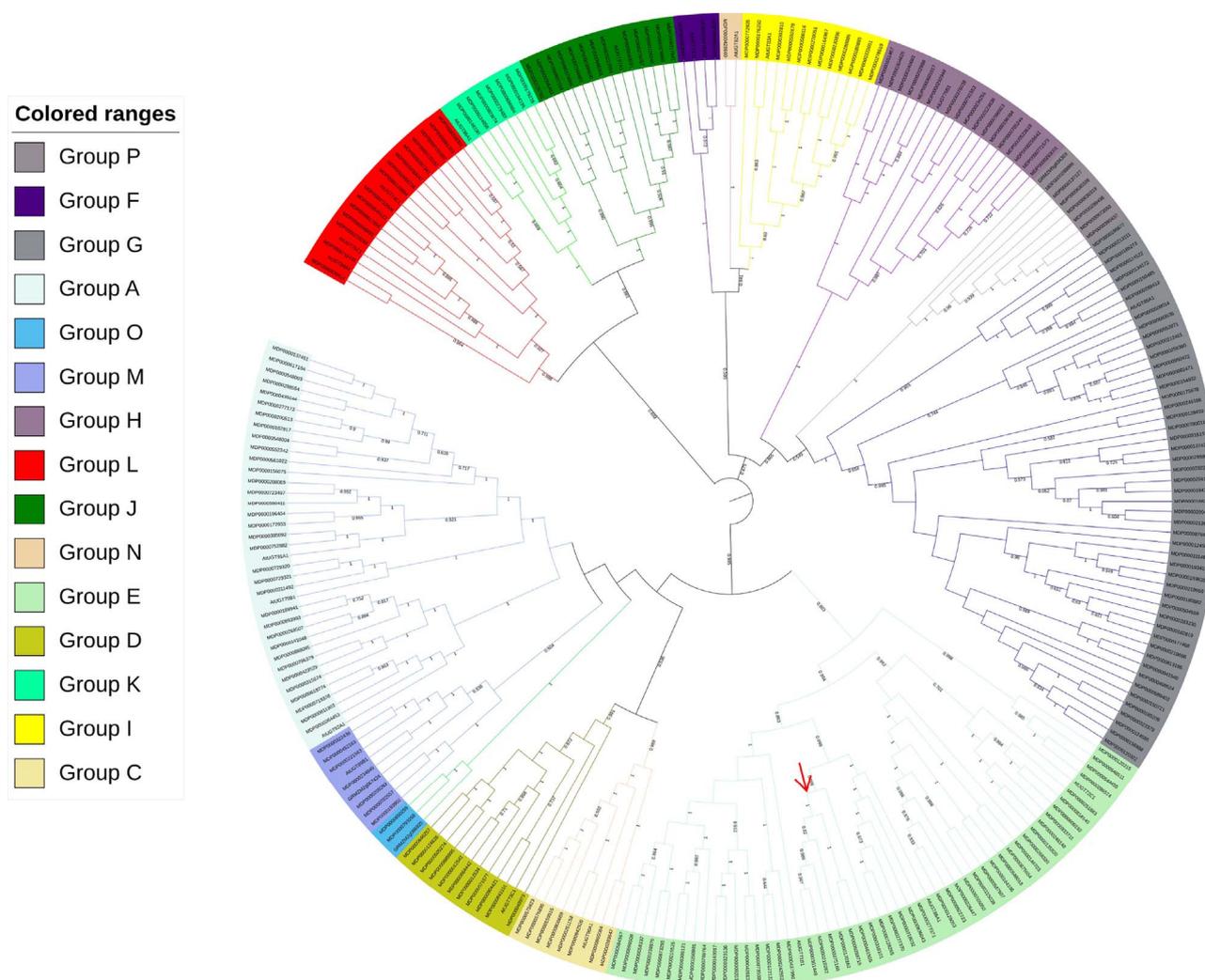


Fig. 2. Phylogenetic analysis of UGTs from apple. The NJ tree was created with MEGA 5.0 program (bootstrap value set at 1000), using full-length sequences of 237 apple UGTs, 18 *Arabidopsis*, and 3 maize UGT proteins. Bootstrap values > 50% are shown above nodes. Sub-group of UGT88F subfamily is indicated by arrow.

MDP0000839414 (Fig. 3B-a) and MDP0000219282 (Fig. 3B-b) were well-correlated with their phloridzin profiles (Fig. 3A).

A comparable trend in expression was observed for MDP0000219282 and MDP0000839414 in ‘Royal Gala’, but different profiles for phloridzin were shared between ‘Royal Gala’ and ‘Adams’ (Figs. 3 and 4). Using leaves from 41 species and cultivars that strongly diverged in their phloridzin levels (eight separate groups based on classification of phloridzin and DHCs concentrations, Supplemental Tables 7 and 8), we monitored the expression of MDP0000219282 and MDP0000839414 (Fig. 6A). The average expression level of MDP0000219282 in each of the eight groups was positively correlated with their mean phloridzin concentrations (Fig. 6B-b). There, the correlation coefficient ( $R^2$ ) was 0.983 and an extremely wide range in transcript abundance was noted (Fig. 6B-f). Expression was much less for individual genotype members in Group 1, a phenomenon that closely reflected their phloridzin accumulations. Although the genotypes within Groups 1 and 2 contained both trilobatin and sieboldin, the amount of total DHCs was much higher in the former group (Fig. 6, Supplemental Table 5). In contrast, no significant correlation was determined between expression of MDP0000839414 and phloridzin accumulations (Fig. 6B-a), and its expression was much higher in some genotypes within Group 1 than in members of the other groups, including Group 8 (Supplemental Table 5). A similar relationship was found between expression of MdCHS (chalcone synthase) genes and total DHCs (Fig. 6B-e, Supplemental Table 5). Therefore, all these results indicated that the encoded glycosyltransferase of MDP0000219282 has an

important role in the conversion of phloretin to phloridzin.

#### 3.4. Further identification of UGT88F subfamily members

Six other UGTs – MDP0000318101, MDP0000461555, MDP00002888715, MDP0000170162, MDP0000375160 and MDP0000361449–were closely clustered with MDP0000219282 (Fig. 2). To determine whether the activity of P2’GT was part of this sub-group, we used RT-PCR and obtained the CDS from ‘Royal Gala’ for all of them except MDP0000318101, MDP00002888715, and MDP0000361449. Originally isolated from ‘Rebella’, *M. pumila*, and ‘Golden Delicious’ apple genotypes, MDP0000219282 was referred to as *UGT88F1*, *MdPGT1*, and *MdP2’GT*, respectively. Therefore, this subfamily was named after UGT88F according to the system of the UGT Nomenclature Committee because those four UGTs shared > 78% amino acid sequence identity (Table 2) [44,45]. These were submitted to the GenBank database with the following accession numbers: MDP0000219282 (MdUGT88F1, KX639791), MDP0000375160 (MdUGT88F4, KX639792), MDP0000461555 (MdUGT88F6, KX639793), and MDP0000170162 (MdUGT88F8, KX639794).

Further qRT-PCR analysis showed that the expression patterns of MDP0000461555 and MDP0000219282 matched with the phloridzin profiles in ‘Adams’ (Fig. 3B). For MDP0000461555, the pattern was not totally consistent, with expression being preferential in the branches and bark and relatively low in the leaves and flowers (Fig. 3B-e).

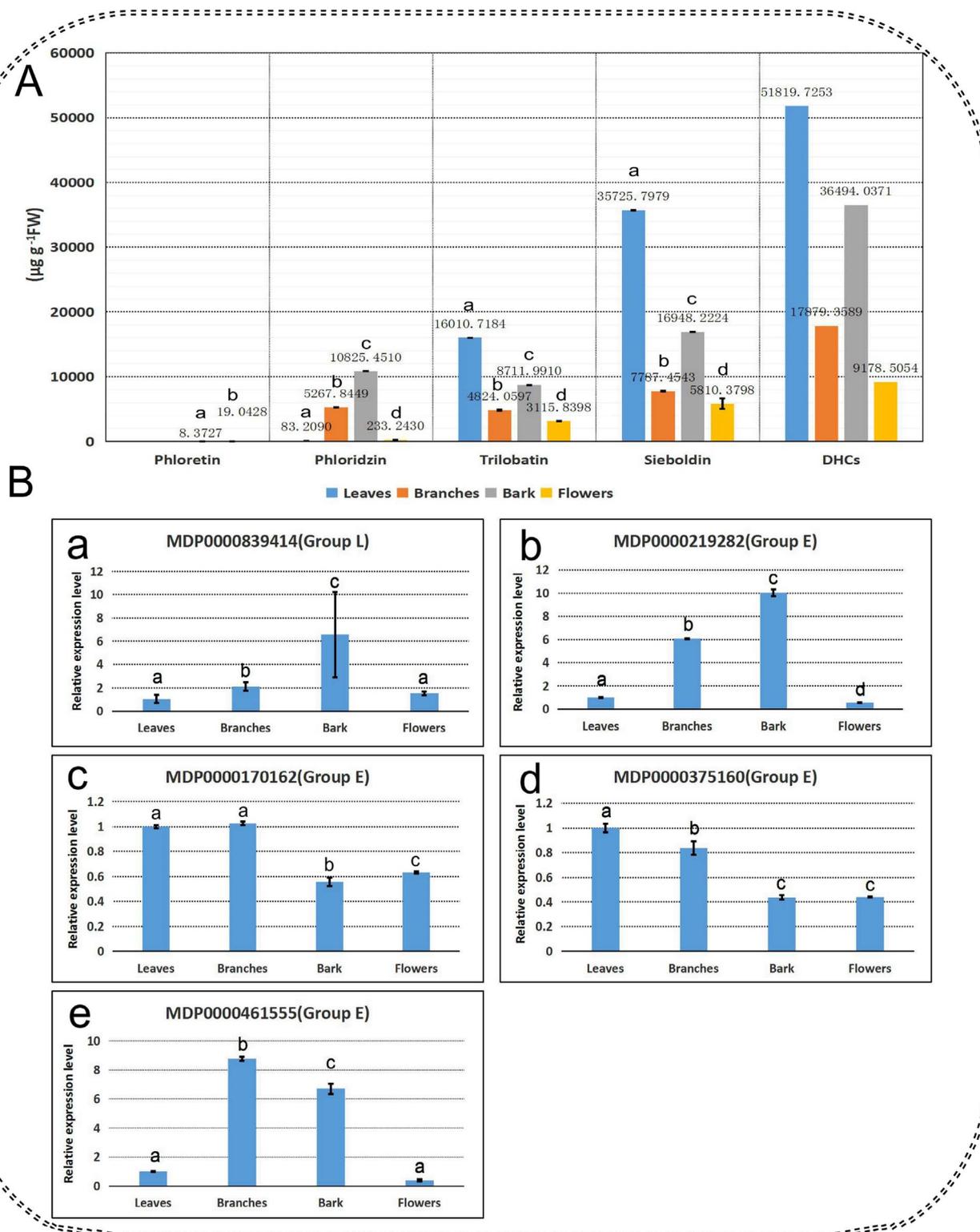


Fig. 3. Spatial profiles of dihydrochalcones (DHCs) and UGTs in *Malus* spp. cv. Adams. (A) Concentrations of phloretin, phloridzin, trilobatin, sieboldin, and total DHCs in leaves, branches, bark, and flowers. (B) Expression patterns of UGTs in corresponding tissues. Data are means  $\pm$  SD. Values not followed by same letter are significantly different ( $P < 0.05$ ).

However, a comparable expression pattern was also detected in ‘Royal Gala’ (Fig. 4B-e). For further verification, we investigated its expression in leaves from 41 genotypes. Although the mean level of expression was significantly correlated with phloridzin concentration, we found no significant difference in expression among our various groups (Fig. 6B-c). In addition, expression profiles were quite similar for MDP0000375160 and MDP0000170162 in ‘Adams’, with both showing

relatively high expression in leaves and branches (Fig. 3B). A similar pattern was shared between MDP0000375160 and MDP0000170162 in ‘Royal Gala’, although transcript accumulations were slightly higher in the branches and bark than in other tissue types (Fig. 6B).

Our RT-PCR results indicated no expression of MDP0000219282, MDP0000375160, and MDP0000461555 in ‘Adams’ leaves. However, all of them, as well as MDP0000170162, were detected in the leaves of

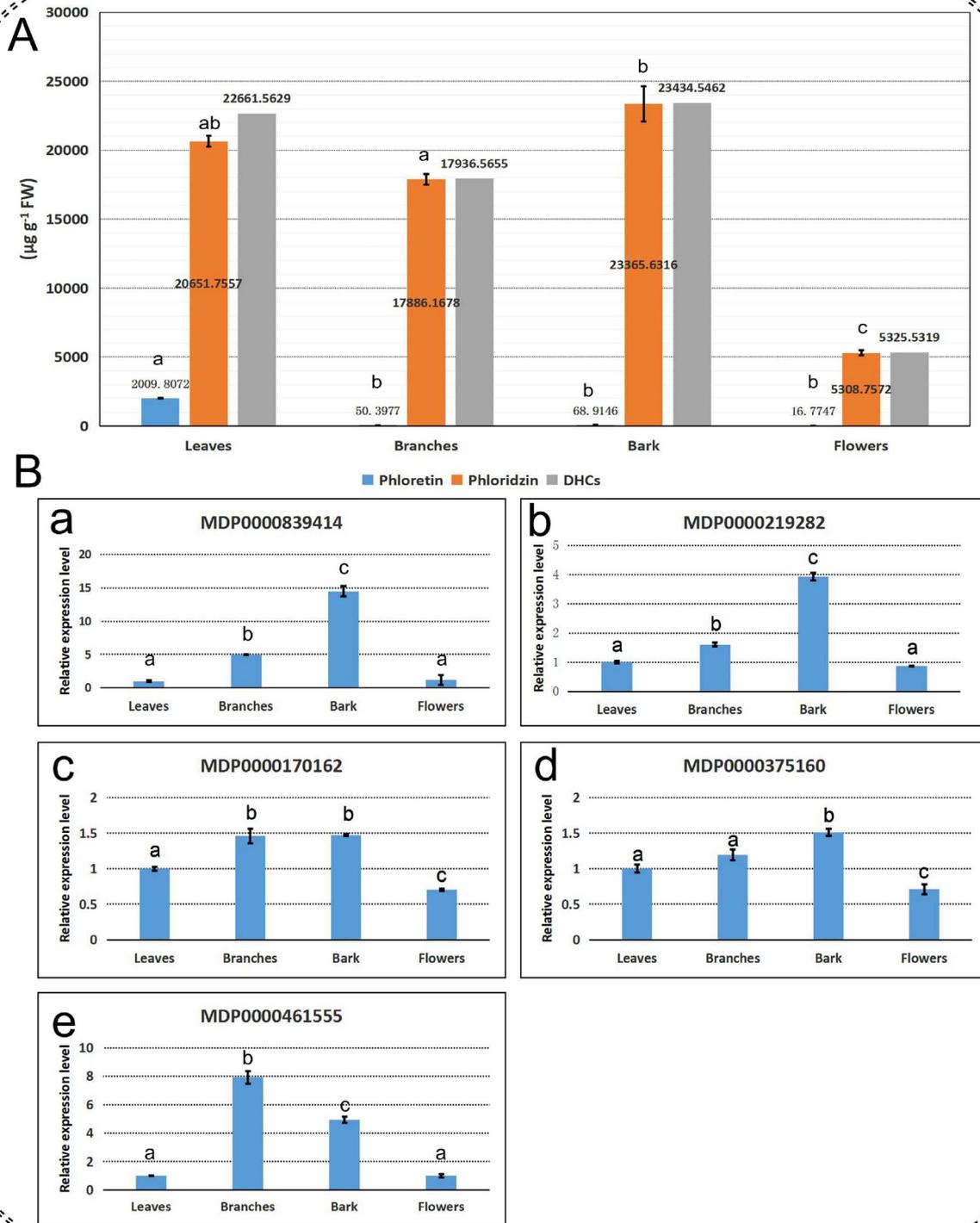


Fig. 4. Spatial profiles of dihydrochalcones (DHCs) and selected UGTs in *Malus × domestica* cv. Royal Gala. (A) Concentrations of phloretin, phloridzin, and total DHCs in leaves, branches, bark, and flowers. (B) Expression patterns in corresponding tissues. Data are means ± SD. Values not followed by same letter are significantly different ( $P < 0.05$ ).

‘Royal Gala’ (data not shown). Further qRT-PCR analysis indicated that MDP0000219282 transcripts were only abundant in the branches of ‘Adams’ (Fig. 7).

### 3.5. Recombinant protein assays and sequence analysis of UGT88F subfamily

To confirm that the true P2’GTs are defined in the UGT88F

subfamily, we expressed the recombinant proteins ectopically in a transformed *E. coli* strain from which extracted total proteins were independently incubated with UDP-glucoses and phloretin. Our HPLC analysis of the supernatants showed that phloridzin was present in the recombinant proteins for MdUGT88F1 and MdUGT88F4 but not for MdUGT88F6 and MdUGT88F8. Furthermore, trilobatin was not detected in any of those systems (Fig. 8). The average expression level of *MdUGT88F4* in each group was positively correlated ( $R^2 = 0.79$ ) with

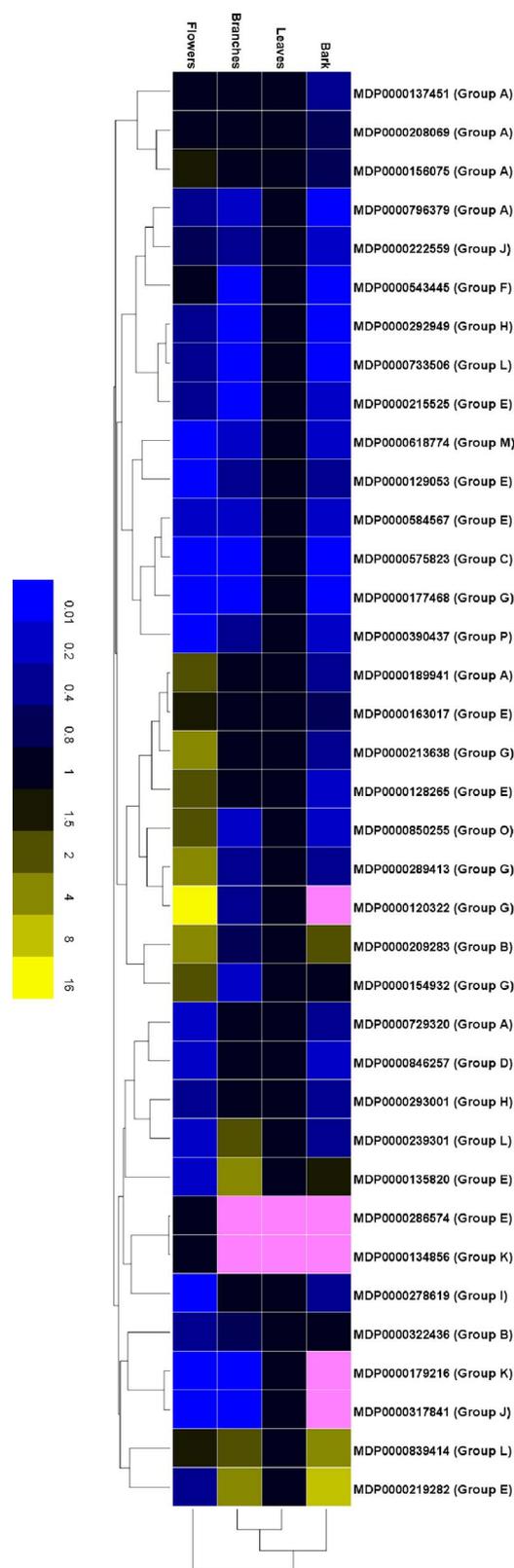


Fig. 5. Spatial profiles of *UGTs* in different tissues from *Malus* spp. cv. Adams. Except for MDP0000286574 and MDP0000134856, expression level is set to '1' for each gene in leaves. Bar indicates relative expression level. Pink block means no detected expression.

their mean phloridzin concentration (Fig. 6).

MdUGT88F1 and MdUGT88F4 shared 91.72% identity in their amino acid sequences, and had entirely identical PSPG motifs. Relatively low similarities were found between MdUGT88F1 and

MdUGT88F6/MdUGT88F8 as well as between the latter two genes and MdUGT88F4. However, sequence identities were relatively higher between MdUGT88F4 and MdUGT88F6/MdUGT88F8 than when those latter two were compared with MdUGT88F1. Because identity was slightly higher between MdUGT88F6 and MdUGT88F8 (Table 1, Fig. 9), we propose that MdUGT88F1 and MdUGT88F4 arose from one duplication, while MdUGT88F6 and MdUGT88F8 came from another duplication, which led to their diversity in functions. We found it interesting that two residues in MdUGT88F1 – K<sup>160</sup> and T<sup>161</sup> – were not present in MdUGT88F4, MdUGT88F6, or MdUGT88F8 (Fig. 9). The orthologs of MdUGT88F1 and MdUGT88F4 in 'Adams' were cloned by RT-PCR and genomic-PCR, respectively. Multiple alignments based on their amino acid sequences demonstrated that they shared 98.55% and 95.43% sequence identity with MaUGT88F1 (Accession No. KX639795) and MaUGT88F4 (KX639796), respectively, from 'Adams'. The presence of K<sup>160</sup>T<sup>161</sup> is a signature that can be used to distinguish those two paralogs in *Malus* species and cultivars (Fig. 9). Based on these results, we could conclude that MdUGT88F1 and MdUGT88F4 are involved in phloridzin production.

### 3.6. Analysis of dihydrochalcone and qRT-PCR expression in transgenic plants

To confirm further that the true P2'GTs are defined in MdUGT88F1 and MdUGT88F4, we designed two vectors containing a 238-bp fragment to specially targeted these two genes by RNAi in 'GL-3' plants. From this, we obtained one pHellsgate 2 (Ri-3) and two pK7WIWG2D (Ri1-1 and Ri3-3) transgenic plants from independent transformation experiments after genomic-PCR identification with primers 282F/R crossing with primers L1/2 (L1 and L2 were designed based on the sequence of vectors) (Fig. 10A). Those three transgenic plants initially grew normally in tissue culture, but differed slightly in phenotype two months later. These phenotypes included stunting and reduced internode length (Ri1-1 and Ri3-3) and reduced leaf size on top of the plants (Ri-3). The expanding leaves of the three transgenic plants and 'GL-3' controls were collected after growing for 35 d in a subculture medium. Two major dihydrochalcones (i.e., phloridzin and phloretin) were qualified by HPLC analysis. The phloridzin levels in transgenic plants were 50.94% to 66.52% of that in the controls (Fig. 10A). As we had expected, the phloretin was highly accumulated in the knockdown lines, with levels being up to 3.71-fold greater than in the controls (Fig. 11B). Therefore, the ratio of phloridzin to phloretin was significantly decreased in all transgenic plants (Fig. 11D). Although the dihydrochalcone trilobatin was not detected in the 'GL-3' plants, it did accumulate in Line Ri-3 (47.36  $\mu\text{g g}^{-1}$  FW), but not in Ri1-1 or Ri3-3 (Fig. 11C). This indicated that low P2'GT activity altered the flux of phloretin substrate in those transgenics. Further qRT-PCR demonstrated that the expression of genes in the *MdUGT88F* family differed in those three lines. For example, expression of *MdUGT88F1* and *MdUGT88F4* ranged between 16.01% and 69.44% and between 36.98% to 133.73%, respectively. Though not expected, the expression of *MdUGT88F6* and *MdUGT88F8* was also significantly interrupted at the transcriptional level (Fig. 10B).

## 4. Discussion

2 Although large quantities of phloridzin are accumulated in apple trees, its biosynthetic pathway has only recently been elucidated. Those steps include 1) NADPH-dependent formation of 4-dihydrocoumaroyl-CoA from *p*-coumaroyl-CoA by dehydrogenase; 2) formation of phloretin from 4-dihydrocoumaroyl-CoA and malonyl-CoA by the common CHS [46]; and 3) glycosylation of phloretin to phloridzin, as mediated by P2'GT [1,10]. The first step successfully competes with common flavonoid formation for the same hydroxycinnamic acid intermediate. Although some potential genes have been documented, identity of the dehydrogenase gene remains unclear [10,46].

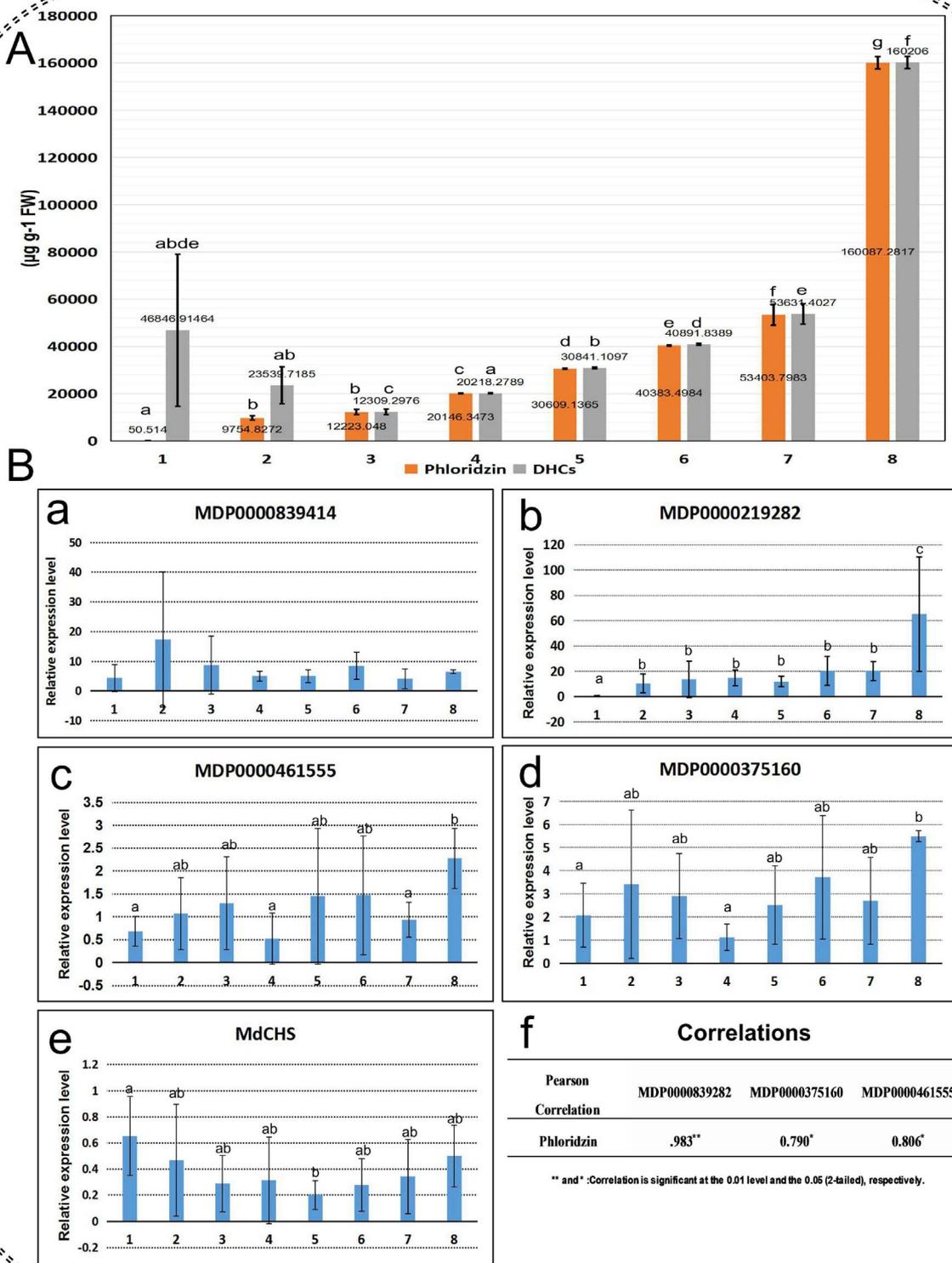


Fig. 6. Levels of dihydrochalcones (DHCs) and expression in 8 groups of genotypes showing divergence in phloridzin accumulations. (A) Concentrations of phloridzin and total DHCs. (B) Expression of MDP0000839414 (a), MDP0000219282 (b), MDP0000461555 (c), MDP0000375160 (d), and MdCHS (e) in Groups 1 through 8, and correlations with phloridzin concentrations (f). Data are means ± SD. Values not followed by same letter are significantly different (P < 0.05).

3 The second step appears to be non-specific because it shares the same CHS for chalcone production. Silencing of *CHS* leads to the loss of key flavonols, e.g., naringenin chalcone and phloridzin, from the leaves and fruit of many plants, including apple [47]. Therefore, *CHS* genes are not suitable candidates for studying the physiological relevance of

phloridzin. For the third step, compelling evidence suggests that P2/GT is a key enzyme catalyzing the rate-limiting step in the phloridzin biosynthetic pathway [10,48,49]. Knowledge about this glycosylation, however, is limited. In *Malus* species, six glycosyltransferases – MdPGT1, UGT88F1, MdP2/GT, UGT71K1, UGT A15, and MdPh-4'-OGT

**Table 2**

Percent identities shared among amino acid sequences in members of UGT88F subfamily. Identities were determined with DNAMAN software.

	MdUGT88F1	MdUGT88F4	MdUGT88F6	MdUGT88F8
MdUGT88F1	100%	91.72%	79.92%	82.4%
MdUGT88F4		100%	82.74%	85.46%
MdUGT88F6			100%	86.9%
MdUGT88F8				100%

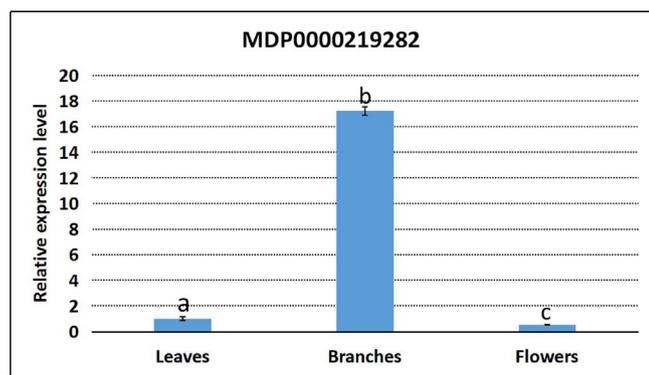


Fig. 7. qRT-PCR analysis of MDP 0000219282 in leaves, branches, and flowers from *Malus* spp. cv. Adams. Data are means  $\pm$  SD. Values not followed by same letter are significantly different ( $P < 0.05$ ).

– have roles in converting phloretin into phloridzin [7,9,32,33,40].

4 In-vitro P2'GT activity has also been reported for three glycosyltransferases from *Pyrus communis* (UGT88F2, ACZ44838; UGT71K2, ACZ44837; and UGT16, ACZ44836) [9]; as well as for one glycosyltransferase from *Dianthus caryophyllus* (DicGT4, AB191248.1) [50] and one bacterial glycosyltransferase (YijC, AAU40842) [51]. However, only four of them – MdPGT1, UGT88F1, MdP2'GT, and UGT88F2 – can selectively act on glycosylate phloretin at the 2'-OH position to

produce analytical amounts of phloridzin. In addition, it is still not clear whether these P2'GTs reported from *Malus* have functionality in apple plants. Because no strong correlation between expression levels and phloridzin concentrations has been published, we conducted our genome-wide identification of P2'GTs using *Malus* genotypes that diverged widely in their accumulations of phloridzin.

The composition of dihydrochalcones differs among genotypes of *Malus* [9,25,28]. This contrast among species is mainly caused by different genetic constitutions [52]. We noted that leaf concentrations of phloridzin, the predominant DHC in most genotypes, varied widely, ranging from 10,195.68  $\mu\text{g g}^{-1}$  FW (ZK13, *M. hybrida* 'Red Splendor') to 161,956.20  $\mu\text{g g}^{-1}$  FW (*M. kansuensis* (Batal.) Schneid. ZA14). The proportion of phloridzin in leaves has previously been reported to account for up to 18% of the total dry weight in that tissue [1,2]. Here, the amount of phloridzin made up approximately 40.5% of the total dry weight in our samples when the water content in mature leaves was almost 60% (data not shown). Among the 10 types of DHCs already described from *Malus* species, only trilobatin and sieboldin occur in the largest quantities, along with relatively high amounts of phloretin and phloridzin [10]. Our investigation revealed that 36 genotypes simultaneously accumulated trilobatin and sieboldin as well as phloridzin, a phenomenon that seems to conflict with previous reports [10] that have never mentioned the presence of all three within the same species. Most members of the sieboldiana group contained both sieboldin and phloridzin. That group includes five cultivars (ZA11, ZA12, ZA13, ZB15, and ZD13) of *M. sieboldii* (Reg.) Rehd. and one (ZA3) of *M. floribunda* Siebold here. The exceptions were ZA12 and ZA13, neither of which had sieboldin. In *Malus trilobata*, phloridzin was replaced completely by trilobatin. The latter occurred with phloridzin only in some progeny of *M. × domestica* crossed with *M. trilobata*. Even so, the three DHCs were also reported together in *M. × domestica* cv. Evereste [53,54]. We noted with interest that eight genotypes each contained less than 100  $\mu\text{g g}^{-1}$  FW phloridzin. In addition to the replacement of phloridzin with varying amounts of trilobatin and sieboldin, the latter is thought to be converted mainly from the former, rather than from 3-hydroxyphloretin. This proposal is based on the lack of large quantities of

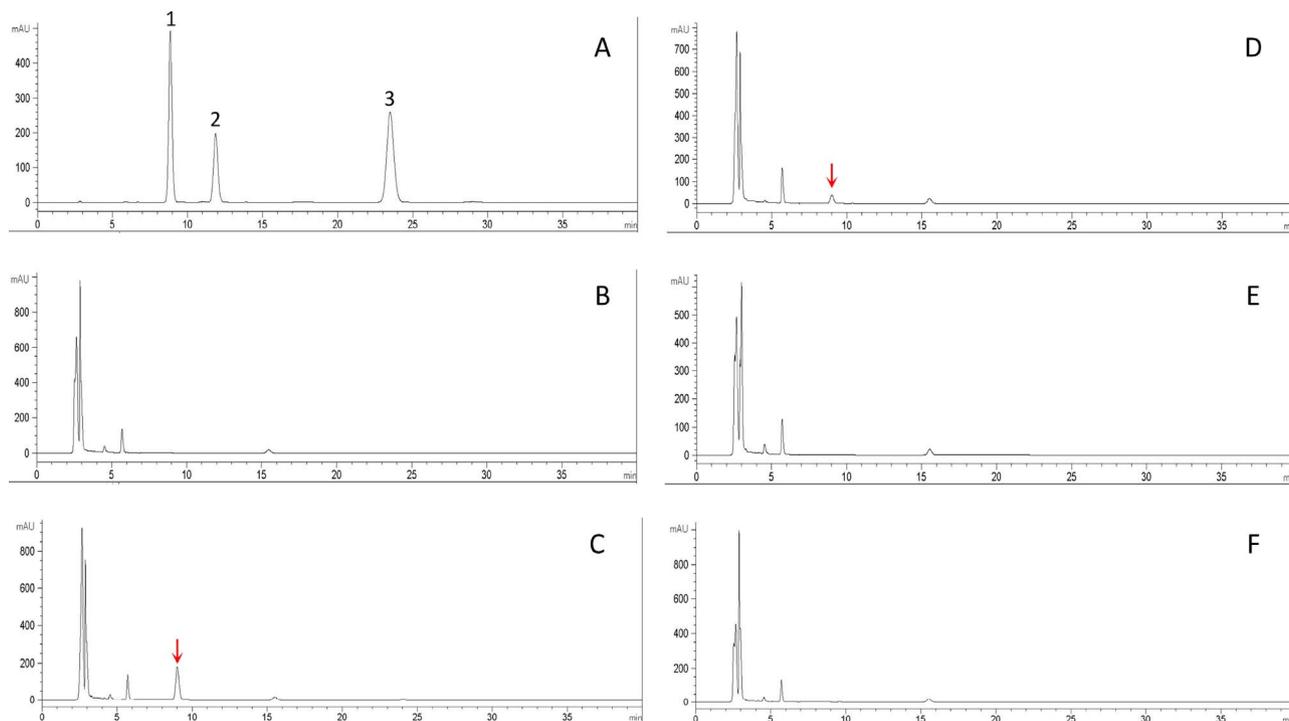


Fig. 8. Detection via HPLC of phloridzin in different supernatants from assays of recombinant protein activity. A, HPLC spectrum of phloridzin (1), trilobatin (2), and phloretin (3) authentic sample. B-F, HPLC spectrum of pET-32a (+), MDP0000219282, MDP0000375160, MDP0000461555, and MDP0000170162 recombinant protein assays, respectively. Arrows indicate phloridzin peaks.

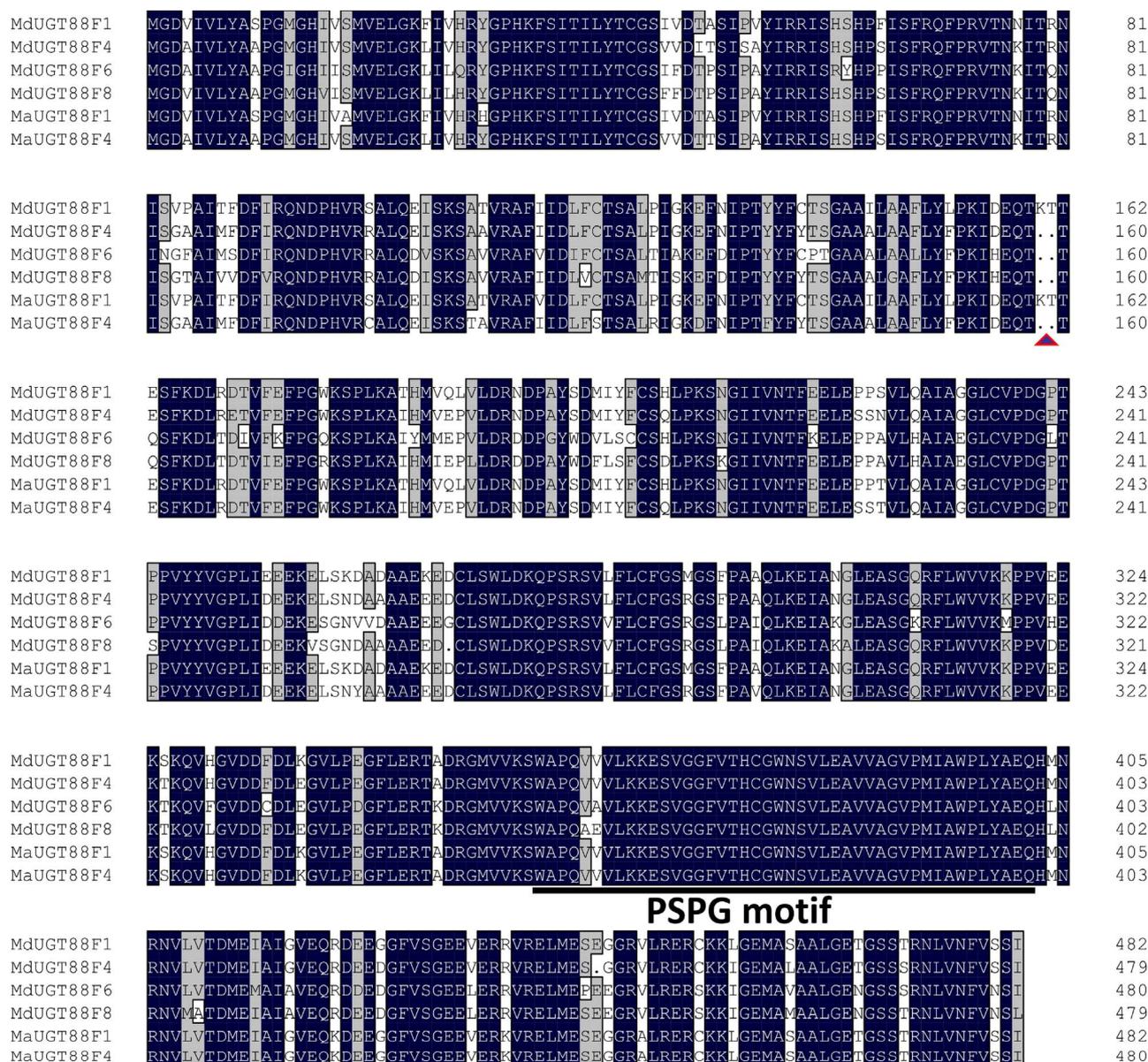


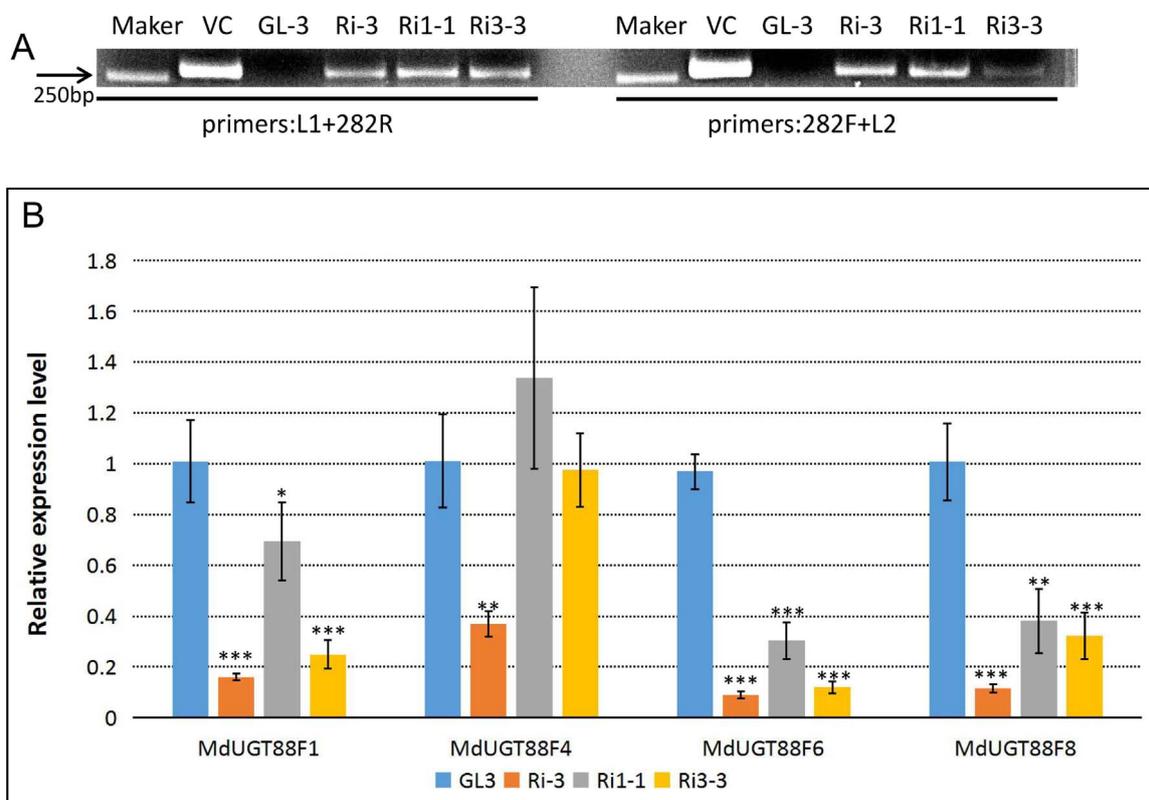
Fig. 9. Alignment of sequences obtained with DNAMAN program. PSPG motif is underlined; triangles indicate K<sup>160</sup> and T<sup>161</sup> residues. Black shading represents identical sequence; grey shading, > 75% sequence similarity.

sieboldin detected independently of trilobatin. Furthermore, progeny of an Evereste × MM106 cross display the DHC composition (and concentration) of one of the parents for either phloridzin, trilobatin, and sieboldin (Evereste profile) or phloridzin alone (MM106 profile) [54]. The concentration of trilobatin is also strongly correlated with that of sieboldin [54]. All of these findings indicate that the activity of P2'GT and, in particular, P4'GT, directly determines DHC profiles, and that F3'H (flavonoid 3'-hydroxylase)/C3H (chalcone-3-hydroxylase) shows high preference to trilobatin in *Malus*.

The etymology of the word 'phloridzin' (from the Greek *phloios* bark + *rhiza* root) implies that it is abundant in the bark [32]. However, we also detected high concentrations in the leaves and branches, but not in the fruit. Although leaf levels were low in some genotypes, total DHCs were still abundant in both leaves and branches of the above-mentioned genotypes. Whereas phloridzin was the predominant DHC in most samples, it was replaced by others, primarily trilobatin and sieboldin, in some genotypes. The DHC profile was also jointly determined by P2'GT and P4'GT activity. Hence, some species harboring extremely high levels of trilobatin (and sieboldin) and much less phloridzin enabled us to

identify the true P2'GTs. Our results showed that phloridzin was rarely found in the leaves and flowers of 'Adams' apple but was accumulated in relatively large amounts in its branches and bark. Based on the level of total DHCs in that cultivar, we suggest that phloridzin is replaced by trilobatin and sieboldin in those leaves and flowers.

Plant UGTs utilize UDP-activated sugars as the major donor molecule. Both P2'GT and P4'GT are members of that protein family. Our phylogenetic analysis showed that the 237 apple UGTs could be clustered into 16 groups, a classification that is in accord with a previous summary [43]. However, even though our results can serve as a good reference point, functional annotations are available for only few of those 237 UGTs. Therefore, further characterizations are needed, because most have, to-date, been inferred based on conserved UGTs identified through the characterization of other plant species. To improve our understanding of the roles of UGTs in various apple tissues, we must continue our monitoring of gene expression patterns. Here, 37 UGTs from each subclade were examined in corresponding samples from 'Adams'. Although all of them were expressed, albeit to varying extents, MDP0000286574 and MDP0000134856 showed tissue-

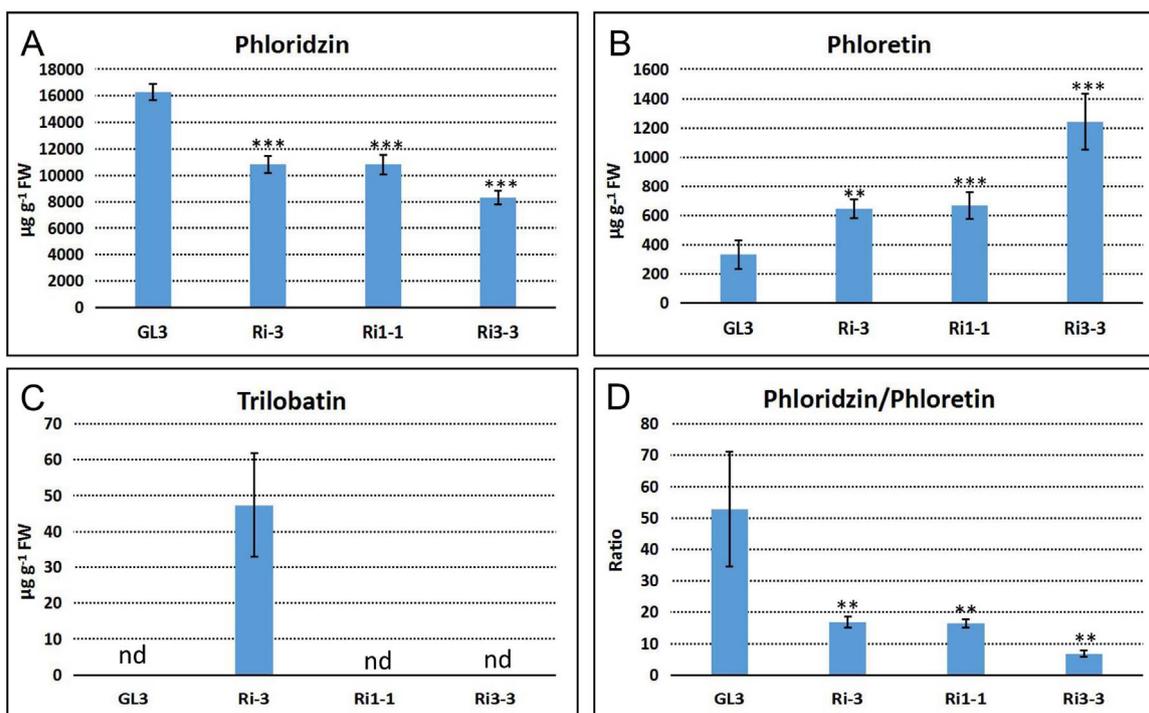


**Fig. 10.** Identification and gene expression of transgenic lines. (A) Genomic-PCR identification of transgenic plants with different combinations of primers; 'VC' represents vector control; (B) Relative expression levels of *MdUGT88F* family in leaves of knockdown lines. Data are means  $\pm$  SD. \*\*\* indicates P-value of < 0.001; \*\* < 0.01; \* < 0.05.

preferential expression, which is consistent with a previous report [41].

Expression of *MDP0000219282* and *MDP0000839414* was strongly correlated with the accumulation of phloridzin in 'Adams'. Meanwhile, comparable expression patterns for those genes were also found in

'Royal Gala', which presented a completely different DHC profile. Therefore, the phloridzin profile could not be reflected briefly by UGT, and was instead jointly determined by the activities of three enzymes in the biosynthesis pathway. Nevertheless, when compared with



**Fig. 11.** Levels of phloretin, phloridzin, and trilobatin in transgenic lines. HPLC was used to determine concentrations of phloridzin (A), phloretin (B), and trilobatin (C). Ratio of phloridzin to phloretin (D) was also calculated in expanding leaves from transgenic lines and 'GL-3' control. Data show mean  $\pm$  SD. \*\*\* indicates P-value of < 0.001; \*\* < 0.01. 'nd', not determined.

phloridzin, the range in variation for amounts of total DHCs was narrower (Supplemental Table 1). As expected, the total DHCs did not individually mirror the expression level of *MdCHS*. Phloridzin concentrations were much lower in members of Group 1 than in the other groups, and the amount of phloridzin in Group-1 samples relative to the others ranged from 2.199 e-3% (ZD7/ZA14) to 1.063% (ZB15/ZK16). Therefore, we believe that this extremely low level of phloridzin is closely related to weak activity by P2'GTs in Group-1 members. Further qRT-PCR analysis with those 41 genotypes verified that only the expression level of *MDP0000219282* was closely related to the amount of phloridzin. Our Blast revealed that *MdPGT1* (from *Malus pumila*), *UGT88F1* (*M. × domestica* 'Rebella'), and *MdP2'GT* (*M. × domestica* 'Golden Delicious') are the proteins coded by *MDP0000219282* in those genotypes. Each appears to accept phloretin, either exclusively or preferentially, as a glycosyl receptor and then selectively glycosylate phloretin at the 2'-OH position to produce phloridzin in vitro. We also confirmed that *MdUGT88F1*, the *MDP0000219282*-coding enzyme in 'Royal Gala', can convert phloretin to phloridzin. The same conversion is mediated by *UGT71K1* (*MDP0000163017*) and *UGT71A15* (*MDP0000215525*) (data not shown). However, the expression profiles of the latter two genes deviate from the phloridzin profiles in 'Adams' apple. Apart from their broader substrate acceptance, poor selectivity of a glycosylation position for *UGT71A15* on phloretin means that they are not the key P2'GT in *Malus* [9]. Although *MdPh-4'-OGT* (*MDP0000617956*) has also been characterized with P2'GT activity in vitro, broader substrate acceptance and extremely low production of phloridzin suggested that this enzyme is not a native P2'GT [40]. We also found that *MdPh-4'-OGT* clusters with *MDP0000839414* in Group L. In that group, the expression patterns of three members (*MDP0000733506* (Fig. 5), *MDP0000229280* (data not shown), and *MDP0000718430* (data not shown)) do not match with the profile of phloridzin in 'Adams'. The accumulation of phloridzin is determined by both the biosynthesis pathway and the lesser-known degradation pathway, although the latter appeared to have only a negligible effect on our study results. This suggests that *MDP0000219282* is the key P2'GT that promotes phloridzin synthesis in *Malus*.

*MDP0000219282* clusters with six other proteins into a subgroup of Group E. However, we could isolate only three of those proteins via RT-PCR, thereby implying that the *UGT88F* subfamily has expanded in parallel with the events of whole-genome duplication that happened after those plant lineages separated. In fact, successive rounds of gene duplication have occurred in other gene families, resulting in a rapid expansion in member numbers and the origins of novel gene functions [55]. Such paleo-duplication events and a relatively recent genome-wide duplication have been revealed previously [40]. In particular, the additional duplication of chromosomal fragments has had a great impact on the amplification of apple genes. Functional divergence among duplicated genes is one of the most important sources of evolutionary innovation in a complex organism. The results from our sequence and phylogenetic analyses suggests that *MdUGT88F1* and *MdUGT88F4* are a paralogous pair, and the latter is still active in phloridzin formation in *Malus*. Although a role for *MdUGT88F4* (*MDP0000375160*) has been described previously [9], we are the first to characterize its involvement in P2'GT activity. The delay in this finding might be due to differences in assay strategies [9]. Here, the reaction product was directly detected in HPLC, which obviously improved the sensitivity of phloridzin. Although the pattern of *MaUGT88F4* (*MDP0000375160*) expression was poorly correlated with phloridzin profiles, an association between the two was still apparent for some of our tested genotypes. In addition, *MdUGT88F4* shared extremely high sequence identity, including the PSPG motif, as well as a similar spatial expression pattern, with *MdUGT88F1*, in 'Royal Gala', which suggested function redundancy between them. Therefore, *MdUGT88F4* is an important, but not key, P2'GT in *Malus*. At the same time, their slight sequence differences are a sign of yet unexplained divergence. Although the expression pattern of *MDP0000461555* reflected the pattern of phloridzin

accumulations in 'Adams', both our recombinant protein assay and expression analysis subsequently proved that they were only weakly associated with phloridzin synthesis. It is possible that all members of the *UGT88F* subfamily originated from one ancestor and underwent radiation during evolution. Whereas *MDP0000461555* retained the expression pattern and lost P2'GT catalytic activity in that cultivar, its paralog, *MDP0000170162*, was possibly differentiated, based on our failure to find either an expression pattern or P2'GT activity. This suggests that differential or tissue-specific expression can be acquired in a new copy after duplication. In fact, approximately two-thirds of all duplicate gene pairs in *Arabidopsis* show divergent expression [56].

Although the P2'GT activities were defined in *MdUGT88F1* and *MdUGT88F4* by in vitro enzyme assay and analysis of correlation of expression, more powerful evidence was necessary. Therefore, we chose gene silencing because of the extremely high amounts of phloridzin in 'GL-3' plants. Fortunately, the phloridzin concentration was significantly decreased, and the aglycone phloretin was correspondingly increased in all transgenic lines where *MdUGT88F1* was significantly silenced. Trilobatin was present only in Ri-3, but not in Ri-1 or Ri-3. The presence of trilobatin was closely related with the lower P2'GT activity and relatively higher P4'GT activity. However, further qRT-PCR found that expression of *MdUGT88F4* was reduced only in Ri-3 and was normal in the other two lines. This suggested that *MdUGT88F4* is important, but not essential, for total P2'GT activity. During the preparation of this manuscript, other researchers have reported that silencing of *MdUGT88F1* can decrease the level of phloridzin and increase the trilobatin concentration in 'Royal Gala' [57]. However, that earlier study showed lower concentrations of phloretin than those measured here, perhaps because of differences in sampling times. For our experiments we selected leaves from *in vitro* plants while the other one relied upon leaves collected from greenhouse-grown seedlings. This indicated that silencing of *MdUGT88F1* in apple can reduce the production of phloridzin, regardless of developmental stage. Although *MdUGT88F4* was not monitored in the earlier work [57], we believe it was significantly silenced. There, a larger fragment (500 bp, spanning positions 429–928 bp) of *MdUGT88F1*, sharing high identity (94.60%) with *MdUGT88F4*, was designed to silence *MdUGT88F1* whereas we applied a smaller fragment (238 bp, spanning 395–632) that had a relatively lower identity (93.28%) with *MdUGT88F4* (Fig. 12). In addition, we found that the expression two other genes in the *MdUGT88F* family, i.e., *MdUGT88F6* and *MdUGT88F8*, was significantly disturbed, but that phenomenon was not mentioned in the previous research [57]. We cannot attribute that behavior to off-target effects because expression of *MdUGT88F4* was comparable in Ri1-1 and Ri3-3. It is more likely that expression of *MdUGT88F6* and *MdUGT88F8* was reduced in parallel with the accumulation of phloretin. As more lower identities between the silencing fragment and *MdUGT88F6* (80.25% and 84.60%, respectively)/*MdUGT88F8* (82.35% and 86.60%, respectively) were found here and in the previous work (Fig. 12). Downregulation was observed for *PAL*, *C4H*, *F3H*, *DRF*, and *FLS*, which that research group attributed to the highly transient accumulation of phloretin [57].

In conclusion, we performed a genome-wide expression analysis of the *UGT* superfamily in several species and cultivars of apple based on a large-scale investigation of phloridzin concentrations. This enabled us to identify two P2'GTs – *MdUGT88F1* and its paralog *MdUGT88F4*. Silencing them in 'GL-3' apple caused those plants to accumulate less phloridzin but more phloretin (and trilobatin). These findings will be beneficial to our future examinations of the physiological roles of dihydrochalcones in apple.

#### Author contributions

K.Z., X.Q.G., and F.W.M. designed the experiments; K.Z. performed the experiments; L.Y.H. was involved in collecting data; P.M.L. provided important information and suggestions; K.Z. wrote the manuscript; and F.W.M. and X.Q.G. critically revised the article.

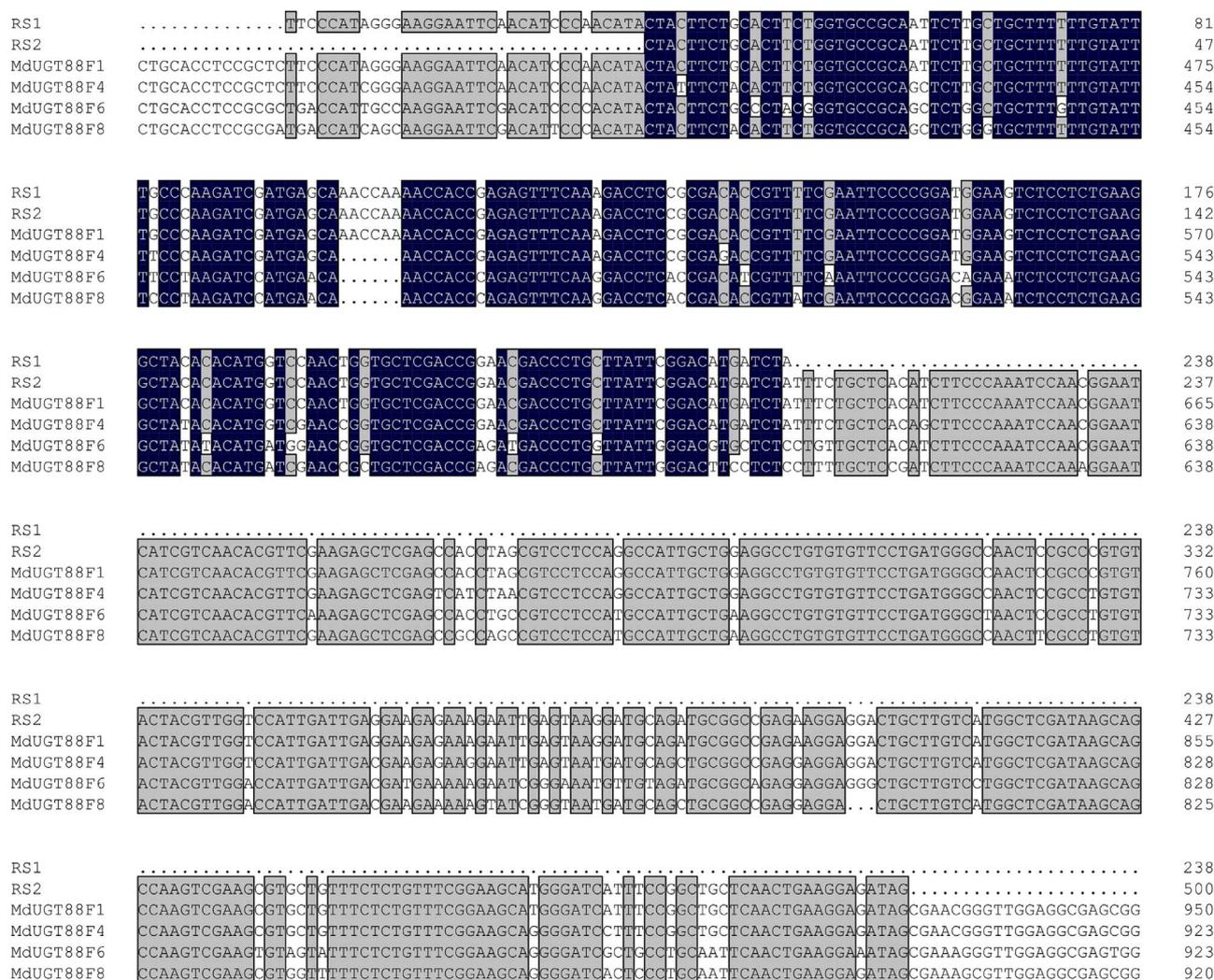


Fig. 12. Alignment of sequences obtained with DNAMAN program. RS1 and RS2 were the silencing sequence applied here and in other study [57], respectively. Black shading represents identical sequence; grey shading, > 75% sequence similarity.

**Acknowledgements**

This work was supported by the National High Technology Research and Development Program of China (863 Program) (2011AA100204), and by the earmarked fund for China Agriculture Research System (CARS-27). The authors are grateful to Dr. Zhihong Zhang for providing tissue-cultured apple plants, to Priscilla Licht for help in revising our English composition.

**Appendix A. Supplementary data**

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.plantsci.2017.10.003>.

**References**

[1] C. Gosch, H. Halbwirth, J. Kuhn, S. Miosic, K. Stich, Biosynthesis of phloridzin in apple (*Malus domestica* Borkh.), *Plant Sci.* 176 (2009) 223–231.  
 [2] M.M. Petkovšek, F. Stampar, R. Veberic, Seasonal changes in phenolic compounds in the leaves of scab-resistant and susceptible apple cultivars, *Can. J. Plant Sci.* 89 (2009) 745–753.  
 [3] J.Y. Cho, S.H. Ji, J.H. Moon, K.H. Lee, K.H. Jung, K.H. Park, A novel benzoyl glucoside and phenolic compounds from the leaves of *Camellia japonica*, *Food Sci. Biotechnol.* 17 (2008) 1060–1065.  
 [4] P. Hilt, A. Schieber, C. Yildirim, G. Arnold, I. Klaiber, J. Conrad, U. Beifuss, R. Carle, Detection of phloridzin in strawberries (*Fragaria x ananassa* Duch.) by HPLC-PDA-MS/MS and NMR spectroscopy, *J. Agric. Food Chem.* 51 (2003) 2896–2899.

[5] E. Hvattum, Determination of phenolic compounds in rose hip (*Rosa canina*) using liquid chromatography coupled to electrospray ionisation tandem mass spectrometry and diode-array detection, *Rapid Commun. Mass Spectrom.* 16 (2002) 655–662.  
 [6] N. Rui-Lin, T. Tanaka, J. Zhou, O. Tanaka, Phlorizin and trilobatin, sweet dihydrochalcone-glucosides from leaves of *Lithocarpus liseifolius* (Hance) Rehd. (Fagaceae), *Agric. Biol. Chem.* 46 (1982) 1933–1934.  
 [7] C. Gosch, H. Flachowsky, H. Halbwirth, J. Thill, R. Mjka-Wittmann, D. Treutter, K. Richter, M.V. Hanke, K. Stich, Substrate specificity and contribution of the glycosyltransferase UGT71A15 to phloridzin biosynthesis, *Trees* 26 (2012) 259–271.  
 [8] L.Z. Lin, J.M. Harnly, Phenolic compounds and chromatographic profiles of pear skins (*Pyrus* spp.), *J. Agric. Food Chem.* 56 (2008) 9094–9101.  
 [9] C. Gosch, H. Halbwirth, B. Schneider, D. Hölscher, K. Stich, Cloning and heterologous expression of glycosyltransferases from *Malus x domestica* and *Pyrus communis*, which convert phloretin to phloretin 2'-O-glucoside (phloridzin), *Plant Sci.* 178 (2010) 299–306.  
 [10] C. Gosch, H. Halbwirth, K. Stich, Phloridzin Biosynthesis, distribution and physiological relevance in plants, *Phytochemistry* 71 (2010) 838–843.  
 [11] B. Lata, A. Trampczynska, J. Paczesna, Cultivar variation in apple peel and whole fruit phenolic composition, *Sci. Hortic.* 121 (2009) 176–181.  
 [12] X.Z. Zhang, Y.B. Zhao, C.M. Li, D.M. Chen, G.P. Wang, R.F. Chang, H.R. Shu, Potential polyphenol markers of phase change in apple (*Malus domestica*), *J. Plant Physiol.* 164 (2007) 574–580.  
 [13] M.M. Petkovšek, F. Stampar, R. Veberic, Increased phenolic content in apple leaves infected with the apple scab pathogen, *J. Plant Pathol.* 90 (2008) 49–55.  
 [14] S. Guo, L. Guan, Y. Cao, C. Li, J. Chen, J. Li, J. Liu, S. Li, B. Wu, Diversity of polyphenols in the peel of apple (*Malus* sp.) germplasm from different countries of origin, *Int. J. Food Sci. Technol.* 51 (2016) 222–230.  
 [15] T. Ridgway, J. O'Reilly, G. West, G. Tucker, H. Wiseman, Potent antioxidant properties of novel apple-derived flavonoids with commercial potential as food additives, *Biochem. Soc. Trans.* 24 (1996) 391S.  
 [16] H.P. Vasantha Rupasingha, A. Yasmin, Inhibition of oxidation of aqueous emulsions of omega-3 fatty acids and fish oil by phloretin and phloridzin, *Molecules* 15 (2010)

- 251–257.
- [17] M. Liaudanskas, P. Viškelis, R. Raudonis, D. Kviklys, N. Uselis, V. Janulis, Phenolic composition and antioxidant activity of *Malus domestica* leaves, *Sci. World J.* (2014) 1–10.
- [18] C. Puel, A. Quintin, J. Mathey, C. Obled, M.J. Davicco, P. Lebecque, S. Kati-Coulibaly, M.N. Horcajada, V. Coxam, Prevention of bone loss by phloridzin an apple polyphenol, in ovariectomized rats under inflammation conditions, *Calcified Tissue Int.* 77 (2005) 311–318.
- [19] E. Makarova, P. Górnaś, I. Konrade, D. Tirezite, H. Cirule, A. Gulbe, I. Pugajeva, M. Dambrova, Acute anti-hyperglycaemic effects of an unripe apple preparation containing phlorizin in healthy volunteers: a preliminary study, *J. Sci. Food Agric.* 95 (2015) 560–568.
- [20] J.R. Ehrenkranz, N.G. Lewis, C.R. Kahn, J. Roth, Phlorizin a review, *Diabetes-Metab. Res.* 21 (2005) 31–38.
- [21] M. Kobori, S. Masumoto, Y. Akimoto, H. Oike, Phloridzin reduces blood glucose levels and alters hepatic gene expression in normal BALB/c mice, *Food Chem. Toxicol.* 50 (2012) 2547–2553.
- [22] M. Najafian, M.Z. Jahromi, M.J. Nowrozezhad, P. Khajeaian, M.M. Kargar, M. Sadeghi, A. Arasteh, Phloridzin reduces blood glucose levels and improves lipids metabolism in streptozotocin-induced diabetic rats, *Mol. Biol. Rep.* 39 (2012) 5299–5306.
- [23] J. Raa, Polyphenols and natural resistance of apple leaves against *Venturia inaequalis*, *Neth. J. Plant Pathol.* 74 (1968) 37–45.
- [24] A. Slatnar, M.M. Petkovšek, H. Halbwirth, F. Stampar, K. Stich, R. Veberic, Enzyme activity of the phenylpropanoid pathway as a response to apple scab infection, *Ann. Appl. Biol.* 156 (2010) 449–456.
- [25] M. Gaucher, T.D. de Bernonville, S. Guyot, J.F. Dat, M.N. Brisset, Same ammo, different weapons: enzymatic extracts from two apple genotypes with contrasted susceptibilities to fire blight (*Erwinia amylovora*) differentially convert phloridzin and phloretin in vitro, *Plant Physiol. Biochem.* 72 (2013) 178–189.
- [26] M.M. Petkovšek, A. Slatnar, F. Stampar, R. Veberic, Phenolic compounds in apple leaves after infection with apple scab, *Biol. Plant.* 55 (2011) 725–730.
- [27] O.S. Hutabarat, H. Flachowsky, I. Regos, S. Miosic, C. Kaufmann, S. Faramarzi, M.Z. Alam, C. Gosch, A. Pei, K. Richter, M.V. Hanke, D. Treutter, K. Stich, H. Halbwirth, Transgenic apple plants overexpressing the chalcone 3-hydroxylase gene of *Cosmos sulphureus* show increased levels of 3-hydroxyphloridzin and reduced susceptibility to apple scab and fire blight, *Planta* 243 (2016) 1213–1224.
- [28] M. Gaucher, T.D. de Bernonville, D. Lohou, S. Guyot, T. Guillemette, M.N. Brisset, J.F. Dat, Histolocalization and physico-chemical characterization of dihydrochalcones: insight into the role of apple major flavonoids, *Phytochemistry* 90 (2013) 78–89.
- [29] L.L. Lairson, B. Henrissat, G.J. Davies, S.G. Withers, Glycosyltransferases Structures, functions, and mechanisms, *Biochemistry* 77 (2008) 521.
- [30] M. Weis, E.K. Lim, N.C. Bruce, D.J. Bowles, Engineering and kinetic characterisation of two glycosyltransferases from *Arabidopsis thaliana*, *Biochimie* 90 (2008) 830–834.
- [31] Y. Li, P. Li, Y. Wang, R. Dong, H. Yu, B. Hou, Genome-wide identification and phylogenetic analysis of Family-1 UDP glycosyltransferases in maize (*Zea mays*), *Planta* 239 (2014) 1265–1279.
- [32] H. Jugdé, D. Nguy, I. Moller, J.M. Cooney, R.G. Atkinson, Isolation and characterization of a novel glycosyltransferase that converts phloretin to phlorizin, a potent antioxidant in apple, *FEBS J.* 275 (2008) 3804–3814.
- [33] T. Zhang, J. Liang, P. Wang, Y. Xu, Y. Wang, X. Wei, M. Fan, Purification and characterization of a novel phloretin-2'-O-glycosyltransferase favoring phloridzin biosynthesis, *Sci. Rep.* 6 (2016) 35274, <http://dx.doi.org/10.1038/srep35274>.
- [34] M. Yahyaa, R. Davidovich-Rikanati, Y. Eyal, A. Shechter, S. Marzouk, E. Lewinsohn, M. Ibdah, Identification and characterization of UDP-glucose: phloretin 4'-O-glycosyltransferase from *Malus × domestica* borkh, *Phytochemistry* 130 (2016) 47–55.
- [35] Y. Zhang, P. Li, L. Cheng, Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh, *Food Chem.* 123 (2010) 1013–1018.
- [36] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-ΔΔCT</sup> method, *Methods* 25 (2001) 402–408.
- [37] Y. Shao, Y. Qin, Y. Zou, F. Ma, Genome-wide identification and expression profiling of the SnRK2 gene family in *Malus prunifolia*, *Gene* 552 (2014) 87–97.
- [38] W. Deng, Y. Wang, Z. Liu, H. Cheng, Y. Xue, Heml. A toolkit for illustrating heatmaps, *PLoS One* 9 (2014) e111988.
- [39] H. Dai, W. Li, G. Han, Y. Yang, Y. Ma, H. Li, Z. Zhang, Development of a seedling clone with high regeneration capacity and susceptibility to *Agrobacterium* in apple, *Sci. Hortic.* 164 (2013) 202–208.
- [40] R. Velasco, A. Zharkikh, J. Affourtit, A. Dhingra, A. Cestaro, A. Kalyanaraman, P. Fontana, S.K. Bhatnagar, M. Troggio, D. Pruss, S. Salvi, M. Pindo, P. Baldi, S. Castelletti, M. Cavaiuolo, G. Coppola, F. Costa, V. Cova, A. Dal Ri, V. Goremykin, M. Komjanc, S. Longhi, P. Magnago, G. Malacarne, M. Malnoy, D. Micheletti, M. Moretto, M. Perazzolli, A. Si-Ammour, S. Vezzulli, E. Zini, G. Eldredge, L.M. Fitzgerald, N. Gutin, J. Lanchbury, T. Macalma, J.T. Mitchell, J. Reid, B. Wardell, C. Kodira, Z. Chen, B. Desany, F. Niazi, M. Palmer, T. Koepke, D. Jiwan, S. Schaeffer, V. Krishnan, C. Wu, V.T. Chu, S.T. King, J. Vick, Q. Tao, A. Mraz, A. Stormo, K. Stormo, R. Bogden, D. Ederle, A. Stella, A. Vecchiotti, M.M. Kater, S. Masiero, P. Lasserre, Y. Lespinasse, A.C. Allan, V. Bus, D. Chagné, R.N. Crowhurst, A.P. Gleave, E. Lavezzo, J.A. Fawcett, S. Proost, P. Rouzé, L. Sterck, S. Toppo, B. Lazzari, R.P. Hellens, C.E. Durel, A. Gutin, R.E. Bumgarner, S.E. Gardiner, M. Skolnick, M. Egholm, Y. Van de Peer, F. Salamini, R. Viola, The genome of the domesticated apple (*Malus × domestica* Borkh.), *Nat. Genet.* 42 (2010) 833–839.
- [41] V.T. Barvkar, V.C. Pardeshi, S.M. Kale, N.Y. Kadoo, V.S. Gupta, Phylogenomic analysis of UDP glycosyltransferase 1 multigene family in *Linum usitatissimum* identified genes with varied expression patterns, *BMC Genom.* 13 (2012) 175.
- [42] Y. Li, S. Baldauf, E.K. Lim, D.J. Bowles, Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*, *J. Biol. Chem.* 276 (2001) 4338–4343.
- [43] L. Caputi, M. Malnoy, V. Goremykin, S. Nikiforova, S. Martens, A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land, *Plant J.* 69 (2012) 1030–1042.
- [44] P.I. Mackenzie, I.S. Owens, B. Burchell, K.W. Bock, A. Bairoch, A. Bélanger, S. Fournel-Gigleux, M. Green, D.W. Hum, T. Iyanagi, D. Lancet, P. Louisot, J. Magdalou, J.R. Chowdhury, J.K. Ritter, H. Schachter, T.R. Tephly, K.F. Tipton, D.W. Nebert, The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence, *Pharmacogenet. Genom.* 7 (1997) 255–269.
- [45] J. Ross, Y. Li, E.K. Lim, D.J. Bowles, Higher plant glycosyltransferases, *Genome Biol.* 2 (2001) 1.
- [46] A.P. Dare, S. Tomes, J.M. Cooney, D.R. Greenwood, R.P. Hellens, The role of enoyl reductase genes in phloridzin biosynthesis in apple, *Plant Physiol. Biochem.* 72 (2013) 54–61.
- [47] A.P. Dare, S. Tomes, M. Jones, T.K. McGhie, D.E. Stevenson, R.A. Johnson, D.R. Greenwood, R.P. Hellens, Phenotypic changes associated with RNA interference silencing of chalcone synthase in apple (*Malus × domestica*), *Plant J.* 74 (2013) 398–410.
- [48] K.T. Watts, P.C. Lee, C. Schmidt-Dannert, Exploring recombinant flavonoid biosynthesis in metabolically engineered *Escherichia coli*, *ChemBioChem* 5 (2004) 500–507.
- [49] Y. Yamazaki, D.Y. Suh, W. Sithithaworn, K. Ishiguro, Y. Kobayashi, M. Shibuya, Y. Ebizuka, U. Sankawa, Diverse chalcone synthase superfamily enzymes from the most primitive vascular plant, *Psilotum nudum*, *Planta* 214 (2001) 75–84.
- [50] S.R. Werner, J.A. Morgan, Expression of a *Dianthus* flavonoid glycosyltransferase in *Saccharomyces cerevisiae* for whole-cell biocatalysis, *J. Biotechnol.* 142 (2009) 233–241.
- [51] R.P. Pandey, T.F. Li, E.H. Kim, T. Yamaguchi, Y.I. Park, J.S. Kim, J.K. Sohng, Enzymatic synthesis of novel phloretin glucosides, *Appl. Environ. Microb.* 79 (2013) 3516–3521.
- [52] J. Tang, L. Tang, S. Tan, Z. Zhou, The study of variation of phloridzin content in six wild *Malus* species, *J. Food Nutr. Res.* 3 (2015) 146–151.
- [53] T.D. de Bernonville, S. Guyot, J.P. Paulin, M. Gaucher, L. Loufrani, D. Henrion, S. Derbré, D. Guilet, P. Richomme, J.F. Dat, M.N. Brisset, Dihydrochalcones Implication in resistance to oxidative stress and bioactivities against advanced glycation end-products and vasoconstriction, *Phytochemistry* 71 (2010) 443–452.
- [54] T.D. de Bernonville, M. Gaucher, S. Guyot, C.E. Durel, J.F. Dat, M.N. Brisset, The constitutive phenolic composition of two *Malus × domestica* genotypes is not responsible for their contrasted susceptibilities to fire blight, *Environ. Exp. Bot.* 74 (2011) 65–73.
- [55] J. Saha, C. Chatterjee, A. Sengupta, K. Gupta, B. Gupta, Genome-wide analysis and evolutionary study of sucrose non-fermenting 1-related protein kinase 2 (SnRK2) gene family members in *Arabidopsis* and *Oryza*, *Comput. Biol. Chem.* 49 (2014) 59–70.
- [56] G. Haberer, T. Hindemitt, B.C. Meyers, K.F.X. Mayer, Transcriptional similarities, dissimilarities, and conservation of cis-elements in duplicated genes of *Arabidopsis*, *Plant Physiol.* 136 (2004) 3009–3022.
- [57] A.P. Dare, Y.K. Yauk, S. Tomes, T.K. McGhie, R.S. Rebstock, J.M. Cooney, R.G. Atkinson, Silencing a phloretin-specific glycosyltransferase perturbs both general phenylpropanoid biosynthesis and plant development, *Plant J.* 91 (2017) 237–250.