Plant Mol Biol (2007) 64:499–518
DOI 10.1007/s11103-007-9167-6

A functional genomics approach to (iso)flavonoid glycosylation in the model legume *Medicago truncatula*

Luzia V. Modolo · Jack W. Blount · Lahoucine Achnine · Marina A. Naoumkina · Xiaoqiang Wang · Richard A. Dixon

Received: 15 February 2007 / Accepted: 21 March 2007 / Published online: 17 April 2007

Abstract Analysis of over 200,000 expressed sequence tags from a range of *Medicago truncatula* cDNA libraries resulted in the identification of over 150 different family 1 glycosyltransferase (*UGT*) genes. Of these, 63 were represented by full length clones in an EST library collection. Among these, 19 gave soluble proteins when expressed in *E. coli*, and these were screened for catalytic activity against a range of flavonoid and isoflavonoid substrates using a high-throughput HPLC assay method. Eight UGTs were identified with activity against isoflavones, flavones, flavonols or anthocyanidins, and several showed high catalytic specificity for more than one class of (iso)flavonoid substrate. All tested UGTs preferred UDP-glucose as sugar donor. Phylogenetic analysis indicated that the *Medicago* (iso)flavonoid glycosyltransferase gene sequences fell into a number of different clades, and several clustered with UGTs annotated as glycosylating non-flavonoid substrates. Quantitative RT-PCR and DNA microarray analysis revealed unique transcript expression patterns for each of the eight UGTs in *Medicago* organs and cell suspension cultures, and comparison of these patterns with known phytochemical profiles suggested in vivo functions for several of the enzymes.

Electronic supplementary material The online version of this article (doi:10.1007/s11103-007-9167-6) contains supplementary material, which is available to authorized users.

L. V. Modolo · J. W. Blount · L. Achnine · M. A. Naoumkina · X. Wang · R. A. Dixon (✉)
Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
e-mail: radixon@noble.org

Present Address:
L. Achnine
GenApps, Inc., 4262 Colby Road, Winchester, KY 40391, USA

Keywords Flavonoids · Isoflavonoids · Glycosylation · Glycosyltransferase · Regio-selectivity

Abbreviations

EST	Expressed sequence tag
UGT	Family 1 uridine diphosphate glycosyltransferase
MeJA	Methyl jasmonate
YE	Yeast elicitor

Introduction

Members of the flavonoid class of plant natural products have broad activities in plant defense, symbiosis, development and pollination (Stafford 1997; Shirley 1998; Aoki et al. 2000; Forkmann and Martens 2001). They also exhibit a wide range of beneficial effects on human health, including protection against cardiovascular disease and certain forms of cancer (Formica and Regelson 1995; Kanadaswami et al. 2005). Various flavonoids possess antimicrobial (Farmakalidis and Murphy 1984; Kape et al. 1991), antioxidant (Borek 2006), antiviral (Kaul et al. 1985), and antiallergic/antinflammatory (Gil et al. 1994) activities. The isoflavonoids are also estrogenic (Bingham et al. 1998; Dixon 2004), and constitute a major class of pre-formed and inducible antimicrobial defense compounds in the Leguminosae (Dixon and Sumner 2003).

Flavonoids and isoflavonoids are commonly found in nature as glycoconjugates in which the glycosyl residue can be crucial for bioactivity. A better understanding of this relationship may lead to development of new and more active and/or effective glycodrugs (Kren and Martinkova

2001). For example, the flavonol quercetin, one of the major health-beneficial antioxidants derived from plant sources (Formica and Regelson 1995; Kanadaswami et al. 2005) (Fig. 1, **20**), is usually found glycosylated on one or more of its five hydroxyl groups to increase its solubility and stability. The position of conjugation has a significant impact on its biological activity and bioavailability, and therefore on its potential health benefits for humans (Day et al. 2000, 2003). Glycosides are generally more water-soluble than their corresponding aglycones, and this influences the pharmacokinetic properties of the compound. Thus, absorption of quercetin glucosides occurs in the small intestine at higher efficiency than that of the aglycone (Manach et al. 2004).

Family 1 uridine diphosphate (UDP) glycosyltransferases (UGTs) are inverting glycosyltransferases that transfer UDP-activated sugar moieties to specific acceptor molecules. Plants contain very large *UGT* gene families, and over 100 putative *UGT* genes have been identified in *Arabidopsis thaliana* (Ross et al. 2001). Plant UGTs can be divided into three distinct clades (Paquette et al. 2003). The members of the plant-specific clade are characterized

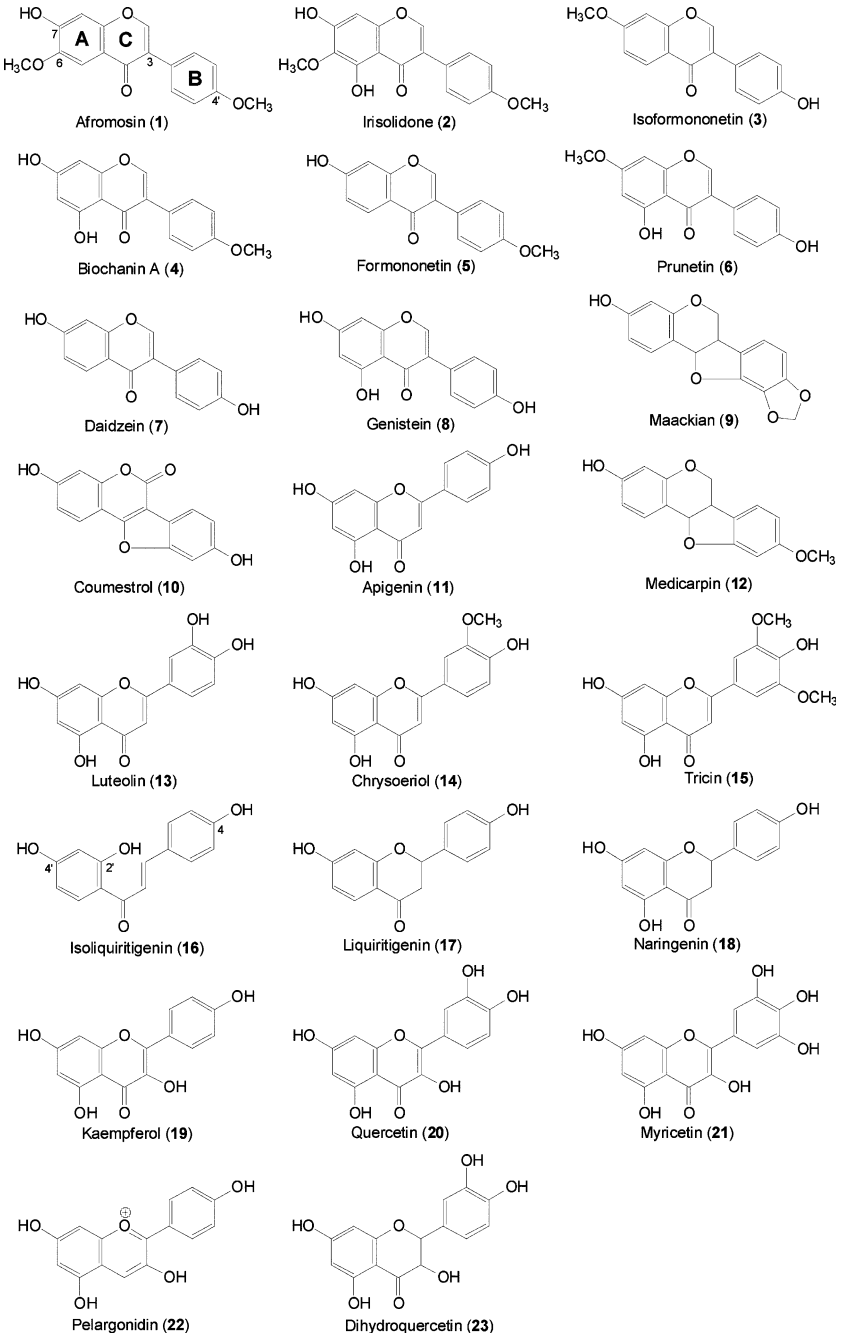


Fig. 1 Acceptor substrates used for determining the substrate preferences of *Medicago* (iso)flavonoid UGTs. Note that the ring position numbering convention shown for (**1**) is identical for all the flavonoid classes except for the chalcone (**16**) and the pterocarpans (**9, 12**). Compounds are referenced by bold numbers in the text

by a highly conserved consensus signature motif denoted as the PSPG (Putative Secondary Plant Glycosyltransferase) motif (Hughes and Hughes 1994; Paquette et al. 2003). Many of these plant UGTs are involved in glycosylation of natural products (Vogt and Jones 2000; Dangl and Jones 2001).

The relationship between primary sequence and substrate specificity of plant UGTs is not yet clear. For example, among 11 *Arabidopsis* UGTs that catalyzed glycosylation at the 3-hydroxyl of quercetin, identity at the amino acid level ranged from only 20% to 72%. Betanidin (a betalain) 6-*O*- and 5-*O*-glucosyltransferases from *Dorotheanthus bellidiformis* share only 19% amino acid sequence identity (Vogt 2002). Thus, there are currently no reliable methods to identify the substrate specificity or regio-selectivity of a plant small molecule glycosyltransferase from sequence information alone.

Collectively, legumes accumulate a wide variety of glycosylated natural products (Dixon and Sumner 2003). In the model legume *Medicago truncatula*, these are primarily flavonoids (flavones, anthocyanins), isoflavonoids (isoflavones, pterocarpans) and triterpene saponins (Huhman and Sumner 2002; Dixon and Sumner 2003), suggesting the presence of a range of UGTs specific for the biosynthesis of these compounds in this species.

We here describe a genomics approach, coupled with high throughput activity screening, to identify a range of *Medicago* UGTs with activity against various flavonoid and isoflavonoid natural products, including quercetin. Interestingly, these *Medicago* (iso)flavonoid *UGT* genes, although encoding proteins with overlapping specificities for certain acceptor substrates, fall into several phylogenetically distinct clades, and are differentially expressed in various tissues of the plant. Most of the enzymes exhibit multiple activities in vitro, and could be developed as useful reagents for the enzymatic synthesis of various bioactive (iso)flavonoid glucosides. Their in vivo activities were predicted based upon comparisons of transcript and metabolite profiles in different tissues.

Materials and methods

Plant materials

M. truncatula plants were grown in Turface MVP^{TM} (Profile Products, Buffalo Grove, IL) under controlled conditions in the greenhouse (28°C, 40% relative humidity, 16 h daylength). Roots, leaves, flowers, stems, and buds were collected from 8 week-old to 9 week-old plants, frozen in liquid nitrogen, and stored at –80°C until use. For infection studies, leaves of 8 week-old to 9 week-old *M. truncatula* plants were wounded by running a tracing wheel across the adaxial surface of the leaf, and inoculated with a spore suspension (6×10^5 spores/ml) of *Phoma medicaginis* in 0.1% Tween 20. The plants were bagged and maintained in a growth chamber as described (Deavours and Dixon 2005). Leaves and stems were harvested 72 h post-inoculation, frozen in liquid nitrogen, and stored at –80°C until use. *M. truncatula* cell suspension cultures were grown and treated with yeast elicitor (YE) or methyl jasmonate (MeJA) as described previously (Suzuki et al. 2005).

Chemicals

UDP-[glucose-(U)^{14}C] (300 mCi/mmol) was purchased from American Radiolabeled Chemicals, Inc. (St. Louis, MO, USA). Benzoic acid, biochanin A, caffeic acid, catechin, *trans*-cinnamic acid, coniferyl alcohol, *trans p*-coumaric acid, coumestrol, dihydroquercetin, epicatechin, esculetin, ferulic acid, *p*-hydroxybenzoic acid, kaempferol, luteolin, quercetin, salicylic acid, scopoletin, and vanillin from Sigma-Aldrich (St. Louis, MO, USA); dihydrokaempferol from Apin (Abingdon, Oxon, UK); apigenin, chrysoeriol, cyanidin, daidzein, delphinidin, formononetin, genistein, isoliquiritigenin, liquiritigenin, myricetin, naringenin, pelargonidin, and prunetin from Indofine (Hillsborough, NJ, USA); and afromosin from Chromadex (Santa Ana, CA, USA). Tricin was kindly provided by Dr. Tom Mabry (University of Texas at Austin).

Field grown alfalfa roots as a source of (–)-medicarpin were extracted in acetone (1:2, w/v). The extracts were concentrated to the aqueous phase under vacuum, extracted three times with an equal volume of ethyl acetate, and the pooled ethyl acetate fractions taken to dryness under vacuum. The residues were then resuspended in HPLC grade methanol (J.T. Baker). Medicarpin was purified by HPLC following published methods (Kessmann et al. 1990), except that 1% H_3PO_4 was replaced by H_2O in the gradient systems. Maackiain was purified from red clover roots as described previously (VanEtten et al. 1983).

Cloning and expression of UGTs

After obtaining the complete nucleotide sequences of the full length *UGT*s in the EST collection, a number were cloned into Gateway compatible cassettes. The PCR product was amplified and subsequently introduced using BP recombination into pDONR201 plasmid (Invitrogen) to create an Entry clone. The latter served as a substrate for the LR recombination reaction to transfer the insert into the destination vector pDEST17 (Invitrogen) for expression in *E. coli.* Primers for Gateway cloning were: GT22D, forward primer GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGTATCCCAAGATCCAAAGGT, reverse

primer GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATTTTGAATTGGCATTTCCATG; GT22E09, forward GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAGGACACTATAGTTCTCTA, reverse GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAGCAAAAAGAAAGGGGGAATT; GT63G, forward GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCATCCGAAGCTTCAATTCA, reverse GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAACAACGTTTGTGTTTCTTAAT.

Cloning of other UGTs was performed using standard procedures (Sambrook et al. 1989). The *M. truncatula* EST clone NF083F04ST (GT83F, corresponding to TIGR TC105632) was PCR amplified from pBluescript II SK+ (Stratagene, La Jolla, CA, USA) with addition of BamHI and NotI sites (5′-CGGATCCATGTCTACCTTCAAAAATG-3′, upstream primer; 5′-TGCGGCCGCACTAGTGACAATTTG-3′, downstream primer). The PCR product was purified, ligated to pGEMTeasy vector (Promega, Madison, WI, USA), sequenced, excised and re-cloned between the BamHI and NotI sites of pET28a(+) (Novagen, Madison, WI, USA) with a hexahistidine tag and a thrombin cleavage site. *E. coli* BL21 (DE3) cells harboring the expression construct were grown to an OD_{600} of 0.4–0.5, and expression was initiated by addition of isopropyl 1-thio-β-D-galactopyranoside (IPTG) at a final concentration of 0.2 mM, with further incubation with shaking overnight at 16°C.

GTs 29C, 29H, 67A and 99D were expressed and purified using a similar strategy. In the case of GT29C, the primers were 5′-CGGATCCATGAAGGATACTTTAGTTCTTTAC-3′ (upstream) and 5′-TGCGGCCGCTCAACGAAGAAGGAAAGGGGAATT-3′ (downstream). For GT29H, the primers were 5′-CGGATCCATGTCTATGAGTGATATAAACAAG-3′ (upstream) and 5′-TGCGGCCGCTCAGTTGCTTCCTGTAATATCATC-3′ (downstream); for GT67A, 5′-CGGATCCATGGGAAACTTTGCAAACAGAAAAC-3′ (upstream) and 5′-TGCGGCCGCTTAATTTTGTTTAAGCAACACATC-3′ (downstream); for GT99D, 5′-CGGATCCATGGAATCCTTTGGTGTGAAAGTT-3′ (upstream) and 5′-TGCGGCCGCCTAAGTTACTGATACCAATTTCTT-3′ (downstream).

High-throughput screening of putative UGT substrates

Equal amounts of each purified recombinant UGT protein were incubated with four independent solutions, each containing eight different compounds that could be baseline-resolved by HPLC and had clearly distinguishable UV absorption spectra (Table S1). The mixtures (0.1 mM of each component) were incubated with 1–5 μg of recombinant protein, 50 mM Tris–HCl, pH 7.0, 14 mM β-mercaptoethanol and 5 mM UDP-glucose in a final volume of 200 μl at 30°C for 1–3 h. Reactions were stopped by adding trichloroacetic acid (TCA) to a final concentration of 24 mg/ml, and the reaction mixtures were extracted with ethyl acetate (500 μl). The organic fractions were collected, dried, and re-suspended in 50 μl of HPLC grade methanol. Samples (20 μl) were analyzed by HPLC (Agilent HP1100 series HPLC fitted with a diode array detector, quaternary pump, and autosampler). The assays were monitored at 230, 254, 280, 315, and 530 nm. Solvent A was 1% phosphoric acid in milli-q water and solvent B was HPLC grade acetonitrile. The flow rate was 1 ml/min. The column was either a Metachem Waters Spherisorb 5 μ or a Varian Metasil Basic 5 μ 250 × 4.6 mm C18 ODS reverse phase column with a corresponding pre-column purchased from Varian (Walnut Creek, CA). The system was equilibrated at 5% B, and samples eluted with 5% B for 10 min followed by a multi-step gradient from 5% to 10% B (5 min), 10% to 17% B (15 min), 17% to 23% B (5 min), 23% to 50% B (35 min), and from 50% to 100% B (4 min). Isocratic elution with 100% B (acetonitrile) was then maintained for 10 min with a flow rate of 1 ml/min.

Enzyme assays with individual substrates

Enzyme reactions were performed with 1–15 μg of enzyme in a total volume of 50 or 200 μl containing 50 mM Tris–HCl pH 7.0, 1.0–5.0 mM UDP-glucose, UDP-galactose or UDP-glucuronic acid, and 100–250 μM acceptor substrate at 30°C. Reactions were stopped with TCA and products analyzed by HPLC as described above.

For kinetic analysis of GT83F, purified enzyme (1.25 μg) was added to reaction mixtures (50 μl final volume) containing 50 mM Tris–HCl pH 7.0, 10 μM UDP–[U–^{14}C]-glucose (0.3 Ci/mmol), 490 μM UDP-glucose (unlabelled), and 0–500 μM acceptor substrate. Reactions were stopped with TCA after 15 min incubation at 30°C. Samples were extracted with 250 μl of ethyl acetate, and 200 μl were taken for liquid scintillation counting (Beckman LS6500). For kinetic analysis of GT67A, purified enzyme (8.5 μg) was added to reaction mixtures (50 μl final volume) containing 50 mM Tris–HCl pH 7.0, 20 μM UDP–[U–^{14}C]-glucose (0.3 Ci/mmol), 980 μM UDP-glucose (unlabelled), 5 mM $MnCl_2$, 500 μM ATP, and 0–500 μM substrate. Reactions were stopped with TCA after 30 min incubation at 30°C. Similar procedures were performed for determination of K_m values for UDP-glucose, except unlabelled sugar donor was used, and analysis of product formation was by HPLC as described above. Data were analyzed using Hyper32 software (http://www.liv.ac.uk/~jse/software.html).

Assays of reverse catalysis by GT83F were done by incubating 1.25 μg of protein with 250 μM flavonoid 3-*O*-glucosides or 7-*O*-glucosides in 50 mM Tris–HCl pH

7.0 with or without 5.0 mM UDP. After 3 h of incubation at 37°C, samples were extracted with ethyl acetate and processed as described for the UGT activity measurement.

Phylogenetic analysis

Amino acid sequences were analyzed by the neighbor-joining method using the distance matrix derived from the multiple alignment of ClustalW (http://www.align.genome.jp). The phylogenetic tree was drawn using the software PhyloDraw version 0.8 (http://www.pearl.cs.pusan.ac.kr/phylodraw).

Primer design for real time qRT-PCR experiments

PCR primers (Table S2) were designed using Primer3 software (http://www.frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Design criteria were Tm values of 60 ± 1°C, PCR amplicon lengths of 60–150 bp, and 18–24 nucleotide primers with GC content of 40–60%. Primer quality was checked using NetPrimer software (http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html). The specificity of the primer pair sequences was checked against the *Medicago truncatula* transcript database using nucleotide-nucleotide BLAST (http://www.ncbi.nlm.nih.gov/BLAST/).

Preparation of RNA

Total RNA was isolated from tissues or cultured cells using TRIreagent (Molecular Research Center, Inc, Cincinatti, OH, USA) according to the manufacturer's instructions. RNA samples were treated with Turbo DNA-free DNase I (Ambion) according to the manufacturer's instructions, and checked for genomic DNA contamination by PCR using the primers 5′-GTCCTCTAAGGTTTAATGAACCGG-3′ (upstream) and 5′-GAAAGACACAGCCAAGTTGCAC-3′ (downstream) designed to amplify an intron sequence of the *M. truncatula* ubiquitin gene (TC# 102473), and the primers 5′-ATTGCCTGCCCAAGAGTGTAAG-3′ (upstream) and 5′-CAGCCAAGTTGCACAAAACAAC-3′ (downstream) designed to amplify an intron/exon fragment of the same gene. RNA integrity was evaluated with an Agilent 2100 Bioanalyzer using RNA nano chips. RT reactions were done using the DNase I-treated RNAs and SuperScript III reverse transcriptase (Invitrogen), according to the manufacturer's instructions.

Real-Time qRT-PCR and DNA microarray analyses

PCR reactions were performed in an optical 384-well plate with an ABI PRISM 7900 HTsequence detection system (Applied Biosystems, Foster City, CA, USA), using SYBR Green to monitor dsDNA synthesis. Reactions contained 2.5 μl of SYBR Green Master Mix reagent (Applied Biosystems), 0.5 μl of cDNA, and 200 nM of each gene-specific primer in a final volume of 5 μl. PCR reactions were performed as described elsewhere (Czechowski et al. 2005). Data were analyzed using the SDS 2.2.1 software (Applied Biosystems). PCR efficiency (E) was estimated using the LinRegPCR software (Ramakers et al. 2003) and the transcript levels were determined by relative quantification (Pfaffl 2001) using the actin gene (TC# 107326) as a reference.

DNA microarray analysis was performed utilizing Affymetrix *Medicago* genome arrays with 61K probe sets (http://www.affymetrix.com) as described previously (Deavours et al. 2006).

Results

Selection and expression of *Medicago* family 1 UGTs

Using both keyword searching (glycosyltransferase, glucosyltransferase) and the sequence of the PSPG box as diagnostic features, informatic analysis of the TIGR *M. truncatula* EST database (now hosted at the Computational Laboratory and Functional Genomics Laboratory of the Dana Farber Cancer Institute at Harvard University, http://www.compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=medicago) revealed 164 distinct tentative consensus sequences (TCs) putatively encoding family 1 UGTs. Of these, 63 were subsequently shown to be represented by full length clones in the Noble Foundation's *M. truncatula* EST library collection. The nomenclature of these clones, the GenBank accession #s of the ESTs, and their tentative consensus numbers in the TIGR database, are summarized in Table S2. After obtaining the complete nucleotide sequences of these *UGT*s, 39 of them were cloned into Gateway compatible cassettes and transferred into the pDEST17 vector (Invitrogen) for expression in *E. coli*. A further 12 were subcloned into *E. coli* expression vector pET28a by traditional methods. Twenty seven of the Gateway clones expressed recombinant protein, and in 11 cases the protein was recovered in the soluble fraction. Of the traditionally cloned constructs, eight gave soluble protein.

The 19 soluble UGTs were expressed as hexa-histidine N-terminal fusion proteins and purified from cell supernatants by nickel affinity chromatography. As an example, Fig. 2 shows SDS-PAGE analysis of *E. coli* lysates expressing GT67A and GT83F. The protein was primarily recovered from the soluble fraction of the cells, and was obtained in a highly purified form by affinity purification. The UGTs successfully expressed and purified to

homogeneity by this approach were GT10A, GT12E, GT13H, GT22D (official UGT designation UGT73C8, GenBank Accession # DQ875459), GT22E09 (UGT88E1, DQ875460), GT29C (UGT88E2, DQ875461), GT29H (UGT71G1, AAW56092), GT43G, GT49F (UGT73K1), GT54H, GT55D, GT63G (UGT84F1, DQ875462), GT67A (UGT85H2, DQ875463), GT68E, GT73H, GT83F (UGT78G1, DQ875464), GT86C, GT99D (UGT73P1, DQ875465) and GT106F, where the initial numbers refer to the designation in the Noble Foundation EST collection, and the numbers in parentheses refer to official UGT designations for those enzymes subsequently characterized with regards to substrate specificity, and the GenBank accession number of the complete open reading frame.

High throughput screening for (iso)flavonoid UGT activity

A method was devised for high-throughput assay of recombinant UGTs against a range of plant-derived phenolic compounds (see Materials and Methods). *E. coli*-expressed recombinant hexa-histidine-tagged enzyme was purified using Magne-His beads, and equal amounts of each protein incubated with uridine diphosphate glucose (UDPG) as sugar donor and mixtures (four in total) of putative acceptor molecules each containing eight different compounds that could, within each mixture, be baseline-resolved by HPLC (Table S1, Fig. 3) and had clearly distinguishable UV absorption spectra. If the recombinant enzyme showed activity against any one of the eight substrates within a particular mixture, a new peak (glucoside) appeared on the HPLC trace (Fig. 3A, B) with a UV spectrum nearly identical to that of one of the eight aglycones. The enzyme was then tested with that substrate independently. For example, when GT22E09 was incubated with substrate group II (consisting of caffeic acid, coniferyl alcohol, benzoic acid, dihydrokaempferol [dihydroflavonol], quercetin [flavonol, 20], coumestrol [coumestan, **10**], medicarpin [pterocarpan, **12**] and biochanin A [isoflavone, **4**]), a new peak (labeled k) appeared at a retention time 45.5 min (Fig. 3B), with a nearly identical UV spectrum to that of biochanin A (Fig. 3A, peak g). By the same criteria, additional peaks (h, i, j) represented glucosides of coumestrol, quercetin and medicarpin, respectively. Re-assaying the enzyme with biochanin A as the sole substrate confirmed production of biochanin A-7-*O*-glucoside by comparison with an authentic standard (Fig. S1). Using this screening method, eight out of the 19 UGTs were shown to exhibit activity against at least one (iso)flavonoid acceptor.

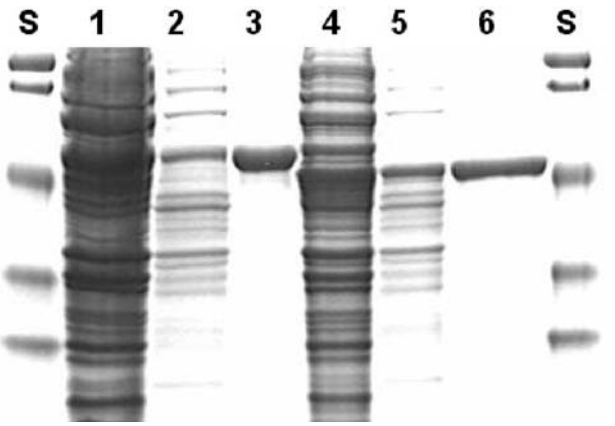


Fig. 2 SDS-PAGE analysis of GT67A (UGT85H2) and GT83F (UGT78G1) proteins. **S**, molecular weight standards; **1**, insoluble fraction from *E. coli* expressing GT67A; **2**, soluble fraction containing GT67A (crude extract); **3**, purified GT67A; **4**, insoluble fraction from *E. coli* expressing GT83F; **5**, soluble GT83F (crude extract); **6**, purified GT83F

Substrate specificities of *Medicago* (iso)flavonoid UGTs

On the basis of preliminary experiments in which kinetic parameters were measured for two of the UGTs, conditions were determined for comparative screening of the eight (iso)flavonoid UGTs against a total of 23 independent substrates (Fig. 1), which included the (iso)flavonoids analyzed in the preliminary high throughput screens (Table S1), plus the chalcone isoliquiritigenin (**16**), the flavone chrysoeriol (**14**), the flavonol myricetin (**21**), and the isoflavones isoformononetin (**3**), afromosin (**1**) and irisolidone (**2**). Glucosides of the two latter compounds were recently identified as important secondary metabolites of *M. truncatula* cell suspension cultures (Farag et al. 2007). The concentration of UDPG was set at 5 mM, and the acceptor substrate concentration was set at 100 μM for the assays with the 23 individual substrates. In preliminary work with GT22D and GT22E09, higher concentrations of acceptor substrate were shown to be inhibitory. GT22D exhibited no activity with 1 mM kaempferol, yet had excellent activity when the concentration was reduced to 100 μM. In contrast, GT22E09 had good activity with 100 μM or 1 mM kaempferol. Initial tests with GT22D and GT22E09 with positive acceptor substrates (100 μM) showed linear activity through two hours. For all other UGTs tested, reactions with 100 μM acceptor substrate were also linear for at least 2 h. In some cases after overnight incubations, small amounts of product were sometimes observed for UGT reactions that did not give measurable product within 2 h. This activity was regarded as insignificant and not pursued further.

Table 1 shows the substrate preferences of the eight UGTs previously shown to have activity with at least one (iso)flavonoid compound. The data, which should be

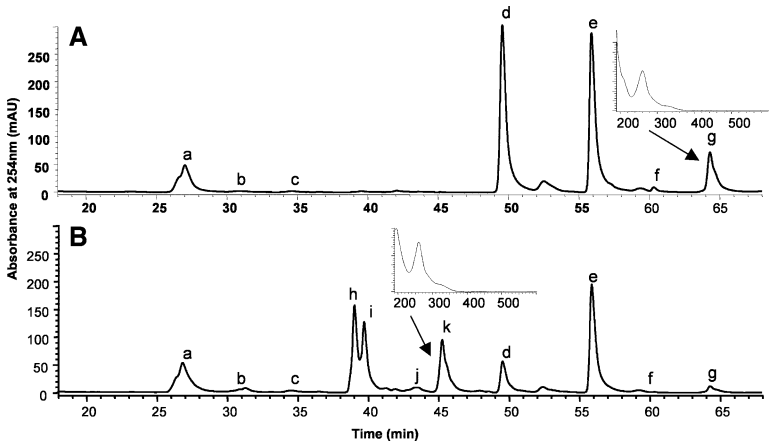


Fig. 3 HPLC analysis of products formed by GT22E09 (UGT88E1) in the presence of UDPG and group II substrates (Table S1). HPLC chromatograms of control protein extract from *E. coli* expressing an empty vector (**A**) and GT22E09 (**B**) incubated with substrate mixture group II. Peaks labeled are: a, caffeic acid; b, coniferyl alcohol; c, benzoic acid; d, quercetin; e, coumestrol; f, medicarpin; g, biochanin A; h, coumestrol glucoside; i, quercetin glucoside; j, medicarpin glucoside; k, biochanin A glucoside. Insets show UV spectra of compounds marked with arrows

viewed as semi-quantitative in the absence of detailed kinetic measurements for every enzyme, are given as % conversion of substrate after incubation (in most cases) of 2 μg of pure enzyme with 5 mM UDPG and 100 μM donor substrate for 2 h. All UGTs had activity with at least three of the 23 substrates, and some had activity with as many as 19 substrates but with varying conversion rates.

Three of the enzymes (GTs 22D, 22E09 and 29C) showed an overall preference for isoflavonoids, being able to glycosylate various isoflavones, including pterocarpan phytoalexins. Two enzymes (GT63G and GT99D) exhibited weak activity only against certain flavonols; it is possible that flavonoids are not the natural substrates for these enzymes. Of the eight enzymes, only GT83F was able to glycosylate anthocyanidins, although other (iso)flavonoids were better in vitro substrates.

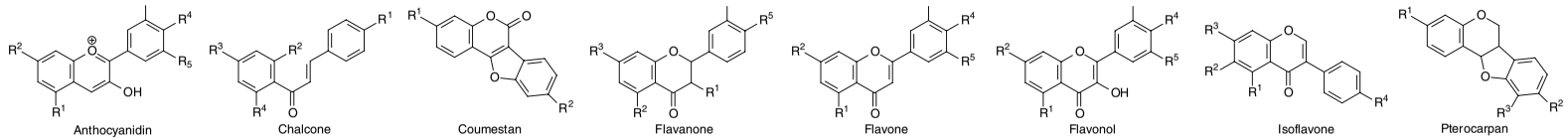
By comparing the retention times of glycosylated products with those of authentic standards, it was possible to determine the preferred positions of glycosylation for many of the enzyme: substrate combinations. These are summarized in Table S3. Isoflavonoids were always glycosylated at the 7-position (A-ring), and occasionally at the 4′-position for compounds with no free 7-hydroxyl. Flavonoids could be glycosylated at multiple positions by some of the enzymes, with different enzymes showing preference for the 7, 4′, 3′, or 3-positions.

Kinetic analysis of GT67A and GT83F

Because of their wide activity profiles against anthocyanidins, flavones, flavonols and isoflavones (Table 1, Table S4), GT67A and GT83F were selected for in-depth kinetic analysis of substrate specificity. The crystal structure and biochemical properties of GT67A are reported elsewhere (Li et al. 2007), and its kinetics are only briefly summarized here. The K_m for UDP-glucose was determined as 485 μM. GT67A exhibited similar, very low K_m values for the isoflavone biochanin A (4.8 μM), the chalcone isoliquiritigenin (9.9 μM) and the flavonol kaempferol (2.9 μM). However, GT67A had higher substrate preference for flavonol than for isoflavone (K_{cat}/K_m = 179.6 s^{-1} M^{-1} for kaempferol, and 28.5 for biochanin A). No glycosylated products were observed when UDP-galactose or UDP-glucuronic acid were used as sugar donor, except in assays with biochanin A, where a single glucosylated product was formed at barely detectable levels.

As with GT67A, GT83F displayed typical Michaelis-Menten kinetics for the various acceptor substrates (Fig. S2, Table 2). A comparison of substrate structure-activity relationships revealed that the carbonyl group on ring C of the (iso)flavonoid acceptor is not critical, since the enzyme glycosylates anthocyanidins, which lack this group (Tables 1, 2). Substitution of the B-ring hydroxyl group at C4' with a methyl group increases the catalytic specificity (K_{cat}/K_m ratio) for the substrate when no hydroxyl group is present at C5 of the A ring. Thus, GT83F prefers formononetin (**5**) to daidzein (**7**) by about 2-fold. Although the K_m values for biochanin A (4′-methoxy genistein, **4**) and genistein (**8**) are different (11.7 and 36.7 μM, respectively), GT83F exhibits similar overall specificity for these compounds under the experimental conditions used. Likewise, only slight differences in specificity were observed between genistein (**8**) and daidzein (**7**), even though the K_m for the former is 23-fold higher than for the latter.

Table 1 Activities of (iso)flavonoid glycosyltransferases from *Medicago truncatula*. Results are given as % of substrate converted



Substrate	GT22D	GT22E09	GT29C	GT29H	GT63G	GT67A	GT83F	GT99D
Anthocyanidin								
Pelargonidin ($R^1 = R^2 = R^4 = OH$; $R^3 = R^5 = H$)	NA	NA	NA	NA	NA	NA	46.0[a]	NA
Chalcone								
Isoliquiritigenin ($R^1 = R^2 = R^3 = OH$; $R^4 = H$)	57.6	48.7	9.8	59.7	NA	5.7[b]	6.0	NA
Coumestan								
Coumestrol ($R^1 = R^2 = OH$)	NA	50.9	50.9	46.5	NA	NA	76.6	NA
Flavanone								
Dihydroquercetin ($R^1 = R^2 = R^3 = R^4 = R^5 = OH$)	NA	NA	NA	NA	NA	NA	NA[c]	11.4
Liquiritigenin ($R^1 = R^2 = R^4 = H$; $R^3 = R^5 = OH$)	34.1	NA	NA	9.2	NA	NA	NA	NA
Naringenin ($R^1 = R^4 = H$; $R^2 = R^3 = R^5 = OH$)	90.6	NA	5.1	26.2	NA	NA	18.1	NA
Flavone								
Apigenin ($R^1 = R^2 = R^4 = OH$; $R^3 = R^5 = H$)	18.1	13	17.8	6	NA	NA	16.2	NA
Chrysoeriol ($R^1 = R^2 = R^4 = OH$; $R^3 = OCH_3$; $R^5 = H$)	23.4	NA	NA	NA	NA	NA	26.9	NA
Luteolin ($R^1 = R^2 = R^3 = R^4 = OH$; $R^5 = H$)	46.9	53.3	NA	29.3	NA	NA	21.3	NA
Tricin ($R^1 = R^2 = R^4 = OH$; $R^3 = R^5 = OCH_3$)	35.4	NA	NA	3.2	NA	NA	10.5	NA
Flavonol								
Kaempferol ($R^1 = R^2 = R^4 = OH$; $R^3 = R^5 = H$)	57.5	72.3	8.4	28	NA	27.7	56.4	20.8
Myricetin ($R^1 = R^2 = R^4 = R^3 = R^5 = OH$)	18.3	47.6	9	40.2	19.9	15.2	71.4	NA
Quercetin ($R^1 = R^2 = R^3 = R^4 = OH$; $R^5 = H$)	14.5	41.4	NA	7	16.4	40.4	93	24.4
Isoflavone								
Afromosin ($R^1 = H$; $R^2 = R^4 = OCH_3$; $R^3 = OH$)	37.6	8.2	35.2	NA	NA	NA	10.9	NA
Biochanin A ($R^1 = R^3 = OH$; $R^2 = H$; $R^4 = OCH_3$)	89.2	93.7	87.6	86.5	NA	25[d]	61.2	NA
Daidzein ($R^1 = R^2 = H$; $R^3 = R^4 = OH$)	53	30.3	67.2	17.3	NA	NA	8.8	NA
Formononetin ($R^1 = R^2 = H$; $R^3 = OH$; $R^4 = OCH_3$)	58.7	87.4	81.8	58.6	NA	NA	50	NA
Genistein ($R^1 = R^3 = R^4 = OH$; $R^2 = H$)	89	84.9	90.3	82.7	NA	NA	94	NA
Irisolidone($R^1 = R^3 = OH$; $R^2 = R^4 = OCH_3$)	36.1	7.1	50.7	10	NA	NA	31.5	NA
Isoformononetin ($R^1 = R^2 = H$; $R^3 = OCH_3$; $R^4 = OH$)	NA	37.3	NA	NA	NA	NA	Na	NA
Prunetin ($R^1 = R^4 = OH$; $R^2 = H$; $R^3 = OCH_3$)	38.6	73.1	34.6	NA	NA	NA	NA	NA

Table 1 Activities of (iso)flavonoid glycosyltransferases from *Medicago truncatula*. Results are given as % of substrate converted

Substrate	GT22D	GT22E09	GT29C	GT29H	GT63G	GT67A	GT83F	GT99D
Pterocarpan								
Medicarpin (R^1 = OH; R^2 = OCH_3; R^3 = H)[e]	55.4	100	86.8	50.3	NA	NA	71.8	NA
Maackian (R^1 = OH; R^2 = R^3 = OCH_2O)	NA	63.2	NA	NA	NA	NA	NA	NA

[a] Yield obtained in experiments using 1.25 μg GT83F, 100–250 μM pelargonidin, and 1.0–3.0 h of reaction

[b] Yield obtained in experiments using 8.5 μg GT67A, 100–250 μM isoliquiritigenin, and 0.5–1.0 h of reaction

[c] Low activity (two unknown products) was measured in experiments using 6.0 μg GT83F, 100 μM dihydroquercetin, and 3.0 h of reaction. Assays using GT83F and dihydroquercetin gave non-reproducible results

[d] Yield obtained in experiments using 8.5–15.0 μg GT67, 100–250 μM biochanin A, and 0.5–1.0 h of reaction

[e] Medicarpin was included at 50 μM

Table 2 Kinetic parameters for GT83F (UGT78G1) toward different (iso)flavonoids

Substrate	V_{max} (μmol/min)	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m (s^{-1} M^{-1})
Kaempferol	1.5830	89.8	5.3×10^{-2}	589.5
Quercetin	0.8557	28.7	2.9×10^{-2}	997.0
Pelargonidin	0.2025	52.0	6.8×10^{-3}	130.2
Cyanidin	0.1698	71.8	5.7×10^{-3}	79.1
Apigenin	0.0688	1.5	2.3×10^{-3}	1,534.2
Genistein	1.2050	36.7	4.0×10^{-3}	1,098.0
Daidzein	0.0530	1.6	1.8×10^{-3}	1,107.7
Biochanin A	0.3666	11.7	1.2×10^{-2}	1,047.8
Formononetin	0.0879	1.5	2.9×10^{-3}	1,959.8

* Indicates the position of glycosylation

The position of the aromatic B ring only slightly affects substrate binding/turnover (Table 2). Thus, the K_{cat}/K_m ratio for apigenin (a flavone, **11**) is 1,534 whereas the K_{cat}/K_m ratio is 1,098 for its corresponding isoflavone (genistein, **8**). GT83F glycosylates apigenin (**11**), genistein (**8**), biochanin A (**4**), daidzein (**7**) and formononetin (**5**) at the 7-*O*-position (Table S4). However, the V_{max} of GT83F for the flavonols kaempferol (**19**) and quercetin (**20**), both of which have OH groups at C7 and C3, is much higher when compared with that for most of the other substrates, and the flavonols are glycosylated at the 3-position. However, anthocyanidins (cyanidin and pelargonidin (**22**)), which also possess a 3-OH group, were poor substrates for GT83F.

Three different sugar donors (UDP-glucose, -galactose and -glucuronic acid) were compared in incubations of GT83F with seven different acceptors (Fig. 4). With flavonol (apigenin), isoflavones (genistein, biochanin A and formononetin) and anthocyanidins (pelargonidin), UDPG was the only effective donor. However, significant activity was observed for the flavonols kaempferol and quercetin with either UDPG or UDP-galactose (Fig. 4).

GT83F was also analyzed for its ability to operate in the reverse direction to deglycosylate (iso)flavonoid glucoderivatives. The enzyme could convert biochanin A 7-*O*-glucoside, genistein 7-*O*-glucoside, kaempferol 3-*O*-glucoside, and quercetin 3-*O*-glucoside, but not formononetin 7-*O*-glucoside, into the corresponding aglycone (Fig. 5). Deglycosylation occurred only in the presence of UDP.

Phylogenetic analysis of *Medicago* UGTs

A phylogenetic tree was constructed to show the relationships of all 63 full-length *Medicago* UGTs with the complete Arabidopsis UGT superfamily (Fig. S3). The *Medicago* sequences clustered with the Arabidopsis sequences throughout the cladogram, except for the group of *Medicago* UGTs at the bottom of the tree, which appear to represent members of a clade not present in Arabidopsis (Fig. S3). Fifty nine out of the 63 *M. truncatula* UGTs clustered within almost all known *Arabidopsis* UGT major groups (A through N) as defined by the phylogenetic analysis of Ross et al. (2001). *Arabidopsis* phylogenetic small groups C (UGT90A1, UGT90A2 and UGT90A4) and K (UGT86A1 and UGT86A2) did not have any corresponding *M. truncatula* orthologs. Interestingly, GT83F was the only *M. truncatula* UGT that clustered within group F containing *Arabidopsis* UDP-glucose:flavonol-3-*O*-glycoside-7-*O*-glucosyltransferase UGT78D1 (Jones et al. 2003).

The sequences of the eight functionally characterized *M. truncatula* UGTs were then aligned with those of all known plant UGTs for which acceptor substrates have been identified. Strikingly, *Medicago* UGTs that share common in vitro activities against certain flavonoid or isoflavonoid acceptors were not always closely related phylogenetically (Fig. 6), or on the basis of overall amino acid sequence identity (Table S5). For example, GT67A and GT83F are distantly related, but both are active with the isoflavone biochanin A (**4**), and the flavonols kaempferol (**19**) and quercetin (**20**). However, whereas the isoflavone formononetin (**5**) was the best substrate of those tested with GT83F, GT67A could not glycosylate this compound. GT67A clusters with UGT85A1 (Figs S3, S6), an Arabidopsis UGT that glycosylates cytokinins (preferred substrate trans-zeatin) (Hou et al. 2004), and a *Stevia rebaudiana* UGT that glycosylates diterpenes (Richman et al. 2005), but GT67A did not glycosylate the cytokinins N^6-benzyladenine, dihydrozeatin, *N,N*-dimethyladenine, N^6-isopentenyladenine, kinetin or *trans*-zeatin (data not shown). GT83F clusters with Arabidopsis UGTs that glycosylate anthocyanidins and/or flavonoids, although one of these enzymes uses UDP-rhamnose rather than UDP-glucose, and anthocyanidins are only poor substrates for GT83F in vitro when compared with isoflavones. These observations highlight our current inability to predict plant UGT function from sequence information alone.

Tissue-specific expression patterns of *Medicago* (iso)flavonoid UGTs

The in vitro substrate specificity of a UGT does not necessarily reflect the enzyme's in vivo function (Achnine et al. 2005). To address potential in vivo functions for the eight *Medicago* UGTs with in vitro activity against (iso)flavonoids, the expression patterns of these genes were evaluated by real time quantitative RT-PCR (qRT-PCR) using primers specific for each gene (Fig. 7, Table S6). Tissues examined were stem, flower, root, leaf, and bud, and root-derived cell suspension cultures were also analyzed. Leaves were also wounded and inoculated with spores of the leaf spot pathogen *Phoma medicaginis*, a treatment that induces accumulation of isoflavonoid aglycones (He and Dixon 2000; Deavours and Dixon 2005). Cell suspension cultures were treated with YE, an inducer of isoflavonoid biosynthesis, or MeJA, an inducer of triterpene saponin biosynthesis (Achnine et al. 2005; Suzuki et al. 2005). Transcript levels in elicited cell suspension cultures, and in other plant organs (nodules, petioles, vegetative shoots and seeds), were also measured by Affymetrix DNA array hybridization (see Materials and Methods) (Fig. 8).

Most of the assessed *Medicago* UGTs were expressed in all plant organs examined by qRT-PCR, with the exception

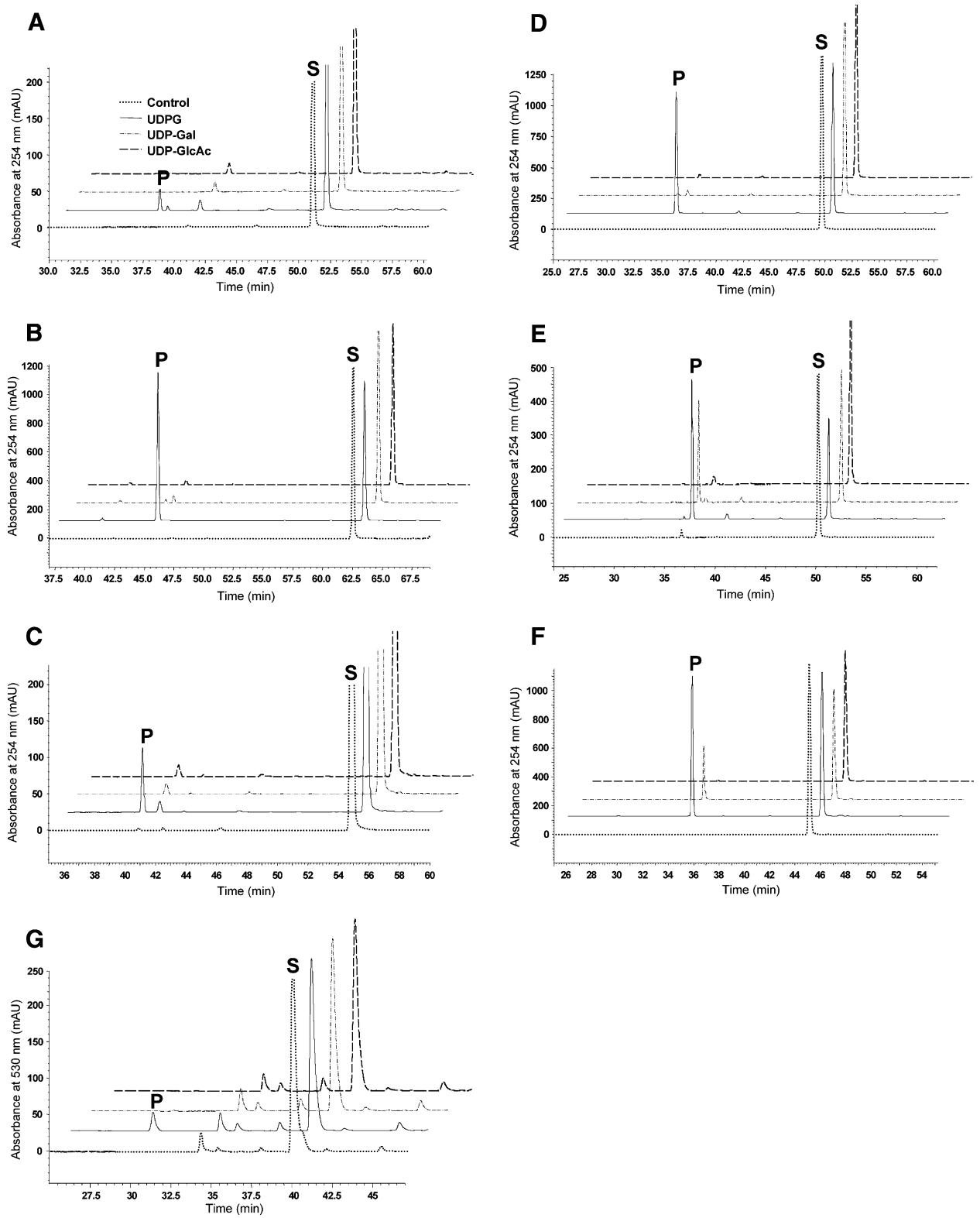


Fig. 4 HPLC analysis of substrates and products from incubations of (iso)flavonoids with GT83F (UGT78G1) in the presence of different sugar donors. Reactions containing acceptor substrate (250 μM), UDP-sugar (2.5 mM), and GT83F (1.25 μg) were incubated for 3 h at 30°C. (**A**) apigenin as acceptor substrate; (**B**) biochanin A; (**C**) formononetin; (**D**) genistein; (**E**) kaempferol; (**F**) quercetin; (**G**) pelargonidin. S, substrate; P, product. Dotted lines, reactions incubated without enzyme; Solid black lines, reactions incubated in the presence of GT83F and UDPG; Dashed-dotted lines, reactions incubated in the presence of GT83F and UDP-galactose; Dashed lines, reactions incubated in the presence of GT83F and UDP-glucuronic acid

of GT22E09, which was expressed in stems, buds and, primarily, roots (Fig. 7). GT22E09 was also expressed in root-derived cell suspension cultures, where its transcript levels were weakly induced by YE and repressed by MeJA (Fig. 7). This expression pattern was confirmed by microarray analysis (Fig. 8).

GT63G exhibited low level constitutive expression in all the tissues analyzed, but was strikingly induced by YE in cell suspension cultures (25 and 32-fold on the basis of Affymetrix and qRT-PCR, respectively, Figs. 7, 8). Its expression was also weakly induced in response to infection of leaves with *P. medicaginis*. GT67A was the only UGT to show no expression in cell suspension cultures, although GT99D transcripts were only detected in cell cultures after MeJA elicitation, with higher transcript levels at 24 h post-elicitation than at 2 h post-elicitation (Fig. 8). GT67A was preferentially expressed in flowers, whereas GT99D was most highly expressed in stems, flowers and buds, and its expression in leaves was strongly reduced following inoculation with *P. medicaginis*.

The remaining UGTs were expressed in all tissues studied by qRT-PCR, but at different relative overall levels. Expression of GT29H was down-regulated following *Phoma* infection of both leaf and stem tissue. Microarray analysis revealed strong expression of GT29H in petioles and pods (Fig. 8). GT29C was most highly expressed in roots, and up-regulated by the wounding process applied as a control for *P. medicaginis* infection in leaves, with further transcript accumulation following fungal inoculation. Consistent with wound induction in leaves, GT29C expression was up-regulated by the wound signal MeJA in cell cultures (Fig. 7, 8). In contrast, GT83F was most strongly expressed in buds, and its expression was strongly down-regulated by wounding of leaves, and also by application of MeJA to cell cultures. The expression level of GT22D was low in cell cultures,

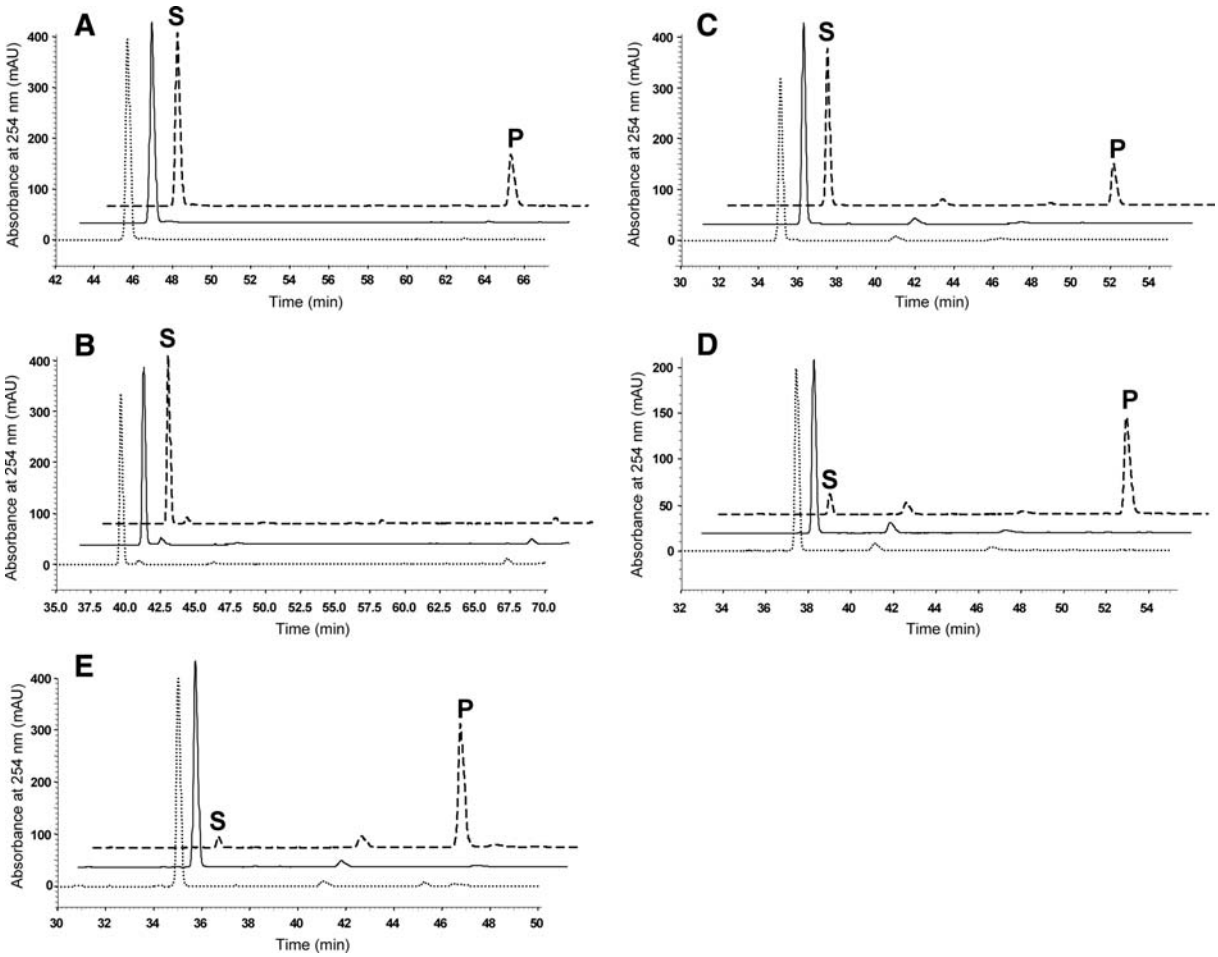


Fig. 5 HPLC analysis of (iso)flavonoid aglycones formed by GT83F (UGT78G1) operating in the reverse direction in the presence of UDP. Substrates are: (**A**) biochanin A 7-*O*-glucoside; (**B**) formononetin 7-*O*-glucoside; (**C**) genistein 7-*O*-glucoside; (**D**) kaempferol 3-*O*-glucoside; and (**E**) quercetin 3-*O*-glucoside. The peak labelled as S is the substrate and peak P is the product. Dotted lines, reactions incubated without enzyme; Solid black lines, reactions incubated in the presence of GT83F; Dashed lines, reactions incubated in the presence of both GT83F and UDP

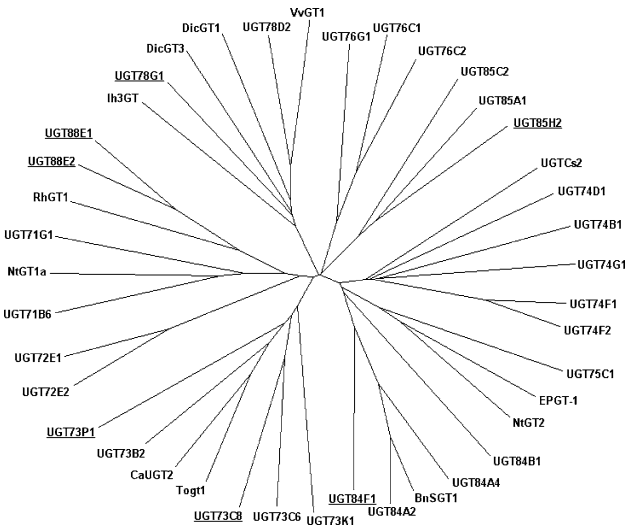


Fig. 6 Unrooted phylogram comparison of the amino acid sequences of *Medicago truncatula* UGTs 83F (UGT78G1), 22D (UGT73C8), 22E09 (UGT88E1), 29C (UGT88E2), 63G (UGT84F1), 67A (UGT85H2), 99D (UGT73P1), 29H (UGT71G1), and 49F (UGT73K1), with those of other functionally characterized plant glycosyltransferases. *Arabidopsis thaliana*: UGT71B6 (substrate: abscisic acid); UGT72E1 and UGT72E2 (coniferyl and sinapyl aldehyde); UGT73B2 (flavonoid 3-*O*); UGT73C6 (flavones/coumarins); UGT74B1 (thiohydroximate); UGT74D1 (jasmonic acid); UGT74F1 and UGT74F2 (anthranilate/salicylic acid); UGT75C1 (anthocyanindin 5-*O*); UGT76C1, UGT85A1, and UGT76C2 (cytokinins), UGT78D2 (anthocyanidins/flavonoids); UGT84A2 and UGT84A4 (sinapate); GT84B1 (indole 3-acetic acid); *Brassica napus:* BnSGT1 (sinapate); *Catharantus roseus*: CaUGT2 (curcumin); *Dianthus caryophyllus*: DicGT1 and DicGT3 (flavonoid 3-*O*); *Crocus sativus*: UGTCs2 (crocetin); *Eucalyptus perriniana*: EpGT-1 (monoterpene); *Iris hollandica*: Ih3GT (anthocyanidins 3-*O*); *Nicotiana tabacum*: NtGT1a, NtGT2, and Togt1 (coumarins and/or flavonols); *Rosa hybrida*: RhGT1 (anthocyanidins); *Stevia rebaudiana*: UGT74G1, UGT76G1, and UGT85C2 (diterpenes); *Vitis vinifera:* VvGT1 (anthocyanins)

where it was down-regulated by YE and up-regulated by MeJA (Figs. 7, 8).

(Iso)flavonoids play important roles in nodulation of legumes (Fisher and Long 1992; Subramanian et al. 2006). Microarray analysis revealed that, of the eight UGTs studied here, GTs 22E09, 22D, 29H and 29C were the most highly expressed in root nodules throughout their development. GT29H and GT83F were the most highly expressed among the eight UGTs in developing seed (Fig. 8).

Discussion

Identification of *Medicago* (iso)flavonoid UGTs

To date, plant glycosyltransferases active with flavonoids or isoflavonoids have been identified by one of two approaches- activity based purification (sometimes followed by molecular cloning) (Durren and McIntosh 1999; Fukuchi-Mizutani et al. 2003), or sequence-based identification followed by in vitro expression and assay (Kramer et al. 2003; Ogata et al. 2004; Achnine et al. 2005). In both cases, actual proof of in vivo function is usually lacking, and likely hard to determine in view of the redundant and promiscuous activities among the many members of the plant UGT superfamily.

The non-biased expression and screening approach utilized in the present work led to the identification of eight *Medicago* UGTs that were able to glycosylate flavonoids and/or isoflavonoids. Each enzyme was effective against more than one class of (iso)flavonoid, but had a distinct pattern of substrate preference. The screening method using pooled acceptor substrates is efficient, but may result in false negatives if one or more compounds inhibit glycosylation of a bona fide substrate. If the frequency of activity with (iso)flavonoids (8 out of 19 clones randomly selected only on the basis of solubility of expressed protein) were to extend to the whole *Medicago* UGT family, well over 50 UGTs would be predicted to show activity with (iso)flavonoids in vitro.

All (iso)flavonoid UGTs described in this study prefer UDP-glucose as sugar donor. However, GT83F is able to transfer galactose, but not glucuronic acid, to kaempferol and quercetin, but not to other substrates such as isoflavonoids. It has been suggested that the preference of plant UGTs for UDPG or UDP-galactose is determined by the nature of the last residue in the PSPG signature motif, and changing this residue from histidine to glutamine in the *Aralia cordata* UDP-galactose: anthocyanin galactosyltransferase switches the preferred donor from UDP-galactose to UDP-glucose (Kubo et al. 2004). Our data, indicating dual specificity of GT83F depending upon the nature of the acceptor, suggest that the molecular basis of donor preference is more complex than suggested by the analysis of the *A. cordata* transferase.

Comparisons of *Medicago* UGTs to previously identified (iso)flavonoid UGTs

UGT73F1 is an isoflavonoid-specific glucosyltransferase from the legume *Glycyrrhiza echinata*. *G. echinata* cell cultures produce high levels of formononetin 7-*O*-glucoside following exposure to yeast elicitor, and UGT73F1, which is induced under these conditions, is most effective against formononetin (**5**) and daidzein (**7**) (Nagashima et al. 2004). However, this enzyme showed weak activity against quercetin and a range of other flavonoids and related phenolic compounds. In contrast, five of the eight *Medicago* enzymes reported here were likewise active with formononetin (**5**) and daidzein (**7**), but also showed significant activity against quercetin (**20**) and/or kaempferol

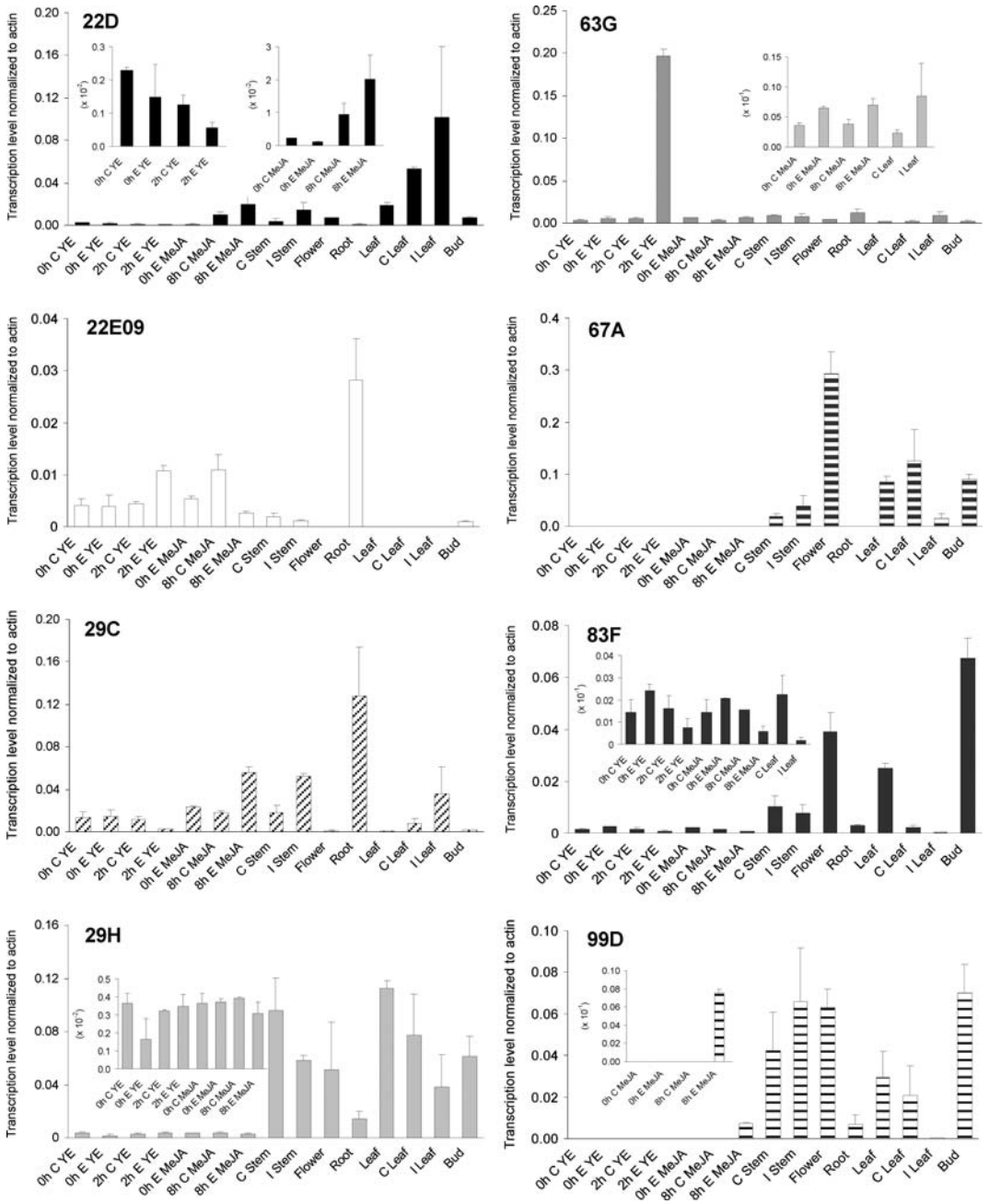


Fig. 7 Transcript levels of (iso)flavonoid UGTs in *M. truncatula* tissues and cell cultures quantified by qRT-PCR. Cell cultures were elicited with yeast elicitor (E) or methyl jasmonate (MeJA) for 2 h and 8 h, respectively. Roots, leaves and buds were harvested from unchallenged plants. Leaves were also wounded (C), and then inoculated (I) with a suspension of *Phoma medicaginis* for 72 h. Stems were sprayed with a solution of 0.1% Tween 20 (C) or inoculated (I) with a spore suspension of *P.medicaginis* in Tween 20. cDNAs obtained from RT-PCR were quantified relative to the abundance of actin. Insets show transcript levels plotted on a more sensitive scale in those cases where they were lower that in other organs in the parent plot

(**19**). They were also active against other isoflavones (afromosin (**1**) and irisolidone (**2**)) the glycosides of which are present in *Medicago* cell cultures (Farag et al. 2007).

Several other UGTs have been reported to possess activity against multiple classes of flavonoid compounds. Thus, UGT73G1 from onion skins glycosylates a range of flavones, flavanones, isoflavones and the chalcone isoliquritigenin (**16**), whereas UGT73J1 from the same source only appeared to be active with one flavone and the isoflavone genistein (**8**) (Kramer et al. 2003). Arabidopsis UGT73B2 glycosylates flavonols (preferred, at the 3-*O*-position), flavones and flavanones, and similar to GT83F,

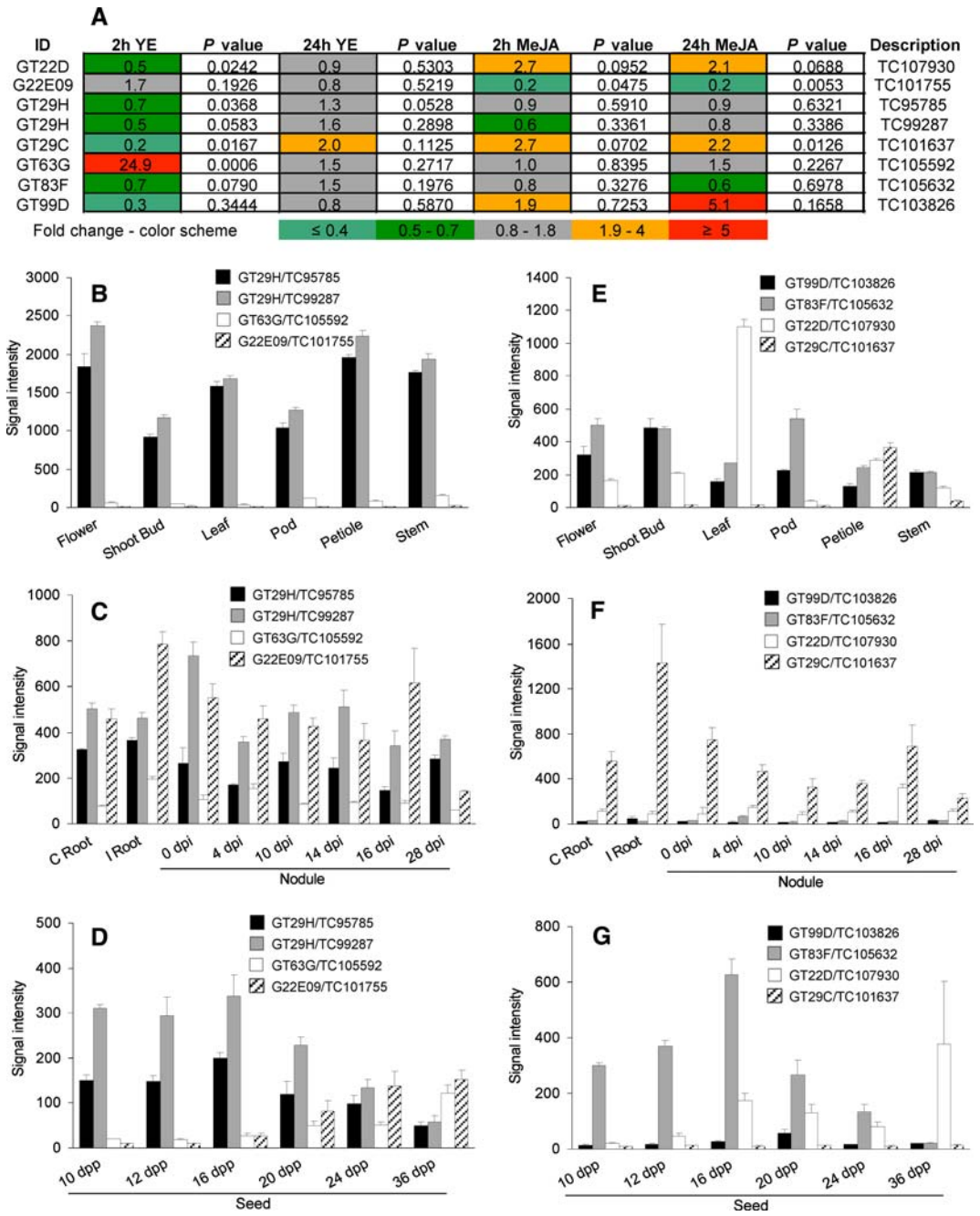


Fig. 8 DNA microarray expression analysis of *Medicago* (iso)flavonoid glycosyltransferase genes. (**A**) transcript changes in cell suspension cultures in response to YE or MeJA. The TC number is from TIGR MtGI v. 8. Numbers represent ratios of treatment versus control samples; (**B**–**G**) transcript levels in different tissues including seeds at various stages of development and nodules derived from *Rizhobium*-inoculated roots. C Root, control roots; I Root, nodule-free inoculated roots; dpi, days post- inoculation; dpp, days post pollination

glycosylates the 7-*O*-position if the 3-*O*-position is not available (Kim et al. 2006). Three out of 18 UGTs cloned from the petals of carnation glycosylated various flavonoids and anthocyanins at the C-ring 3-*O*-position, and two glycosylated chalcone at the 2′-*O*-position, and other flavonoids at various positions (Ogata et al. 2004). Thus, the *Medicago* enzymes reported here appear quite typical as regards their broad substrate preferences in vitro.

Substrate specificity comparisons are more informative if they report K_{cat}/K_m values, but such analyses can be both time consuming and prohibitively expensive when comparing many potential substrates. In the case of GT83F,

K_{cat}/K_m values for quercetin (**20**) and daidzein (**7**) are similar, but the K_m for daidzein is approximately 50-fold less than for quercetin, which has the higher turnover number. These considerations may be important when considering potential in vivo functions of UGTs with overlapping substrate specificities that are co-expressed in the same cell types.

UGT substrate preference and regiospecificity

Many plant secondary metabolite UGTs are more regiospecific than substrate specific (Vogt et al. 1997; Taguchi et al. 2000; Vogt and Jones 2000; Lim et al. 2002; Vogt 2002; Kramer et al. 2003). However, UGTs from carnation petals have been shown to glucosylate some flavonoids on both the 7-*O*- and 3-*O*-positions, and other flavonoids on the 7-*O*-position alone (Ogata et al. 2004). Similarly, GT83F catalyzes glycosylation at the 7-*O*-position of the studied (iso)flavonoids, unless these compounds have a hydroxyl group at C3, in which case this position is substituted. A comparison of substrate structure-activity relationships revealed that the carbonyl group on ring C is not very important for activity of *Medicago* GT83F, since this enzyme was able to catalyze the 3-*O*-glycosylation of anthocyanidins, which lack this group.

Recently, UGTs from *Arabidopsis thaliana* were extensively analyzed both in vitro and in vivo with quercetin as acceptor; from a total of 29 UGTs which exhibited activity with quercetin, 14 catalyzed glycosylation either on the C-ring 3-hydroxyl or the A-ring 7-hydroxyl, the most commonly reported glycosylation positions for quercetin derivatives isolated from plants. Eleven of the UGTs produced two products, glycosylating the 7- and 3-hydroxyls, 7-and 3′-hydroxyls, 7 and 4′-hydroxyls, or 3- and 4′-hydroxyls. The remaining enzymes produced three or four monoglucosides substituted on different hydroxyl groups (Lim et al. 2004). Some of these UGTs also glycosylated additional substrates such as caffeic acid (UGT71C1) (Lim et al. 2003), indole-3-acetic acid (UGT84B1) (Jackson et al. 2001) or salicylic acid (UGT74F1) (Lim et al. 2002). Similarly, GT29H (UGT71G1) glucosylates triterpenes in addition to quercetin and genistein (Achnine et al. 2005).

As seen by a comparison of the kinetic constants of GT83F for apigenin (a flavone) and genistein (the corresponding isoflavone), the position of the aromatic ring B (ie. attached to the 2 or 3 positions of the C ring) in flavonoid derivatives has little effect on regiospecificity or turnover of this enzyme. Thus, the K_{cat}/K_m ratio for apigenin is 1,534 whereas the enzyme has a K_{cat}/K_m ratio of 1,098 for the corresponding isoflavone. It should be noted that the apigenin conjugates found in *Medicago* are, however, primarily glucuronides rather than glycosides (Deavours and Dixon 2005), and the in vivo significance of the activity of GT83F with apigenin is therefore not clear. Although GT83F generally prefers to catalyze the glycosylation of (iso)flavonoids at the 7-*O*-position, the V_{max} of GT83F for kaempferol and quercetin (both of which are glucosylated on the OH group at C3) is much higher when compared with the other substrates, except for genistein. Quercetin glycosides are found in *Medicago* flowers (Bisby et al. 1994). However, anthocyanidin floral pigments (cyanidin and pelargonidin) were comparatively poor substrates for 3-*O*-glucosylation by GT83F (UGT 78G1) in vitro. Arabidopsis UGT78D2, which effectively glycosylates both anthocyanins and flavonols at the 3-*O*-position, has been implicated in anthocyanin biosynthesis in vivo on the basis of genetic analysis (Tohge et al. 2005).

The hydroxyl group at C5 of the A-ring is important for activity of GT67A toward isoflavones, since no activity was found with daidzein or formononetin, the corresponding C5-deoxy derivatives of genistein and biochanin A, respectively. The presence of a methoxyl group (–OCH_3) at C4′ (ring B) of isoflavones improved substrate activity for GT67A; thus, the extent of glycosylation of biochanin A was around 4-fold higher than for genistein. However, introducing steric hindrance around position C7 by substitution at C6 (eg. irisolidone (**2**)) prevented activity. In contrast, both irisolidone and afromosin (with a methoxyl group at C6) were efficiently glucosylated at the 7-*O*-position by GTs 22D and 29C.

GT67A appears to be more regioselective for flavonols with a single hydroxyl group on ring B, since quercetin was converted to two different glucoderivatives whereas a single 3-*O*-glucosylated product was formed with kaempferol. Dihydroquercetin, liquiritigenin or naringenin were not substrates for GT67A, demonstrating the importance of the double bond in ring C for substrate binding or conversion. Unlike GT83F, GT67A did not glycosylate pelargonidin; this could be due to lack of an oxygen function at C4, or to the residual positive charge on the anthocyanidin C-ring structure (flavilium cation). However, presence of C-ring is not essential, since GT67A acted on the chalcone isoliquiritigenin.

Clearly, much is yet to be learned concerning structure-activity relations in the plant UGT superfamily. To date, only three crystal structures are available for plant family 1 UGTs (*Medicago* UGT71G1 (Shao et al. 2005), grape *Vv*GT1 (Offen et al. 2006) and *Medicago* GT67A (UGT85H2) (Li et al. 2007). Molecular docking and mutagenesis studies are beginning to define aspects of the active site pocket that control substrate specificity and regioselectivity (Shao et al. 2005; He et al. 2006), but primary sequence alone is relatively uninformative. Likewise, phylogenetic analysis is not particularly helpful as a predictor of possible function (or at least in vitro catalytic

activity), as seen from the observation that *Medicago* UGTs with good activity against several flavonoids and isoflavonoids cluster within several clades of the phylogenetic tree, often with enzymes from other species identified as having activity against non-flavonoid substrates. However, no independent study has attempted to assay multiple UGTs against all possible substrates, and such an analysis might reveal shared catalytic capabilities among phylogenetically related enzymes from diverse species. Nevertheless, *Medicago* GT67A clusters with several Arabidopsis UGTs that glycosylate cytokinins (preferred substrate *trans*-zeatin), but this enzyme does not glycosylate *trans*-zeatin or any of five other cytokinins tested.

Reversibility of the UGT reaction

Glaser and Brown described the reversibility of the native chitin synthetase reaction in one of the first reports of in vitro GT activity (1957), and it was subsequently shown that the reaction catalyzed by the parsley UDPG: flavonol 3-*O*-glucosyltransferase was freely reversible (Sutter and Grisebach 1975). Recent work has shown that reactions catalyzed by both previously uncharacterized prokaryotic GTs (CalG1 and CalG4) and well-studied GTs (GtfD and GtfE) are also reversible (Zhang et al. 2006). However, to the best of our knowledge, the reversibility of plant UGT reactions has not been investigated other than for the parsley flavonol UGT. Our results show that at least one other plant UGT reaction is also reversible, at least in vitro. The physiological significance of this finding is not clear, particularly since plants possess a diverse family of β-glucosidases, specialized for the de-glycosylation of sugar conjugates to release active aglycones upon physiological need. In the case of *M. truncatula*, 45 TCs and over 30 singletons representing β-glucosidases can be found in the TIGR database (http://www.compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb = medicago) with members expressed in almost all plant tissues and under various stress conditions. The reversibility of the glycosyltransfer reaction may have important practical applications in vitro for making valuable natural products and for processes such as the creation of rare NDP-sugars, moving sugars from one scaffold to another, or substituting one sugar on a core scaffold for another.

Predicting UGT function based on tissue-specific expression

It has occasionally been possible to determine the in vivo biochemical function of a plant UGT by analysis of mutants or transgenic lines in which the enzyme's expression has been eliminated or reduced (Jones et al. 2003; Grubb et al. 2004). However, in many cases this approach does not work, possibly as a result of the redundancy in substrate preference among UGT family members (Bowles et al. 2006) as clearly seen in the present work. Knowing which tissue(s) a UGT is expressed in can provide clues as to function if the metabolite composition of that tissue and the in vitro substrate preferences of the enzyme are known. Recently, a combination of transcriptomic and metabolomic analysis has been used to predict UGTs involved in anthocyanin glycosylation in Arabidopsis (Tohge et al. 2005).

Glucoconjugates of afromosin (**1**), biochanin A (**4**), isoliquiritigenin (**16**), formononetin (**5**), and medicarpin (**12**), as well as conjugates of liquiritigenin (**17**) and tricin (**15**), have been detected in roots of *M. truncatula* (Harrison and Dixon 1993; Farag et al. 2007). Among the UGTs described here, GT22E09 and GT29C were the most highly expressed in roots, and were active in vitro with formononetin, medicarpin, afromosin, biochanin A, and isoliquiritigenin (**16**), but not liquitiritigenin (**17**) or tricin (**15**). Thus, GT22E09 and GT29C could function in the glycosylation of endogenous isoflavones in *Medicago* roots. Such a role in isoflavone metabolism is less clear from the results of expression in cell suspension cultures, where GT22E09 was induced by YE and repressed by MeJA treatment, whereas GT29C was repressed by YE and induced by MeJA.

M. trucatula flowers accumulate flavonols including quercetin (**20**), myricetin (**21**), and kaempferol (**19**) (Bisby et al. 1994). GT67A and GT99D are highly expressed in flower tissues, and the corresponding proteins have activity toward flavonols. In the case of GT67A, kaempferol is preferred over isoflavone (biochanin A) or chalcone (isoliquiritigenin) (Li et al. 2007). Thus, GT67A and GT99D could potentially function in the biosynthesis of floral flavonol glycosides.

GT63G gene expression was markedly up-regulated in cell suspension cultures within 2 h of exposure to YE. This response is associated with accumulation of isoflavone aglycones and glycosides (Suzuki et al. 2005; M. Farag and L.W. Sumner unpublished results); however, GT63G showed no activity against isoflavones, being weakly active against flavones, glycosides of which were not detected in the cultures. The in vivo substrate of GT63G can not therefore be predicted at present.

It has been shown that GT29H (UGT71G1) is induced during the later stages of the response to MeJA in *Medicago* cell cultures (at times later than those analyzed here), and that the enzyme has weak, but probably physiologically significant, activity for the glycosylation of MeJA-induced triterpenes in spite of its preference for flavonoids in vitro (Achnine et al. 2005). This enzyme is also active with isoflavonoids, and its down-regulation in fungally

infected leaves may represent a mechanism to ensure accumulation of the antimicrobial isoflavonoid aglycones coumestrol (**10**), afromosin (**1**), and medicarpin (**12**), rather than their less active glycosides, at or around the sites of infection, in view of the high constitutive expression level of this gene in leaves.

GT99D, which is weakly active with flavonols but inactive with isoflavones, was most highly expressed in stems, flowers and buds, and its expression in leaves was strongly reduced following inoculation with *P. medicaginis*. *GT99D* transcripts were induced by MeJA in cell cultures, suggesting the possibility of involvement in triterpene saponin biosynthesis (Achnine et al. 2005).

Further functional characterization of the present group of *Medicago* (iso)flavonoid UGTs will require genetic analysis using either transgenics (antisense or RNAi lines) or insertion/deletion mutants that are now available (Tadege et al. 2005; Wang et al. 2006), coupled with detailed metabolite measurements in the specific organs in which the UGTs are expressed. The current metabolite data on which the above tentative functional assignments have been based are more qualitative than quantitative, and in vivo activity will likely depend on the relative concentrations of potential acceptors within a particular tissue.

Acknowledgements We thank Drs Kiran Mysore and Rui Zhou for critical reading of the manuscript, Cuc Ly for assistance with artwork, Dr Michael Udvardi for sharing Affymetrix results for tissue-specific expression of *Medicago* UGTs, Klementina Kakar for providing primers for ubiquitin and actin, Dr Vagner A. Benedito for assistance with qRT-PCR, and Dr Yuhong Tang for assistance with microarray analysis. This work was supported by the National Science Foundation Molecular Biochemistry Program (Grant # 0416833 to XW and RAD) and by the Samuel Roberts Noble Foundation.

References

Achnine L, Huhman DV, Farag MA, Sumner LW, Blount JW, Dixon RA (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. Plant J 41:875–887

Aoki T, Akashi T, Ayabe S (2000) Flavonoids of leguminous plants: structure, biological activity, and biosynthesis. J Plant Res 113:475–488

Bingham SA, Atkinson C, Liggins J, Bluck L, Coward A (1998) Phyto-oestrogens: where are we now? Br J Nutr 79:393–406

Bisby FA, Buckingham J, Harborne JB (1994) Plants and their constituents, vol 1. Chapman and Hall, New York

Borek C (2006) Aging and antioxidants. Fruits and vegetables are powerful armor. Adv Nurse Pract 14:35–38

Bowles D, Lim E, Poppenberger B, Vaistij F (2006) Glycosyltransferases of lipophilic small molecules. Annu Rev Plant Biol 57:567–597

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol 139:5–17

Dangl JL, Jones JDG (2001) Plant pathogens and integrated defence responses to infection. Nature 411:826–833

Day AJ, Bao YP, Morgan MRA, Williamson G (2000) Conjugation position of quercetin glucuronides and effect on biological activity. Free Radical Biol Med 29:1234–1243

Day AJ, Gee JM, DuPont MS, Johnson IT, Williamson G (2003) Absorption of quercetin-3-glucoside and quercetin-4′-glucoside in the rat small intestine: the role of lactase phlorizin hydrolase and the sodium-dependent glucose transporter. Biochem Pharmacol 65:1199–1206

Deavours BE, Dixon RA (2005) Metabolic engineering of isoflavonoid biosynthesis in alfalfa (*Medicago sativa* L.). Plant Physiol 138:2245–2259

Deavours BE, Liu C-J, Naoumkina M, Tang Y, Farag MA, Sumner LW, Noel JP, Dixon RA (2006) Functional analysis of members of the isoflavone- and isoflavanone *O*-methyltransferase enzyme families from the model legume *Medicago truncatula*. Plant Mol Biol 62:715–733

Dixon RA (2004) Phytoestrogens. Annu Rev Plant Biol 55:225–261

Dixon RA, Sumner LW (2003) Legume natural products. Understanding and manipulating complex pathways for human and animal health. Plant Physiol 131:878–885

Durren RL, McIntosh CA (1999) Flavanone-7-*O*-glucosyltransferase activity from *Petunia hybrida*. Phytochemistry 52:793–798

Farag M, Huhman DV, Lei Z, Sumner LW (2007) Metabolic profiling and systematic identification of flavonoids and isoflavonoids in *Medicago truncatula* cultured cells and roots using HPLC-UV-ESI-MS and GC-MS. Phytochemistry 68:342–354

Farmakalidis E, Murphy PA (1984) Oestrogenic response of the CD-1 mouse to the soya-bean isoflavones genistein, genistin and daidzin. Food Chem Toxicol 22:237–239

Fisher RF, Long SR (1992) *Rhizobium*-plant signal exchange. Nature 387:655–660

Forkmann G, Martens S (2001) Metabolic engineering and applications of flavonoids. Curr Opin Biotechnol 12:155–160

Formica JV, Regelson W (1995) Review of the biology of quercetin and related bioflavonoids. Food Chem Toxicol 33:1061–1080

Fukuchi-Mizutani M, Okuhara H, Fukui Y, Nakao M, Katsumoto Y, Yonekura-Sakakibara K, Kusumi T, Hase T, Tanaka Y (2003) Biochemical and molecular characterization of a novel UDP-glucose: anthocyanin 3′-*O*-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. Plant Physiol 132:1652–1663

Gil B, Sanz MJ, Terencio MC, Ferrandiz ML, Bustos G, Paya M, Gunasegaran R, Alcaraz MJ (1994) Effects of flavonoids on Naja naja and human recombinant synovial phospholipases A2 and inflammatory responses in mice. Life Sci 54:333–338

Glaser L, Brown DH (1957) The synthesis of chitin in cell-free extracts of *Neurospora crassa*. J Biol Chem 228:729–742

Grubb CD, Zipp BJ, Ludwig-Müller J, Masuno MN, Molinski TF, Abel S (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. Plant J 40:893–908

Harrison MJ, Dixon RA (1993) Isoflavonoid accumulation and expression of defense gene transcripts during the establishment of vesicular arbuscular mycorrhizal associations in roots of *Medicago truncatula*. Mol Plant-Microbe Interact 6:643–654

He X-Z, Dixon RA (2000) Genetic manipulation of isoflavone 7-*O*-methyltransferase enhances the biosynthesis of 4′-*O*-methylated isoflavonoid phytoalexins and disease resistance in alfalfa. Plant Cell 12:1689–1702

He X, Wang X, Dixon RA (2006) Mutational analysis of the *Medicago* glycosyltransferase UGT71G1 reveals residues that control regio-selectivity for (iso)flavonoid glycosylation. J Biol Chem 281:34441–34447

Hou B, Lim EK, Higgins GS, Bowles DJ (2004) N-glucosylation of cytokinins by glycosyltransferases of *Arabidopsis thaliana*. J Biol Chem 279:47822–47832

Hughes J, Hughes MA (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. J Seq Map 5:41–49

Huhman DV, Sumner LW (2002) Metabolic profiling of saponin glycosides in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospary ion-trap mass spectrometer. Phytochemistry 59:347–360

Jackson RG, Lim EK, Li Y, Kowalczyk M, Sandberg G, Hoggett J, Ashford DA, Bowles DJ (2001) Identification and biochemical characterization of an *Arabidopsis* indole-3-acetic acid glucosyltransferase. J Biol Chem 276:4350–4356

Jones P, Messner B, Nakajima JI, Schäffner AR, Saito K (2003) UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*. J Biol Chem 278:43910–43918

Kanadaswami C, Lee LT, Lee PP, Hwang JJ, Ke FC, Huang YT, Lee MT (2005) The antitumor activities of flavonoids. In vivo 19:895–909

Kape R, Parniske M, Werner D (1991) Chemotaxis and *nod* gene activity of *Bradyrhizobium japonicum* in response to hydroxycinnamic acids and isoflavonoids. Appl Environ Microbiol 57:316–319

Kaul TN, Middleton E Jr., Ogra PL (1985) Antiviral effect of flavonoids on human viruses. J Med Virol 15:71–79

Kessmann H, Edwards R, Geno P, Dixon RA (1990) Stress responses in alfalfa (*Medicago sativa* L.) V. Constitutive and elicitor-induced accumulation of isoflavonoid conjugates in cell suspension cultures. Plant Physiol 94:227–232

Kim JH, Kim BG, Ko JH, Lee Y, Hur HG, Lim Y, Ahn JH (2006) Molecular cloning, expression, and characterization of a flavonoid glycosyltransferase from *Arabidopsis thaliana*. Plant Sci 170:897–903

Kramer CM, Prata RTN, Willits MG, De Luca V, Steffens JC, Graser G (2003) Cloning and regiospecificity studies of two flavonoid glucosyltransferases from *Allium cepa*. Phytochemistry 64:1069–1076

Kren V, Martinkova L (2001) Glycosides in medicine: the role of glycosidic residue in biological activity. Curr Med Chem 8:1303–1328

Kubo A, Arai Y, Nagashima S, Yoshikawa T (2004) Alteration of sugar donor specificities of plant glycosyltransferases by a single point mutation. Arch Biochem Biophys 429:198–203

Li L, Modolo LV, Achnine L, Dixon RA, Wang X (2007) Structure of UGT85H2, an (iso)flavonoid uridine diphosphate glycosyltransferase from the model legume *Medicago truncatula*, in review

Lim EK, Ashford DA, Hou B, Jackson RG, Bowles DJ (2004) Arabidopsis glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. Biotech Bioeng 87:623–631

Lim EK, Doucet CJ, Li Y, Elias L, Worrall D, Spencer SP, Ross J, Bowles DJ (2002) The activity of *Arabidopsis* glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. J Biol Chem 277:586–592

Lim EK, Higgins GS, Li Y, Bowles DJ (2003) Regioselectivity of glucosylation of caffeic acid by a UDP-glucose: glucosyltransferase maintained in planta. Biochem J 373:987–992

Manach C, Scalbert A, Morand C, Remesy C, Jimenez L (2004) Polyphenols: food sources and bioavailability. Am J Clin Nutr 79:727–747

Nagashima S, Inagaki R, Kubo A, Hirotani M, Yoshikawa T (2004) cDNA cloning and expression of isoflavonoid-specific glucosyltransferase from *Glycyrrhiza echinata* cell-suspension cultures. Planta 218:456–459

Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. EMBO J 25:1396–1405

Ogata J, Itoh Y, Ishida M, Yoshida H, Ozeki Y (2004) Cloning and heterologous expression of cDNAs encoding flavonoid glucosyltransferases from *Dianthus caryophyllus*. Plant Biotechnol 21:367–375

Paquette S, Møller BL, Bak S (2003) On the origin of family 1 plant glycosyltransferases. Phytochemistry 62:399–413

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

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 13:62–66

Richman A, Swanson A, Humphrey T, Chapman R, McGarvey BD, Pocs R, Brandle J (2005) Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of *Stevia rebaudiana*. Plant J 41:56–67

Ross JR, Li Y, Lim EK, Bowles DJ (2001) Higher plant glycosyltransferases. Genome Biol 2:3004.3001–3004.3006

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X (2005) The structure of UGT71G1, a multifunctional triterpene/flavonoid uridine diphosphate glycosyltransferase from the model legume *Medicago truncatula*. Plant Cell 17:3141–3154

Shirley BW (1998) Flavonoids in seeds and grains: physiological function, agronomic importance and the genetics of biosynthesis. Seed Sci Res 8:415–422

Stafford HA (1997) Roles of flavonoids in symbiotic and defense functions in legume roots. Bot Rev 63:27–39

Subramanian S, Stacey G, Yu O (2006) Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. Plant J 48:261–273

Sutter A, Grisebach H (1975) Free reversibility of the UDP-glucose: flavonol 3-*O*-glucosyltransferase reaction. Arch Biochem Biophys 167:444–447

Suzuki H, Reddy MSS, Naoumkina M, Aziz N, May GD, Huhman DV, Sumner LW, Blount JW, Mendes P, Dixon RA (2005) Methyl jasmonate and yeast elicitor induce differential genetic and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. Planta 220:698–707

Tadege M, Ratet P, Mysore KS (2005) Insertional mutagenesis: a Swiss Army Knife for functional genomics of *Medicago truncatula*. Trends Plant Sci 10:229–235

Taguchi G, Imura H, Maeda Y, Kodaira R, Hayashida N, Shimosaka M, Okazaki M (2000) Purification and characterization of UDP-glucose: hydroxycoumarin 7-*O*-glucosyltransferase, with broad substrate specificity from tobacco cultured cells. Plant Sci 157:105–112

Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J-i, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki M, Saito K (2005) Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing a MYB transcription factor. Plant J 42:218–235

VanEtten HD, Mattews PS, Mercer EH (1983) (+)-Maackiain and (+)-medicarpin as phytoalexins in *Sophora japonica* and identification of the (–) isomers by biotransformation. Phytochemistry 22:2291–2295

Vogt T (2002) Substrate specificity and sequence analysis define a polyphyletic origin of betanidin 5- and 6-*O*-glucosyltransferase from *Dorotheanthus bellidiformis*. Planta 214:492–495

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

Vogt T, Zimmermann E, Grimm R, Meyer M, Strack D (1997) Are the characteristics of betanidin glucosyltransferases from cell-suspension cultures of *Dorotheanthus bellidiformis* indicative of their phylogenetic relationship with flavonoid glucosyltransferases? Planta 203:349–361

Wang H, Li G, Chen R (2006) Fast neutron bombardment (FNB) induced deletion mutagenesis for forward and reverse genetic studies in plants. In: da Silva JT (ed) Floriculture, ornamental and plant biotechnology; advances and topical issues, vol 1. Global Science Books, pp 629–639

Zhang C, Griffith BR, Fu Q, Albermann C, Fu X, Lee I-K, Li L, Thorson JS (2006) Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. Science 313:1291–1294