Characterization of a Glucosyltransferase Enzyme Involved in the Formation of Kaempferol and Quercetin Sophorosides in *Crocus sativus*1[C][W]

Almudena Trapero, Oussama Ahrazem, Angela Rubio-Moraga, Maria Luisa Jimeno, Maria Dolores Gómez, and Lourdes Gómez-Gómez*

Instituto Botánico, Departamento de Ciencia y Tecnología Agroforestal y Genética, Facultad de Farmacia, Universidad de Castilla-La Mancha, Campus Universitario s/n, 02071 Albacete, Spain (A.T., O.A., A.R.-M., L.G.-G.); Fundación Parque Científico y Tecnológico de Albacete, Campus Universitario s/n, 02071 Albacete, Spain (O.A.); Centro Química Orgánica “Lora Tamayo,” Consejo Superior de Investigaciones Científicas, Madrid 28006, Spain (M.L.J.); and Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain (M.D.G.)

UGT707B1 is a new glucosyltransferase isolated from saffron (*Crocus sativus*) that localizes to the cytoplasm and the nucleus of stigma and tepal cells. UGT707B1 transcripts were detected in the stigma tissue of all the *Crocus* species analyzed, but expression analysis of UGT707B1 in tepals revealed its absence in certain species. The analysis of the glucosylated flavonoids present in *Crocus* tepals reveals the presence of two major flavonoid compounds in saffron: kaempferol-3-O-β-D-glucopyranosyl-(1-2)-β-D-glucopyranosyl and quercetin-3-O-β-D-glucopyranosyl-(1-2)-β-D-glucopyranoside, both of which were absent from the tepals of those *Crocus* species that did not express UGT707B1. Transgenic Arabidopsis (*Arabidopsis thaliana*) plants constitutively expressing UGT707B1 under the control of the cauliflower mosaic virus 35S promoter have been constructed and their phenotype analyzed. The transgenic lines displayed a number of changes that resembled those described previously in lines where flavonoid levels had been altered. The plants showed hyponastic leaves, a reduced number of trichomes, thicker stems, and flowering delay. Levels of flavonoids measured in extracts of the transgenic plants showed changes in the composition of flavonoids when compared with wild-type plants. The major differences were observed in the extracts from stems and flowers, with an increase in 3-sophoroside flavonol glucosides. Furthermore, a new compound not detected in ecotype Columbia wild-type plants was detected in all the tissues and identified as kaempferol-3-O-sophoroside-7-O-rhamnoside. These data reveal the involvement of UGT707B1 in the biosynthesis of flavonol-3-O-sophorosides and how significant changes in flavonoid homeostasis can be caused by the overproduction of a flavonoid-conjugating enzyme.

Flavonoids encompass a very large family of phenolic compounds, with more than 9,000 of these known (Harborne and Williams, 2000; Anderson and Markham, 2006; Buer et al., 2010), including anthocyanin, proanthocyanidin, and phlobaphene pigments, as well as the flavonol, flavone, and isoflavone subfamilies that possess specific biological functions in the species in which they accumulate (Grotewold, 2005; Taylor and Grotewold, 2005; Lepiniec et al., 2006; Subramanian et al., 2007). In plants, flavonoids have many diverse functions, including defense, UV protection, auxin transport inhibition, allelopathy, plant-microorganism communication, regulation of reactive oxygen species, and flower coloring to attract pollinators, while in many species they are required for pollen viability (Buer et al., 2010). A number of biological processes, such as transcriptional regulation, signal transduction, and cell-to-cell communication, are also influenced by flavonoids. In addition, flavonols play a crucial role in human nutrition and have pharmaceutical potential, as the intake of flavonols from vegetables, fruit, berries, and beverages has been favorably linked with reduced risks of a number of diseases in large population studies (Harborne and Williams, 2000; Ross and Kasum, 2002; Bischoff, 2008; Cazarolli et al., 2008).

Flavonoids are synthesized in plants via the flavonoid branch of the phenylpropanoid and acetate-malonate metabolic pathway. Their biosynthesis has been investigated extensively, with nearly all the enzymes isolated and functionally characterized (Anderson and Markham, 2006; Tanaka and Filippa, 2006). However, the pathways for sequential modification, such as...
glycosylation, acylation, and methylation, are still relatively unexplored, even though modification produces a huge chemical diversity and is essential for the stable accumulation of flavonoids. Glycosylation renders them more water soluble and less toxic, and it may enable flavonol transport and compartmentation (Bowles et al., 2006). Glycosylation is achieved by UDP-carbohydrate-dependent glycosyltransferases (UGTs) that catalyze the transfer of a carbohydrate from an activated donor sugar onto small molecule acceptors by the formation of a glycosidic bond (Li et al., 2001). These enzymes exhibit a rather broad acceptor tolerance but have been shown to display strict regioselectivity in many cases (Vogt and Jones, 2000). Moreover, the determination of the sugar donor specificity of UGTs appears to have occurred in a lineage-specific manner after the establishment of general regiospecificity for the glycosyl acceptor (Yonekura-Sakakibara et al., 2007, 2012). UGTs also catalyze the attachment of additional sugars to flavonoid glycosides. With respect to flavonoid glycosides, complementary DNAs (cDNAs) encoding anthocyanidin 3-O-glucosyltransferase from *Ipomoea nil* (Morita et al., 2005) and *Veronica persica* (Ono et al., 2010), anthocyanidin 3-O-glucoside:2′′-glucuronolyltransferase from *Bellis perennis* (Sawada et al., 2005), flavanone 7-O-glucoside:2′-rhamnosyltransferase (yielding 7-O-neohesperidoside) from *Citrus maxima* (Frydman et al., 2004), flavonoid 3-O-glucuronyltransferase from *Caltharanthus roseus* (Masada et al., 2009), cyanidin 3-O-galactoside:2′′-xyloxyllitransferase from *Actinidia chinensis* (Montefiori et al., 2011), and anthocyanin 3-O-glucoside:2′′-O-xylosyltransferase (Yonekura-Sakakibara et al., 2012) have been cloned and functionally characterized. Although various flavonol 1,2′-sophorosides have been isolated from various sources of higher plants (Calderón-Montaño et al., 2011) none have been reported for flavonol glucoside 1,2′′-glycosyltransferases.

Saffron (*Crocus sativus*) is one of the most valuable crops worldwide due to its stigmas, widely used as spice, medicinal drugs, and food additives. These stigmas are characterized by the presence of apocarotenoids and flavonoids. Their antioxidant properties, along with their bitter taste, could qualify them as potential organoleptic agents of the spice (Straubinger et al., 2001), and anthocyanin 3-O-glucoside:2′′-O-xylosyltransferase (Yonekura-Sakakibara et al., 2012) have been cloned and functionally characterized. Although various flavonol 1,2′-sophorosides have been isolated from various sources of higher plants (Calderón-Montaño et al., 2011) none have been reported for flavonol glucoside 1,2′′-glycosyltransferases.

**RESULTS**

**Cloning and Analysis of UGT707B1**

To identify glycosyltransferases from saffron stigmas, a homology-based strategy was used, taking advantage of specific glycosyltransferase motifs located in the C-terminal region (Hughes and Hughes, 1994). A cDNA population was prepared by reverse transcription (RT) of poly(A)′ from total RNA isolated from saffron stigmas. DNA fragments were amplified by degenerate primers, and the products obtained were cloned and analyzed. Sequencing of one PCR product revealed homology to glycosyltransferases. The sequence information from this clone allowed the design of PCR-specific primers to obtain the full-length transcripts. We performed 5′ and 3′ RACE using poly(A)′ from saffron stigmas as a template. The gene obtained (1,576 bp; GenBank accession no. HE793682) was intronless, containing a putative open reading frame of 1,413 bp encoding 471 amino acid residues with a calculated molecular mass of 52 kD. The amino acid sequence was named according to the UGT Nomenclature Committee as UGT707B1 (Mackenzie et al., 1997).

Because saffron is a triploid, we employed in silico screening of a large stigma cDNA EST database (http://www.saffrongenes.org/) as an effective method for the identification of potential UGT707B1 alleles. However, no EST clones were identified.

The C terminus of the protein contained the plant secondary product glycosyltransferase (PSPG) box and Kokkalou, 2008). With regard to flavonoid aglycones in tepals, kaempferol dominates, as kaempferol glucosides constitute between 70% and 90% of the total contents of flavonoids; quercetin glucosides vary from 5% to 10%; and glucosides of dihydrokaempferol, isorhamnetin, naringenin, taxifolin, tamarixetin, and myricetin are only minor components (Norbaek et al., 2002).

In order to broaden our understanding on the enzymes responsible for flavonol glycosylation in saffron, we focused attention on the correlation of the accumulation of specific flavonol glycoside end products with the expression of *UGT707B1* in the flower tissues from different *Crocus* species. In addition, to determine the activity of *UGT707B1* in plants, we developed transgenic Arabidopsis (*Arabidopsis thaliana*) lines overexpressing *UGT707B1* and then carried out liquid chromatography-mass spectrometry (MS)-based metabolite profiling to allow for a more comprehensive identification of flavonol metabolites in the transgenic lines. Finally, integration of comprehensive flavonol identification/annotation from the transgenic Arabidopsis lines and *Crocus* species with *UGT707B1* expression analysis identified *UGT707B1* as a gene encoding a flavonol 2′′-O-glycosyltransferase responsible for the formation of flavonoid sophorosides, which interfere with specific developmental plant processes.
signature motif. Analysis of the UGT707B1 sequence for N-terminal targeting signal or C-terminal membrane anchor signal using the SignalP and TargetP Web-based programs predicted a signal peptide localized between amino acids 31 and 32. Prediction of subcellular location with PSORT yielded a chloroplast location. By contrast, the TargetP version 1.1 program predicts the presence of a signal peptide, although not a chloroplast transit peptide, and was predicted to be in the secretion pathway. Furthermore, due to the presence of a signal peptide, UGT707B1 was also analyzed for the presence of glycosylated residues using the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/). This program predicted the presence of two putative glycosylation sites in two Asn residues: Asn-34 (NISI) and Asn-192 (NYTW). Finally, the location of UGT707B1 in flower tissues was investigated using polyclonal antibodies (Fig. 1). The protein seems to be localized preferentially in the nuclei but also was present in the cytosol of tepals, stigmas, and style cells (Fig. 1).

For comparative modeling, UGT707B1 was aligned with MtUGT71G1 and UGT72B1, whose crystal structures have been solved (He et al., 2006; Brazier-Hicks et al., 2007), using the Phyre2 server (http://www.sbg.bio.ic.ac.uk/phyre2/). UGT707B1 displayed 39% and 38% overall identity with MtUGT71G1 and UGT72B1, respectively (Fig. 2A). UGT707B1 had a typical PSPG box sequence including highly conserved key residues for substrate recognition and the catalysis of UGTs: a His residue, His-15 (His-22 in UGT71G1 and His-20 in UGT72B1), that is highly conserved among PSPGs is proposed to act as a key catalytic residue that activates the hydroxy group of the glucosyl acceptor molecule to facilitate glucosidic linkage formation. A well-conserved Asp residue, Asp-116 (Asp-121 in UGT71G1 and Asp-119 in UGT72B1), is hydrogen bonded with the His residue and is proposed to assist in its general acid/base role during catalysis. The residues Glu-374, Glu-390, and Gln-391 that were proposed in previous crystallographic and biochemical studies as important for the catalytic activity (Shao et al., 2005; Offen et al., 2006; Noguchi et al., 2007) are also conserved in UGT707B1(Glu-365, Glu-381, and Gln-382).

Phylogenetic Characterization of UGT707B1

The genomic UGT707B1 gene had no introns, like the other related sequences present in the orthologous group OG7 (Yonekura-Sakakibara and Hanada, 2011). This OG7 group contains multiple UGT families with different functions, including abscisic acid glucosyltransferase (Priest et al., 2006) and monolignol 4-O-glucosyltransferase (Lanot et al., 2006). As shown in a phylogenetic tree (Fig. 2B), UGT707B1 showed the highest amino acid sequence identity to *Sorghum bicolor* XP002460739 (53%), *Brachypodium distachyon* XP_003583056 (53%), rice (*Oryza sativa*) UGT707A3 (52%) and NP_001059726.1 (51%), maize (*Zea mays*) NP00114865, and *Phyllostachys edulis* FP092236.1 (50%). Interestingly, all the UGT707B1-related sequences shown had conserved the glycosylation site NYTW (Supplemental Fig. S1), while the TargetP version 1.1 program predicted the presence of signal peptides in all these sequences. In this monocotyledonous group, only UGT707A3 has been characterized in vitro and acts as a flavonoid-3-O-glucosyltransferase (Ko et al., 2008). When UGT707B1 is compared with the most related sequences from dicotyledonous plants, the highest identity is found with grapevine (*Vitis vinifera*) CAN77507.1 (43%) and poplar (*Populus trichocarpa*) XP002305022 (41%), which are present in the UGT88 group (Fig. 2B). UGT88 is considered to be a functional
Figure 2. The isolated glycosyltransferase from saffron stigmas belongs to a new glycosyltransferase family. A, Amino acid sequence alignment of UGT707B1 against UGT71G1 and UGT72B1. The alignment was performed guided by the conservation of secondary structure predicted for UGT707B1 and observed from the solved crystal structures of UGT71G1 and UGT72B1. α-Helices are highlighted in blue and β-strands in pink. The PSPG motif that interacts with the sugar donor, UDP-Glc, is underlined. B, Nonrooted molecular phylogenetic tree of plant glycosyltransferases. The tree was constructed as described in “Materials and Methods.” The GenBank accession numbers not shown in the tree are given here in parentheses: UGT71A8 (BAF96581), UGT71A9 (BAF96582), UGT71A10 (BAF96583), UGT71A14 (BAG80554), UGT71B6 (AE76551), UGT71F3 (BAF75886), DbB6GT (AAL57240), UGT71G1 (AAW56092), UGT71C1 (ACE08300), UGT707B1 (HE793682), UGT707A3 (BAC83989), UGT88A4 (ABL85471), UGT88A7 (BAG31949), UGT88A9 (ACB56924), GmIF7GT (NP001235161), RsAS (Q4R1I9), UGT8881 (AAO06919), SbB7GT (Q76HR7), AmC4GT (BAA48239), LvC4GT (BAE48240), UGT89C1 (Q89C1), Atf3G7GT (NP181217), Gelf7GT (BAC78438), GT14A05 (ADV71369), Atf7GT (NP567955), DbB5GT (CAB56231), DcC2GT (BAD52006), SbF7GT (BAA83484), FfF3GT (AAD21086), PfF3GT (BAA19659), GfF3GT (BAA12737), PfF3GT (BAA89008), Atf3GT (NP180534), UGT78D2 (AED92377), UGT78A11 (CAN74919), UGT78A12 (BAG82846), HvF3GT (P14726), PiasGT (BAA36421), BAC54093 (ThAsGT), VhAsGT (BAA36423), NgGT2 (BAA88935), PhAsGT (BAA88909), Atf5GT (AAL69494), UGT84A1 (XPO0286202), UGT94B1 (Q5NTH0), BpUGAT (Q5NTH0), UGT94D1 (BAA9027), CmF7G12RT (AFB37772), CmF7G12RT (Q8VGE3), Pha3GRT (CAAS0376), Ipap3G2GT (BAD95582). [See online article for color version of this figure.]
group represented by flavonoid 7-O-glycosyltransferases (the 7-position of isoflavones and the corresponding position of other flavonoids; Hirotani et al., 2000; Noguchi et al., 2007; Ono and Nakayama, 2007). Among the functionally characterized UGTs, UGT707B1 shares the highest identity (40%) with UGT88A4 from Maclura pomifera, which did not exhibit any activity toward flavonoids and isoflavonoids but catalyzed the glucosylation of coumarin substrates, although the in vivo substrate is unknown (Tian et al., 2006). The same identity is shown with UGT88A8 and UGT88A9 from Hieracium pilosella, which showed activity with flavones, flavonols, caffeic acid, esculetin, catechol, resorcinol, and hydroquinone (Witte et al., 2009).

Expression Analysis during Stigma Development and in Other Tissues

As a step toward functional analysis, we examined the expression pattern of UGT707B1 in saffron flowers using a RT-PCR procedure. These analyses revealed that UGT707B1 was expressed in all the flower organs examined (Fig. 3A). In stigmas, UGT707B1 was detected in all developmental stages and reached the

![Figure 3. Expression analysis of UGT707B1 in saffron tissues determined by RT-PCR. A, The level of UGT707B1 expression was analyzed in the stigma tissue of saffron in different developmental stages: yellow, orange, red, 2 d before anthesis (−2da), anthesis (da), 1 d after anthesis (+1da), and 3 d after anthesis (+3da). Expression levels for UGT707B1 were also analyzed in closed and open anthers, tepals, and calli. B, Levels of UGT707B1 expression in tepals, leaves, stamens, and corm. C, Expression levels in style and stigma tissues: white style (section A), yellow style (section B), and stigma (section C). D, Expression analysis of UGT707B1 in the stigmas and tepals of different Crocus species. The levels of the constitutively expressed 18S coding gene were assayed as controls. The PCR products were separated by 1% (w/v) agarose gel electrophoresis and visualized by ethidium bromide staining. [See online article for color version of this figure.]
highest level at the time of anthesis. The expression pattern of UGT707B1 was also examined in vegetative organs, leaves, and corms, and expression was detected in all these tissues except in the corm (Fig. 3B). The expression of UGT707B1 was also analyzed along the style and stigma tissue, with higher expression levels in the style (Fig. 3C), which has been characterized by a higher content in flavonoids (Rubio-Moraga et al., 2010).

The Expression of UGT707B1 in the Tepals of Different Crocus Species Was Associated with the Presence of Kaempferol-3-O-β-D-Glucopyranosyl-(1-2)-β-D-Glucopyranoside

Although saffron is mainly appreciated for its long red stigmas, other crocuses have been exploited by plant breeders and gardeners for their ornamental value. The expression of UGT707B1 was analyzed in several Crocus species that highly differ in stigma and tepal coloration due to differences in flavonoid and anthocyanin contents (Nørbaek et al., 2002). Flavonols, although possessing a weak yellow color as pure substances, are potential copigments, and their sugar units are important for their interactions with anthocyanins (Meng et al., 2011). UGT707B1 was detected in the stigmas at anthesis in all the crocuses analyzed (Fig. 3D; Supplemental Fig. S2). However, when the expression was analyzed in tepals, the transcripts of UGT707B1 were undetectable in Crocus speciosus, Crocus pulchellus (Fig. 3D; Supplemental Fig. S2), and Crocus chrysanthus (data not shown). In order to determine the possible correlation between UGT707B1 expression levels and the presence of specific glucosides, flavonols and anthocyanins were extracted from the tepals of saffron, Crocus carwrightianus ‘Albus’, Crocus ochroleucus, C. speciosus, C. pulchellus, and C. chrysanthus. HPLC-electrospray ionization (ESI)-tandem mass spectrometry (MS/MS) analyses were performed with each sample (Fig. 4). First, the samples were compared and searched for both different and identical peaks in terms of retention time (Rt), UV spectra, and mass data. In saffron, Crocus carwrightianus ‘Albus’, and C. ochroleucus samples, several major peaks were identified as common (Fig. 4A). Using the MS and UV data, these peaks were identified as kaempferol-3-O-β-D-glucopyranosyl(1-2)-β-D-glucopyranoside (kaempferol 3-O-β-glucoside) and quercetin-3-O-β-D-glucopyranosyl(1-2)-β-D-glucopyranoside (quercetin 3-O-β-glucoside; Nørbaek et al., 2002), whereas in C. speciosus, C. pulchellus, and C. chrysanthus, the identified major peaks were different (Fig. 4B) and corresponded to quercetin 3-O-α-(2-O-β-glucosyl)-rhamnoside-7-O-β-glucoside (compound 1) and kaempferol 3-O-α-(2-O-β-glucosyl)-rhamnoside-7-O-β-glucoside (compound 2). The presence of 3-O-β-sophorosides in saffron, Crocus carwrightianus ‘Albus’, and C. ochroleucus, along with their absence from the other analyzed Crocus species in which UGT707B1 transcripts were not detected, suggested a role of UGT707B1 in the formation of these compounds, with a putative flavonol 3-O-glucoside:2′-O-glucosyltransferase activity.

NMR Characterization of the Putative Kaempferol 3-O-β-Sophoroside

In order to elucidate the complete structure of the putative kaempferol 3-O-β-sophoroside, the saffron extract was separated by HPLC and the corresponding chromatographic peak was manually collected to be further analyzed by NMR. HPLC-ESI-MS/MS data indicated the presence of the flavonoid kaempferol and two units of Glc. Then, unequivocal structural elucidation of this compound was carried out by the combined use of one- and two-dimensional 1H, 13C-NMR (gradient correlation spectroscopy and total correlation spectroscopy) and 1H-13C-NMR experiments (multiplicity-edited gradient heteronuclear single-quantum correlation spectroscopy and gradient heteronuclear multiple bond correlation). 13C- and 1H-NMR data confirmed the kaempferol skeleton (Table I). The 1H-NMR spectrum of the compound recorded in D2O at 293 K showed broad signals in the aglycone region. In particular, the low intensity of H-C(8) as compared with other protons of aglycone was attributed to a keto-enol tautomeration. Finally, in order to obtain the resolved multiplets, experiments were carried out at 313 K.

The presence of two units of Glc was confirmed from the observation of 12 carbons (two of each anomeric carbon) in the 103- to 60-ppm range of the 13C-NMR spectrum. The 1H-NMR spectrum showed two doublets in the anomeric region, and its coupling constants indicated a β-configuration for both Glc residues. The total correlation spectroscopy experiment led to the identification of the 1H signals of the two Glc residues, while a multiplicity-edited gradient heteronuclear single-quantum correlation spectroscopy spectrum was used to link the carbon signals to the corresponding proton resonances. Finally, the HMBC experiment showed the linkages present in the structure. Correlations observed between the Glu’”-H1 anomeric proton (4.75 ppm) and the Glu’”-C2 carbon (79.53 ppm) and between the Glu’”-C1 anomeric carbon (102.32 ppm) and the Glu’”-H2 proton (3.77 ppm) revealed the presence of a β-d-glucopyranosyl-(1→2)-β-d-glucopyranoside moiety. In addition, the observed correlation between the Glu’”-H1 anomeric proton (5.10 ppm) and C-3 (133.51 ppm) revealed the linkage with the aglycone moiety. Therefore, compound 2 was identified as kaempferol-3-O-β-D-glucopyranosyl(1→2)-β-D-glucopyranoside. Full assignments are shown in Table I. 13C-NMR chemical shifts were in accordance with the literature data (Markham et al., 1978).

Biochemical Characterization

To identify the substrate specificity of UGT707B1, the cDNA was cloned into the expression vector pGEX-5X-3 (Supplemental Fig. S3). UGT707B1 eluted after glutathione-Sepharose purification did not show appreciable activity over any of the assayed substrates.
including kaempferol-3-glucoside as substrate. Variations in incubation buffers, time, and/or temperature did not improve the results. Thus, we used other expression vectors with different tags: pET100 (Invitrogen), pDEST14 (Invitrogen), and pE-SUMOpro3 (Lifesensors), all three with His tags, and pMAL-p5X.
Overexpression of *UGT707B1* in Arabidopsis

Due to the lack of activity in the in vitro assays over many of the substrates tested, we decided to develop transgenic Arabidopsis plants expressing *UGT707B1*. The cauliflower mosaic virus 35S promoter was used to drive the expression of *UGT707B1* in transgenic Arabidopsis plants. Nine independent lines overexpressing *UGT707B1* were taken to homozygosity through backcrossing to Col-0. RT-PCR was carried out to confirm the steady-state levels of its transcripts in each of the lines. The *UGT707B1*-expressing plants displayed phenotypes that differed from those of the wild-type plants. Figures 5 and 6 show data for six lines selected as non-expressing (lines 9 and 12) and expressing (lines 1, 3, 4, and 5) compared with the wild-type (Fig. 5A). The transgenic expressing lines showed aberrant cotyledon and leaf development, with a peripheral zone that was bent upward compared with ecotype Columbia (Col-0; Fig. 5B). This is referred to as hyponastic growth and is the result of asymmetric growth of the adaxial and abaxial surfaces of the cotyledon (Kang, 1979). Alterations in stem thickness were observed (Fig. 5C), and flowering time was also delayed in the expressing lines for up to 2 weeks (Fig. 5D). The flowering time was defined by the number of rosette leaves when floral buds were clearly visible in Col-0, *UGT707B1*-9.1, *UGT707B1*-12.1, and *UGT707B1*-12.2 lines. The presence of flower buds is highlighted in Figure 5D by light gray bars, while dark gray bars denote the absence of flower buds at the time of the experiment. In fact, when siliques were already present in Col-0 plants, flower buds were still absent from some of the expressing *UGT707B1* transgenic lines (Supplemental Fig. S4). Other phenotypic characteristics were observed in the expressing lines: the trichomes of the first rosette leaves showed an altered morphology with a significant reduction in their number compared with the wild-type and nonexpressing lines (Fig. 6).

Metabolite Profiles in Transgenic Arabidopsis Plants

For the analysis of flavonol accumulation, wild-type, *UGT707B1*-1.2, *UGT707B1*-5.4, and *UGT707B1*-9.1 plants were grown for 7 weeks, and different tissues were separately pooled and analyzed. The targeted metabolite analysis was performed by HPLC-diode array detector (DAD) and HPLC-ESI-MS/MS. Most flavonol glycosides were identified by their UV light absorption spectra and ESI-MS/MS analysis. Recently published data were also used to interpret results (Veit and Pauli, 1999; Kerhoas et al., 2006; Le Gall et al., 2006; Stobiecki et al., 2008; Yonekura-Sakakibara et al., 2008). Flavonol levels were obtained by calculating the area below each HPLC peak per milligram dry weight of plant material. Figure 7 shows the elution profiles of rosette leaves, cauline leaves, stems, and inflorescences at 254 and 340 nm. Several changes were observed between the wild type and the expressing lines. Some flavonols were less abundant than in wild-type tissues, others were elevated in the expressing lines, and new ones appeared in the transgenic lines (Fig. 7; Table II). In rosette and cauline leaves, the major flavonols f3 (mass-to-charge ratio [m/z] 739.05, 593, 428.96, and 283.85; Rt 29 min), f2 (m/z 593.01, 446.92, 430.95, and 284.89; Rt 37.4 min), and f1 (m/z 577.1, 430.89, and 284.88; Rt 42.4 min) were detected in the wild type and in transgenic lines (Fig. 7; Table II). In transgenic lines, an additional compound not present in wild type Col-0, fx (m/z 755.0, 609, 431.02, and 284.85; Rt 33 min) was detected (Fig. 7; Table II). In flowers, changes in the relative levels of flavonols were also appreciated (Fig. 7; Table II). The Arabidopsis flowers contain the most flavonol derivatives in comparison with leaves and stems (Fig. 7). The flavonols f8 (m/z 757, 611, 449,
and 303; Rt 25.6 min), f3, f6 (m/z 611, 449, and 303; Rt 33.3 min), f15, f26 (m/z 627.9, 465, and 303; Rt 35.4 min), f2, f5 (m/z 592.98, 446.94, 300.88, and 178.82; Rt 38.1 min), f14 (623, 476.94, and 303.86; Rt 38.1 min), f21, f1, and f30 (m/z 606.96, 460.95, and 314.9; Rt 43.1 min) were detected in wild-type plants. In transgenic lines, clear differences in the pattern of accumulation were observed. We detected the presence of f8, f3, f26, f14, f21, f1, and f30 at approximately equal levels as in the wild type; f5 and f2 were reduced three to four times in comparison with the wild-type; while f6 and f15 flavonols were not detected. Instead, fx was again detected in this tissue, and high levels of f21 (m/z 608.99, 446.97, and 284.88; Rt 39.4 min) and f26 were observed (Fig. 7; Table II). Stems showed a relatively low number of flavonol types in the wild type, with the presence of the kaempferol derivatives f3, f2, and f1 at high levels (Fig. 7). In the transgenic lines expressing UGT707B1, an important reduction of these flavonols was observed, with up to three times less for f1 and f3 and more than a 30 times reduction for f2 (Table II). Instead, f21 (m/z 609 and 286; Rt 40.1 min) was detected as the second major peak in addition to f1 (m/z 578.83, 433.02, and 287.05; Rt 42.08 min); other minor peaks were also detected as f3 (m/z 740.90 and 594.98; Rt 29.6 min), f2 (m/z 595, 442.8, and 287; Rt 37.8 min), and fx (m/z 755.0, 609, 431.02, and 284.85; Rt 33 min).


UGT707B1 Catalyzes Flavon Sophoroside Biosynthesis

Figure 5. Phenotypes of Arabidopsis transgenic lines expressing UGT707B1. A, RT-PCR of Col-0 wild-type plants (wt) and different transgenic lines transformed with UGT707B1. B, Aberrant cotyledons and reduced trichome number in the transgenic lines. The periphery of cotyledons is bent upward, referred to as hyponastic growth. Also, the leaves of UGT707B1-expressing lines contain fewer trichomes (indicated by white arrows). C, Thicker flower shoot in UGT707B1-expressing lines. D, Rosette leaf numbers at bolting of the Col-0 wild type, UGT707B1-expressing plants (lines 1.2, 1.3, 3.0, 3.1, 3.2, 4.2, 4.3, 4.4, and 5.4), and UGT707B1-nonexpressing plants (lines 9.1, 12.1, and 12.2) grown under long-day conditions. At least 20 plants were scored for each line. Error bars indicate SD. [See online article for color version of this figure.]
The new compound that accumulates in UGT707B1-expressing lines was identified as a kaempferol with a Rha molecule and two Glc residues. Mild acid hydrolysis of fx led to the detection of the aglycone kaempferol and kaempferol 3-sophoroside (f21). The UV study in methanol after the addition of UV-shift reagents (Mabry et al., 1970) showed that fx was a kaempferol derivative, with the hydroxyls at the 7- and 3-positions blocked. Furthermore, we determined the possible structure of fx based on the ions obtained through its fragmentation (Ferreres et al., 2004). The ESI plus full MS resulted in the detection of a main peak of $m/z$ 755.01; the MS/MS of that peak resulted in a fragment of $m/z$ 609.00 (755.01 - 146.01), suggesting the loss of a Rha molecule; the MS/MS/MS of the 609 peak generated a fragment of $m/z$ 284.85 (609 - [2 × 162]; Supplemental Fig. S5). Taking into consideration the presence of 3- and 7-substitutions on the Arabidopsis flavonols, this successive pattern of fragmentation of fx suggests a 7-substitution in the kaempferol molecule with a Rha molecule and a 3-substitution with a sophoroside molecule. Although this compound has not been detected in Arabidopsis Col-0 wild-type plants (Yonekura-Sakakibara et al., 2008; Stracke et al., 2010), its presence has been reported in rosette leaves of the Arabidopsis ecotype C-24 (Stobiecki et al., 2006). Flavonol aglycones were not detected in any of the plant lines analyzed. Table II shows the spectral data of each of the identified flavonol species and the induction/repression factor in the expressing lines.
compared with the wild type. In general, flavonols containing 3-O-sophoroside were increased severalfold. The clearest alteration was found for fx (kaempferol 3-O-sophoroside-7-O-rhamnoside), newly detected in all the tissues analyzed in the transgenic lines that expressed UGT707B1. This new presence can be partially explained by...
by the reduction observed in f2 (kaempferol-3-O-glucoside-7-O-rhamnoside). The second differentially detected compounds were f21 and f26. f21 (kaempferol-3-O-sophoroside) was detected in the wild-type and UGT707B1-1.2 plants. In addition, a putative MYBPZM-binding element was present that is present in the A1 gene, required for 3-deoxy flavonoid and phlobaphene biosynthesis and recognized by the Myb homolog P gene from maize (Grotewold et al., 1994). There are no data about cis-elements recognized by other transcription factors known to modulate the expression of glycosyltransferases in different tissues (Stracke et al., 2007, 2010); therefore, we do not know if such elements could be recognized by R2R3-MYB transcription factors. The promoter sequence showed multiple light-regulatory units (Table III, as have been shown in the promoter sequences of genes of the phenylpropanoid biosynthesis (Hartmann et al., 2005). In Arabidopsis, these light-regulatory units contain two distinct types of cis-regulatory elements, namely the ACGT-containing element (core sequence, CACGT) and the MYB-recognition element (core sequence, ACCTT). It has been demonstrated that the MYB-recognition element acts via binding MYB12, an R2R3-MYB-type transcription factor (Mehrtens et al., 2005). These two elements are present in the promoter region of UGT707B1. In addition, a putative MYBPZM-binding element was present that is present in the A1 gene.

Table II. The flavonol profiles of measuring peak area using a response value of the peak area of the internal standard. The new peaks and increasing peaks in the transgenic UGT707B1-1.2 line are highlighted in boldface. Flavonols are as follows: f1, kaempferol-3-O-rhamnoside; f2, kaempferol-3-O-glucoside-7-O-rhamnoside; f3, kaempferol-3-O-rhamnosyl(1→2)-glucoside-7-O-rhamnoside; f5, quercetin-3-O-rhamnoside-7-O-rhamnoside; f6, quercetin-3-O-glucoside-7-O-rhamnoside; f8, quercetin-3-O-rhamnosyl(1→2)-glucoside-7-O-rhamnoside; f14, isorhamnetin-3-O-glucosyl-7-O-rhamnoside; f21, kaempferol-3-O-sophoroside; f26, quercetin-3-O-sophoroside; f30, isorhamnetin-3-O-rhamnoside-7-O-rhamnoside; f14, kaempferol-3-O-glucosyl(1→2)-glucoside-7-O-rhamnoside. n.d., Not detected.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Compound</th>
<th>Wild Type</th>
<th>UGT707B1-1.2</th>
<th>Wild Type</th>
<th>UGT707B1-1.2</th>
<th>UGT707B1-1.2:Wild Type Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Wild Type</td>
<td></td>
<td>Wild Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rt UV Spectra</td>
<td>Peak Area</td>
<td>Rt UV Spectra</td>
<td>Peak Area</td>
<td></td>
</tr>
<tr>
<td>Rosette leaves</td>
<td>f3</td>
<td>29</td>
<td>268, 348</td>
<td>33.08</td>
<td>29</td>
<td>268, 348</td>
</tr>
<tr>
<td></td>
<td>f6</td>
<td>fx</td>
<td>33</td>
<td>280, 310</td>
<td>17.36</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>f2</td>
<td>f2</td>
<td>37.4</td>
<td>265, 348</td>
<td>21.9</td>
<td>37.4</td>
</tr>
<tr>
<td></td>
<td>f1</td>
<td>f1</td>
<td>42.4</td>
<td>264, 342</td>
<td>146</td>
<td>42.4</td>
</tr>
<tr>
<td>Cauline leaves</td>
<td>f3</td>
<td>29.3</td>
<td>268, 348</td>
<td>221</td>
<td>29.3</td>
<td>268, 348</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>fx</td>
<td>33.2</td>
<td>n.d.</td>
<td>9.8</td>
<td>33.2</td>
</tr>
<tr>
<td></td>
<td>f2</td>
<td>f2</td>
<td>37.7</td>
<td>265, 348</td>
<td>173.5</td>
<td>37.3</td>
</tr>
<tr>
<td></td>
<td>f1</td>
<td>f1</td>
<td>42.4</td>
<td>264, 342</td>
<td>431.6</td>
<td>42.4</td>
</tr>
<tr>
<td>Stems</td>
<td>f3</td>
<td>f3</td>
<td>29.6</td>
<td>268, 348</td>
<td>135.75</td>
<td>29.6</td>
</tr>
<tr>
<td></td>
<td>f6</td>
<td>fx</td>
<td>33.6</td>
<td>271, 340</td>
<td>3.7</td>
<td>33.6</td>
</tr>
<tr>
<td></td>
<td>f2</td>
<td>f2</td>
<td>38</td>
<td>265, 348</td>
<td>174.4</td>
<td>37.8</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>f21</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>40.1</td>
</tr>
<tr>
<td>Flowers</td>
<td>f8</td>
<td>f8</td>
<td>25.6</td>
<td>260, 352</td>
<td>62.7</td>
<td>25.9</td>
</tr>
<tr>
<td></td>
<td>f3</td>
<td>f3</td>
<td>29</td>
<td>268, 348</td>
<td>499</td>
<td>29.2</td>
</tr>
<tr>
<td></td>
<td>i6 + f15</td>
<td>fx</td>
<td>33.3</td>
<td>261, 353</td>
<td>280.8</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td>f26</td>
<td>f26</td>
<td>35.5</td>
<td>260, 352</td>
<td>106.13</td>
<td>35.4</td>
</tr>
<tr>
<td></td>
<td>f2</td>
<td>f2</td>
<td>37.3</td>
<td>263, 352</td>
<td>255.5</td>
<td>37.3</td>
</tr>
<tr>
<td></td>
<td>f5 + f14</td>
<td>f5 + f14</td>
<td>38.1</td>
<td>258, 353</td>
<td>740.7</td>
<td>38.1</td>
</tr>
<tr>
<td></td>
<td>f21</td>
<td>f21</td>
<td>39.4</td>
<td>265, 350</td>
<td>145.5</td>
<td>39.4</td>
</tr>
<tr>
<td></td>
<td>f30</td>
<td>f30</td>
<td>43.1</td>
<td>263, 350</td>
<td>259.17</td>
<td>43.1</td>
</tr>
</tbody>
</table>

Promoter Analysis of UGT707B1

Variation in the flavonol pattern in different organs, and thus the formation and accumulation of specific flavonol compounds, could be caused by differences in the set of genes coding for those enzymes responsible for flavonol modification. In Arabidopsis, there is clear evidence that R2R3-MYB transcription factors are able to activate glycosyltransferase expression (Tohge et al., 2005; Yonekura-Sakakibara et al., 2008; Stracke et al., 2010), controlling flavonol accumulation in an organ- and development-dependent manner. Hence, we isolated and analyzed the promoter region of UGT707B1 in order to identify elements that could be recognized by R2R3-MYB transcription factors. The promoter sequence showed multiple light-regulatory units (Table III, as have been shown in the promoter sequences of genes of the phenylpropanoid biosynthesis (Hartmann et al., 2005). In Arabidopsis, these light-regulatory units contain two distinct types of cis-regulatory elements, namely the ACCT-containing element (core sequence, CACCT) and the MYB-recognition element (core sequence, ACCTT). It has been demonstrated that the MYB-recognition element acts via binding MYB21, an R2R3-MYB-type transcription factor (Mehrtens et al., 2005). These two elements are present in the promoter region of UGT707B1. In addition, a putative MYBPZM-binding element was present that is present in the A1 gene, required for 3-deoxy flavonoid and phlobaphene biosynthesis and recognized by the Myb homolog P gene from maize (Grotewold et al., 1994). There are no data about cis-elements recognized by other transcription factors known to modulate the expression of glycosyltransferases in different tissues (Stracke et al., 2007, 2010); therefore, we do not know if such elements could be present in the promoter of UGT707B1. Nevertheless, in Arabidopsis, the glycosyltransferase enzymes responsible for the formation of K-3GG and Q-3GG are unknown, as are the MYB transcription factors, which differentially influence the spatial accumulation of these specific flavonol derivatives.
DISCUSSION

Molecular and metabolite analysis approaches were used to identify UGT707B1-metabolite correlations in order to determine the role of UGT707B1 in flavonol glucosylation in saffron. Because coupling the analysis of metabolites with coexpression has been used as an efficient method for identifying glycosyltransferase functions (Gachon et al., 2005; Tohge et al., 2005; Yonekura-Sakakibara et al., 2007, 2008, 2012), we have used this strategy on several Crocus species that differ in the expression of a glucosyltransferase gene, UGT707B1, highly expressed in saffron flowers. Furthermore, these Crocus species differed in the activity of 3-O-glucoside-2'-O-glycosyltransferases in tepals. The nonexpressing UGT707B1 group was characterized by high activity of a flavonol 3-O-glucoside-2'-O-rhamnosyltransferase, while the UGT707B1-expressing group had an active flavonol 3-O-glucoside-2'-O-glucosyltransferase. In addition, the use of transgenic Arabidopsis lines expressing UGT707B1 effectively supports the role of UGT707B1 in saffron as a flavonol 3-O-glucoside-2'-O-glucosyltransferase involved in the formation of flavonol 3-O-sophorosides.

The UDP-glucosyltransferase UGT707B1 was identified in saffron stigmas based on the conservation of amino acid domains of plant UDP-glucosyltransferases associated with secondary metabolism (Hughes and Hughes, 1994; Gachon et al., 2005). In the C-terminal domain, UGT707B1 contains a motif that shows similarity with the 44-amino acid-long PSPG sequence found in most known UDP-glucosyltransferases. The motif comprises amino acid residues playing a key role in binding the nucleotide-diphosphate sugar (Osmani et al., 2009). Particularly, Trp (W), Glu (E), and Gln (Q) residues found in the PSPG sequence of UGT707B1 (at positions 22, 43, and 44 of the motif, respectively) have been shown to be conserved in several plant UDP-glucosyltransferases, preferring UDP-Glc as the sugar donor substrate (Shao et al., 2005). In fact, Glc is the unique glycoside present in saffron stigmas ( Rubio-Moraga et al., 2010), where UGT707B1 was expressed throughout all the developmental stages, although its expression was developmentally regulated, with higher levels at the time of anthesis characterized by the highest accumulation of kaempferol sophorosides (Moraga et al., 2009). In addition, UGT707B1 transcripts were detected in other flower parts such as the style, rich in flavonols (Rubio-Moraga et al., 2010), stamens, and tepals, and they were also present in leaves but absent from the corm.

The phylogenetic relationships did not allow the determination of the possible activity of UGT707B1, while only one of the most related sequences, UGT707A3 (with 52% identity), showed a 3-O-glucosyltransferase activity toward kaempferol and quercetin ( Ko et al., 2008). Moreover, the biochemical basis of the enzymes from the most related family, UGT71, is not decisive due to the significant structural diversity of their substrates. For example, Medicago truncatula UGT71G1 and Arabidopsis UGT71C1 have glucosylation activity toward triterpene and coumarin in vitro, respectively ( Lim et al., 2003; Achnine et al., 2005). Arabidopsis UGT71B proteins are able to glucosylate a spectrum of different compounds such as asaciscic acid and hydroxybenzoic acid ( Lim et al., 2002; Priet et al., 2006; Brazier-Hicks et al., 2007), whereas UGT71A9 from Sesamum indicum catalyzed the glucosylation at the 2-hydroxyl group of (+)-sesaminol (Noguchi et al., 2008). Because sequence homology is far from being a definitive argument to describe the precise enzymatic activity of glucosyltransferase enzymes ( Caputi et al., 2012), the functional characterization of these enzymes is still necessary. Furthermore, correlations between transcripts and metabolites could facilitate an efficient narrowing down of candidate products of glucosyltransferase activities (Yonekura-Sakakibara et al., 2008; Kovinich et al., 2011; von Saint Paul et al., 2011).

Expression analysis of UGT707B1 in the stigmas and tepals of other Crocus species that differ in flavonol and anthocyanin accumulation allowed the determination of the possible products of UGT707B1 activity.

Table III. Putative cis-elements identified in the UGT707B1 promoter

<table>
<thead>
<tr>
<th>Category</th>
<th>Cis-Acting Element</th>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light</td>
<td>ACGT-containing element</td>
<td>ACGTGGGA</td>
<td>−778, −630</td>
</tr>
<tr>
<td></td>
<td>ABRE</td>
<td>GACACGTGGC</td>
<td>−729</td>
</tr>
<tr>
<td></td>
<td>ATCT motif</td>
<td>AATCTAATCC</td>
<td>−672</td>
</tr>
<tr>
<td></td>
<td>G box</td>
<td>CACGTG</td>
<td>−727</td>
</tr>
<tr>
<td></td>
<td>G box</td>
<td>ACACGTGGC</td>
<td>−409, −729</td>
</tr>
<tr>
<td></td>
<td>GATA motif</td>
<td>GATAAGA</td>
<td>−468</td>
</tr>
<tr>
<td></td>
<td>GTGCCC motif</td>
<td>CATCGTGGC</td>
<td>−967, −1,199</td>
</tr>
<tr>
<td></td>
<td>I box</td>
<td>GATAAAGTC</td>
<td>−191, −970</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salicylic acid response</td>
<td>SARE</td>
<td>TTCGACCTCCTT</td>
<td>−237</td>
</tr>
<tr>
<td>MYB flavonol activation</td>
<td>MYBPZM</td>
<td>CCGWACC</td>
<td>−967, −1,110</td>
</tr>
<tr>
<td>MYB-binding site involved in drought inducibility</td>
<td>MBS</td>
<td>CAACTG</td>
<td>−302, −467</td>
</tr>
<tr>
<td>Abscisic acid response</td>
<td>CE3</td>
<td>GACGGTGTCC</td>
<td>−729</td>
</tr>
<tr>
<td>MYB-binding site</td>
<td>MBY recognition element</td>
<td>AACCTT</td>
<td>−630, −1,011</td>
</tr>
<tr>
<td>Core promoter element</td>
<td>TATA box</td>
<td>ATATAT</td>
<td>−109</td>
</tr>
<tr>
<td>Common cis-acting element in promoter and enhancer regions</td>
<td>CAAT box</td>
<td>CAAT</td>
<td>−97, −248, −662, −918, −1,092</td>
</tr>
</tbody>
</table>
The transcripts were detected in all the stigmas analyzed but were absent from the tepals of several Crocus species. Consistent with this, a comparison of the glycosylated metabolites from the tepals of these nonexpressing species with those from the UGT707B1-expressing species showed the absence of kaempferol and quercetin 3-O-glucosyl-2’’-O-gluicosides, suggesting that these compounds could be the result of UGT707B1 activity. Furthermore, the amino acid sequence identity (25%–29%) of UGT707B1, with UGTs catalyzing glycosylation at the sugar moiety of a glycoside (Richman et al., 2005; Shibuya et al., 2010; Yonekura-Sakakibara and Hanada, 2011), is not far from that observed among these UGTs.

We were unable to determine the in vitro activity of UGT707B1 using a recombinant enzyme produced in E. coli and several substrates, including kaempferol and kaempferol-3-O-glycoside. Several reasons can explain the lack of activity of the recombinant enzyme. UGTs involved in secondary metabolism pathways are known to be unstable enzymes that occur only in minute amounts in plants (Vogt and Jones, 2000); thus, the first reason can be that UGT707B1 is highly labile, with loss of enzyme activity upon purification. However, when the activity was assayed in culture cells without purification, no products were detected. The second reason could be that the produced recombinant enzyme in E. coli is inactive. UGT707B1 is predicted to contain a signal peptide and to be glycosylated, which suggests that it may play an important role in establishing or maintaining the activity. The third but not last reason is that perhaps the substrates adequate for UGT707B1 activity were not present in the reaction. It cannot be excluded that UGT707B1 might not directly use UDP-Glc as a donor.

Therefore, we decided to functionally characterize UGT707B1 in vivo. The transgenic Arabidopsis lines expressing UGT707B1 showed characteristic phenotypes and were affected in the flavonol glucosylation profile. The observed phenotype of UGT707B1-expressing Arabidopsis lines closely resembles the phenotype previously described in the Arabidopsis rol mutants, affected in the Rha synthase RHMI (Ringli et al., 2008), which accumulate K-3G and kaempferol-3-O-(6”-acetylglucoside) and show hyponastic cotyledon growth and aberrant trichomes on the first rosette leaves, most likely caused by kaempferol-induced modulation of auxin transport (Kuhn et al., 2011). In addition to these phenotypes, the UGT707B1-expressing Arabidopsis lines showed a reduced number of trichomes, a delay in flowering time, and thicker stems. In Arabidopsis, flavonoid biosynthesis and modification have been shown to be involved in trichome formation (Walker et al., 1999; Gou et al., 2011), the transition to flowering (Gou et al., 2011), and flowering time (Wang et al., 2012). The UGT707B1-expressing transgenic plants displayed a strong increase in 3-diglucosylated flavonols, with variations in the different tissues analyzed. Compared with the wild type, there is a reduction in the total amounts of flavonols in the shoots (f3, f2, and f1) along with the presence of new flavonols not detected in the wild-type f21 (K-3GG) and fx (K-3GG-7R), where accumulations in shoots were correlated with the development of the observed phenotype. These compounds were completely absent in lines developing a wild-type shoot phenotype. In flowers, the pattern of flavonol accumulation is more complex (Yonekura-Sakakibara et al., 2008, 2012). In the flowers of UGT707B1-expressing lines, f2 and f5 levels were reduced. By contrast, the levels of f26 (Q-3GG) and f21 (K-3GG) were increased, and fx (K-3GG-7R) appeared as a new flavonol in this tissue. The observed reduction in f2 (K-3G7R) levels in the tissues of the transgenic lines could be explained if this compound is converted by UGT707B1 into fx (K-3GG7R). In addition, the reduction of kaempferol and quercetin 3-O-rhamnosides observed in flowers and stems could be due to the reduced availability of quercetin and kaempferol for their conversion by rhamnosyltransferases, due to the increased transformation of those compounds by UGT707B1 in 3-O-diglucosides.

The K-3GG-7R, Q-3GG, and K-3GG flavonols have been identified previously in the rosette leaves from the Arabidopsis C24 ecotype (Stobiecki et al., 2006), which is not the only ecotype that has shown substantial qualitative and quantitative differences in flavonoid composition in comparison with Col-0 (Tohge et al., 2005). Interestingly, Arabidopsis C24 is a vernization-responsive, moderately late-flowering ecotype in comparison with the early-flowering ecotype Col-0 (Stangeland et al., 2009). The UGT707B1-expressing lines were characterized as well by a delay in flowering transition, suggesting a possible involvement of these glucosylated flavonols in the control of the flowering process. Recently, delayed flowering was reported in flavonol synthase-silenced tobacco (Nicotiana tabacum) lines, which could be due to a decrease in the level of indole acetic acid in the apical region of their shoots (Mahajan et al., 2011), whereas the glycosyltransferase UGT87A2 activity, in a way still unknown, regulates flowering time via the flowering repressor FLOWERING LOCUS C (Wang et al., 2012). Genetic and pharmacologic evidence clearly demonstrate a role for flavonols as negative regulators of auxin transport. The flavonols quercetin and kaempferol have been identified as the most active flavonoids, acting as regulators for the transport of auxins (Jacobs and Rubery, 1988; Brown et al., 2001; Taylor and Grotewold, 2005), and can displace synthetic auxin transport inhibitors in vitro (Jacobs and Rubery, 1988) and in vivo (Kuhn et al., 2011). Furthermore, flavonol biosynthesis was recently shown to be induced by auxin through a TRANSPORT INHIBITOR RESPONSE51 auxin receptor-dependent pathway (Lewis et al., 2011).
nucleus (Wang et al., 2012). In addition, the results presented in this work suggest UGT707B1 as being both a cytoplasmic and nuclear protein alike, although the TargetP program, which has high sensitivity and specificity for plant sequences, showed the presence of a signal peptide in the UGT707B1 sequence that was predicted to be in the secretory pathway. The detection in the nuclei was unexpected due to the absence of a nuclear location signal. Nevertheless, in other well-known plant proteins that are localized in the nuclei, the nuclear location signal domain is not necessary for nuclear localization (Fleck and Harberd, 2002). Typically, small molecules (approximately 50 kD) diffuse freely in and out of the nucleus through the nuclear pores, and the calculated 48 kD of the mature UGT707B1 is smaller than the reported permeability of the nuclear pore complex by passive diffusion (Keminer and Peters, 1999). Although the prevailing model suggests that flavonoids are synthesized in a complex of metabolic enzymes located on the cytoplasmic face of the endoplasmic reticulum (Burbulis and Winkel-Shirley, 1999), flavonols accumulate in the nucleus and cytoplasm of different plant species (Buer and Mudy, 2004; Feucht et al., 2004a, 2004b) as the flavonoid biosynthetic enzymes chalcone synthase, chalcone isomerase (Saslowsky et al., 2005), and flavonol synthase (Kuhn et al., 2011), which suggests that the flavonol biosynthetic machinery is also active in the nucleus, where flavonols might modulate gene expression (Naoumkina et al., 2008; Gilbert and Liu, 2010; Gou et al., 2011; Yin et al., 2012). Interestingly, the UGT87A2 enzyme has been found in the cytosol and nucleus (Wang et al., 2012), an M. truncatula (iso) flavonoid β-glucosidase, which digests prestored (iso) flavonoid glucos conjugates, was found to localize in the nucleus of the cells under stress conditions (Naoumkina et al., 2007), and an isoflavone malonyltransferase involved in the formation of (iso)flavonoid conjugates localizes in the nucleus as well (Yu et al., 2008). Thus, the observed nuclear location of UGT707B1 suggests that, in addition to flavonoid biosynthetic enzymes, flavonoid-modifying enzymes are also present in the nuclei and can modulate the activity of flavonoids in this compartment, controlling different developmental processes in the plant.

Flavonols are involved in a vast array of biological functions in plants, and the presence of different flavonol-glycosylation profiles in plant tissues suggests specific functions of these compounds. As an example, the expression of UGT707B1 in the floral organs of saffron is responsible for the presence or absence of a specific set of flavonols. In tepals, flavonoids could directly contribute to color formation (Nishihara and Nakatsuka, 2011) in order to attract pollinators (Mol et al., 1999). In the stigma tissue, UGT707B1 reaches its highest expression levels at the time of anthesis, coincident with the highest accumulation of kaempferol sophorosides (Moraga et al., 2009). The accumulation of flavonols at the time of anthesis has also been observed in other plant species (Beliaeva and Evdokimova, 2004) and could be associated with floral morphogenesis and pollination (Mo et al., 1992). Nevertheless, one of the most important roles of flavonols is to influence the transport of the plant hormone auxin. In this study, the phenotype displayed by UGT707B1 transgenic Arabidopsis plants suggests the involvement of these flavonols in auxin transport. Auxin gradients have been proposed to determine flower morphology: high auxin levels in the apical region promote differentiation and proliferation of the style and stigma, intermediate levels specify the ovary, and low levels in the basal region specify the gynophore (Ståldal and Sundberg, 2009). In Arabidopsis, alteration of these gradients by inhibition of polar auxin transport leads to an elongation of the style and stigma in the apical part, whereas the ovary size concomitantly decreases (Okada et al., 1991). The resulting plants showed a phenotype that could resemble the morphology of saffron flowers. The female organs of saffron consist of a trilocular ovary (0.5–0.7 cm long), a very long style (10–12 cm), and three red styr branches (3–4 cm long) that are folded to give the stigmas a trumpet-like structure (Grilli and Canini, 2004). Auxins participated in the elongation of the floral tube in Crocus (Stark, 1982), and the different parts of the female organs of saffron differ in their flavonol contents. No flavonoids have been detected in ovaries, and between stigma and style, flavonoids are present at higher levels in the style than in the stigma (Rubio-Moraga et al., 2010). In addition, expression analysis showed higher expression levels of UGT707B1 in the style than in the stigma, and this distribution could have a direct role in the control of the auxin gradient and thus in flower morphology.

Therefore, the precise identification of flavonols in conjunction with expression analysis have led to the identification of a gene function. The data in this study indicate that UGT707B1 is an enzyme involved in the formation of flavonol sophorosides in saffron. Based on the accumulation of certain flavonols in the transgenic Arabidopsis plants, we suggest that UGT707B1 has a flavonol-3-O-glucoside:2”'-glucosyltransferase activity.

**MATERIALS AND METHODS**

**Chemicals and Plant Materials**

Chemicals and reagents were obtained from Sigma-Aldrich unless otherwise stated. Plant tissues and stigmas from Crocus sativus grown under field conditions in Tarazona de La Mancha, Spain, were used throughout the experiments. Crocus ghoulmyji, Crocus major, Crocus cartwrightianus ‘Albus’, Crocus speciosus, Crocus niveus, Crocus palchellus, Crocus chrysanthus, and Crocus ochroleucus were obtained from Pottertons Nursery. Stigmas were collected at the developmental stages described previously (Rubio et al., 2008) and defined as follows: yellow stigma, closed bud inside the perianth tubes (around 0.3 cm in length); orange stigma, closed bud inside the perianth tubes (around 0.4 cm in length); red stigma, closed bud inside the perianth tubes (0.8 cm in length); 2da, 2 d after anthesis, dark red stigma in closed bud outside the perianth tubes (3 cm in length); da, day of anthesis, dark red stigma (3 cm in length); +1da, 1 d after anthesis, dark red stigma; and +3da, 3 d after anthesis, dark red stigma. Tepals, styles, and stamens were collected from
flavonoids at the time of anthesis and together with corons were frozen in liquid nitrogen and stored at −80°C until required.

Seeds of Arabidopsis (Arabidopsis thaliana) wild type Col-0 and transgenic lines were sown in pots containing vermiculite and watered with nutrient solution under a controlled environment with a 16-h-light/8-h-dark cycle at 22°C/18°C.

Cloning of Saffron UGT707B1 cDNA

As a first step in identifying glycosyltransferase genes expressed in saffron stigmas, total RNA and mRNA were isolated from developed saffron stigmas by using Ambion PolyATtract and following the manufacturer’s protocols (Ambion). First-strand cDNA was synthesized by RT from 2 μg of total RNA using an 18-μl oligo(dT) primer and a first-strand cDNA synthesis kit (Amersham Biosciences) according to the manufacturer’s instructions. These cDNAs were used as templates for PCR using degenerate primers designed based on the conserved regions of Ross et al. (2001). The primers used were used to analyze and identify the 5’ and 3’ ends of the glycosyltransferase. For these, 1 μg of poly(A) RNA from stigmas was used to synthesize the 5’ and 3’ ends of the first-strand cDNA using SuperScript II reverse transcriptase, using the 5’-CDs primer and SMART-A oligonucleotide for the 5’-RACE reaction and the SMART-B primer for the 3’-RACE reaction (SMART RACE cDNA Amplification Kit (Clontech-Takara). After dilution, the first-strand reaction product was subjected to PCR for amplification. We used the gene-specific primers UGT707B1-f1 (5’-AACAGAGAGGGATGGTGTTG-3’) and UGT707B1-f2 (5’-TGGAAAGACATCTACAGGAGC-3’) as forward primers and UGT707B1-r1 (5’-CTCTGCTCAGATCCTTCAAA-3’) and UGT707B1-r2 (5’-TGGATCTGCTTGACAGCT-3’) as reverse primers in combination with the universal primer mix from the SMART RACE cDNA Amplification Kit (Clontech-Takara) as the reverse/forward primer with the following cycling program: one cycle at 94°C for 3 min; 10 cycles at 94°C for 20 s, 60°C to 0.2°C per cycle for 20 s, and 72°C for 2 min; 30 cycles at 94°C for 20 s, 60°C for 20 s, and 72°C for 2 min; and a final extension at 72°C for 5 min. The amplified PCR products were analyzed by electrophoresis on a 1% agarose gel. The PCR products were then cloned into pGEM-T (Promega). The ligated DNA was transformed into E. coli strain JM109. The clones (20 colonies) were picked individually and amplified in 3 ml of Luria-Bertani (LB) medium at 37°C overnight. The plasmid DNA from each clone was extracted using a plasmid Miniprep kit (Promega) and then analyzed by EcoRI restriction digestion. Five clones of each size were sequenced with Sp6 and T7 primers using an automated DNA sequencer (ABI PRISM 3730xl; Perkin-Elmer) from Macrogen. Computer-aided sequence similarity searches were made with the BLAST suite of programs at the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov) Motif searches were done using PROSITE (http://expasy.hcuge.ch/sprot/prosite.html), TMPRED (http://www.isrec.isb-sib.ch/software/software.html), SignalP (http://www.cbs.dtu.dk/services/SignalP), and PSORT II (http://psort.nibb.ac.jp). Once the 5’ and 3’ sequences were determined, the full-length cDNA UGT707B1 was amplified from the cDNA and genomic DNA with the following primer sequences, forward primer 5’-ACGATGCAAATCTTCTGTTCT-3’ and reverse primer 3’-TCAGGAGAATCTTATGCTGCA-3’, in order to determine the presence of intron sequences.

Phylogenetic Analysis

To construct the phylogenetic tree, the amino acid sequences were aligned using the BLOSUM62 matrix with the ClustalW (http://www.clustal.org) algorithm-based AlignX module from MEGA version 4.0 (http://www.megasoftware.net/mega.html). The alignments were saved and executed by MEGA version 4.0 to generate a neighbor-joining tree with bootstrapping (5,000 replicates) analysis and handling gaps with pairwise deletion.

Isolation of Promoter Sequences

Genomic DNA was prepared from saffron leaves with the cetyl-trimethyl-ammonium bromide method. The saffron UGT707B1 upstream flanking sequences were isolated with the GenomeWalker Universal Kit (BD Biosciences) using two nonoverlapping gene-specific primers, Pr-1 (5’-AGATGAGTACGGCTGGCAAA-GAAGT-3’) and Pr-2 (5’-AGGAGGAGAGGATCCATGCAA-3’), based on the cDNA sequences obtained. All PCRs were performed using Advantage 2 Polymerase mix (BD Biosciences). Two rounds of PCR amplification using the P-r1 primer with an adapter-specific primer and nested PCR amplification with P-r2 plus another adapter-specific primer gave fragments around 1,400 bp. All PCR products were ligated to pGEM-T with the TA Cloning Kit (Promega). The ligated DNA was transformed into E. coli strain JM109. The clones (20 colonies) were picked individually and amplified in 3 ml of LB medium at 37°C overnight. The plasmid DNA from each clone was extracted using the DNA Plasmid Miniprep Kit (Promega). Plasmids were digested using an automated DNA digestion kit (ABI PRISM 3730xl; Perkin-Elmer) from Macrogen. Promoters were analyzed with the PLACE (Higo et al., 1999) and PlantCARE (Lescot et al., 2002) databases.

Expression Analysis

For RT-PCR, total RNA was isolated from saffron anthers, leaves, tepals, styles, corons, calli, and stigmas (anthesis) and from the tepals and stigmas (anthesis) of the seven other Crocus species and from wild-type and transgenic Arabidopsis plants. This was done by grinding fresh tissue in liquid nitrogen to a fine powder and extracting in 1 ml of Trizol reagent (Gibco-BRL) per 100 mg of tissue fresh weight, according to the protocol of the manufacturer. The RNA was resuspended in 100 μl of RNase-free water and treated with RQ1 RNase-free DNase (Promega). The DNase was heat inactivated before RT-PCR. The RNA was quantified with a spectrophotometer at an optical density of 260 and 280 nm and stored at −80°C. Various initial concentrations of treated RNA, ranging over a 10-fold difference, were used to demonstrate the differential accumulation of the RNA in the tissues analyzed in the RT-PCR experiments. Total RNA samples were reverse transcribed with a first-strand cDNA synthesis kit (Amerham Biosciences) and random primers (Promega). The gene expression levels of Crocus samples were evaluated by PCR using 10 μl of each primer (described above), 200 μM deoxynucleotide triphosphates, and 2 units of Taq polymerase (Invitrogen). After an initial denaturation step for 2 min, PCR was performed for 30 cycles at 94°C for 20 s, 55°C for 20 s, and 72°C for 2 min. As an internal control, the mRNA level of the constitutively expressed ribosomal protein 18 was used (Moraga et al., 2004). The PhotoCaptMw program was used to quantify the intensity of the ethidium bromide-stained DNA bands from the positive images of the gel.

For the analysis of expression in the Arabidopsis samples, the primers used for UGT707B1 amplification were the same as those used above, and the Arabidopsis actin gene was used as a control as described previously (Gómez-Gómez et al., 1999).

Immunocytochemical Analysis

Polyclonal antibodies against UGT707B1 were produced in rabbits (GenScript), for which UGT707B1 was subcloned into a bacterial expression vector with the pColdI-XHis tags. Recombinant purified protein was used for rabbit immunization. Tissue for immunochemical localization was fixed in 4% paraformaldehyde in 100 mM sodium phosphate buffer at pH 7.2. After dehydration in a graded ethanol series, the tissue was infiltrated and embedded in paraffin (Paraplast plus; Sigma-Aldrich). Sections 10 μm thick were obtained and mounted onto poly-L-lysine slides (Sigma-Aldrich). The sections were blocked with 5% bovine serum albumin (w/v) in Tris-buffered saline buffer (20 ml) of Tris-HCl [pH 8.2] and 0.9% NaCl for 30 min, incubated with the UGT707B1 antibody (1:1,000) in Tris-buffered saline buffer (pH 8.2), 0.1% bovine serum albumin, and 0.05% Tween 20 for 2 h, rinsed in the same solution, and incubated with rabbit anti-mouse IgG labeled with alkaline phosphatase for 1 to 2 h. Alkaline phosphatase was detected by the 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium procedure. A purple precipitate forms at the site of localization. Controls included the use of preimmune serum and the elimination of the antibody incubation step. A bright-field microscope was used for sample visualization and photography (E600; Nikon).

Vector Construction and Arabidopsis Transformation

To produce transgenic plants in which the UGT707B1 protein was expressed under the control of the 35S promoter, the UGT707B1 open reading frame was cloned into the Gateway Entry vector pENTR/D TOPO (Invitrogen) and confirmed by sequencing. For stable transformation by Agrobacterium tumefaciens, the UGT707B1 clone was introduced into the Gateway plant transformation destination vector pBGW7 vector (Karimi et al., 2002) using Gateway LR Clonase enzyme mix with pENTR-UGT707B1 according to the manufacturer’s instructions (Invitrogen). The reading frame of the resulting vector, pBGW7-UGT707B1, was confirmed by sequencing, and the vector was transformed
HPLC Analysis of Flavonoids and Anthocyanins Extracted from Crocus Tepals

For the analysis of flavonoids and anthocyanins in tepals from different Crocus species, 30 mg of tissue was extracted with methanol:HC10:water (80:1:20) at 5°C for 24 h and subsequently macerated with a pestle, followed by vigorous vortexing and centrifugation at 10,000 g for 10 min. After centrifugation, the supernatant was collected, filtered, and used for HPLC analysis as described previously (Moraga et al., 2009).

NMR Structural Characterization

Structure elucidation was accomplished by NMR spectroscopy. NMR spectra were recorded at 293, 303, and 313 K, using D2O as the solvent. One-dimensional NMR experiments (1H and 13C) were performed using standard Varian pulse sequences. Two-dimensional 1H, 1H-NMR experiments (gradient correlation spectroscopy and total correlation spectroscopy) were carried out with the following parameters: a delay time of 1 s, a spectral width of 3,906.2 Hz in both dimensions, 4,096 complex points in t2, and four transients for each of 200 time increments, and linear prediction to 512. The data were zero-filled to 4,096 × 4,096 real points. Two-dimensional 1H-13C-NMR experiments (gradient heteronuclear single-quantum correlation spectroscopy and gradient heteronuclear multiple bond correlation) used the same 1H spectral window, a 13C spectral window of 25,133 Hz, 1 s of relaxation delay, 1,024 data points, and 128 time increments, with a linear prediction to 256. The data were zero-filled to 4,096 × 4,096 real points. Typical numbers of transients per increment were four and 16.

HPLC Analysis of Metabolites Extracted from Transgenic Arabidopsis Lines

For analysis of the glucosides of interest in the transgenic plants, rosette leaves, cauline leaves, stems, and flowers from 6-week-old plants were cut, frozen in liquid nitrogen, and lyophilized to determine the dried weight. Flavonol analyses were performed in triplicate as described previously (Stobiecki et al., 2006). The dried material was incubated overnight at 4°C in 500 μL of 80% methanol containing 3 mM scopolin as an internal standard and subsequently macerated with a pestle, followed by vigorous vortexing and centrifugation at 10,000g during 10 min. After centrifugation, the supernatant was collected, filtered, and used for different HPLC analyses. The flavonoid profile analyses were performed by HPLC on a Konik HPLC system equipped with a Suglerlab Inertial ODS 2.5-μm C18 column (250 × 4.6 mm) and connected online to a photodiode array detector with a dynamic range from UV to the visible region (200–700 nm). The column was developed with solvent systems B (90% acetonitrile and 0.05% trifluoroacetic acid) and A (10% acetonitrile and 0.05% trifluoroacetic acid) under the following conditions: 0 to 5 min, isocratic at 5% B; 5 to 40 min, linear gradient from 5% to 25% B; 40 to 48 min, linear gradient up to 100% B; 48 to 60 min, isocratic at 100% B; 60 to 72 min, isocratic at 100% methanol; and for running for 10 min at initial conditions. Using a flow rate of 0.2 mL min⁻¹, compounds were first characterized by their elution time, and their UV absorption spectra were recorded with a photodiode array detector (Konik). The quantitative evaluation was performed on 340 UV-Vis absorbance profiles for flavonol glycosides. Flavonols were quantified by measuring the peak area using as reference the peak area of the internal standard.

HPLC-ESI-MS/MS Analyses

The HPLC-ESI-MS/MS system adopted in this work consisted of a Surveyor HPLC system (Thermo Mod; Finnigan) plus detector and linear ion-trap mass spectrometer detector LXQ (Thermo Mod; Finnigan). This mass spectrometer was equipped with an ESI source. Ion Trap source parameters were as follows: positive mode, ESI source voltage of 4.5 kV, nebulization with nitrogen at 30 psi, dry gas flow of 6.0 L min⁻¹ at temperature of 300°C, and scan range of 150 to 1,000 m/z. Metabolites were identified by their retention times, mass spectra, and product ion spectra in comparison with the data determined for authentic reference materials.

Heterologous Expression in E. coli

The full-length open reading frame of UGT707B1 cDNA was amplified by PCR using Pfu polymerase (Promega). The oligonucleotide sequences for UGT707B1 cloning were as follows: forward primer 5'-GACCGATCCGATGCAAATTTCTCTTGTC-3' and reverse primer 5'-ATCAGGCGCCCTT-TACAGCAGCAATCTTCTTAG-3'. Using these primers, the generated product has a BsmBI site at the 5' end and a NotI site at the 3' end (underlined in the corresponding primers). The PCR product was cloned directionally (Small-NotI) into bacterial glutathione S-transferase expression vector pGEX-5X-3 (GE Healthcare) to create in-frame fusions at the 5' terminus with the glutathione S-transferase coding sequence. The construct was sequenced to confirm that the gene was in the correct reading frame. After transformation into E. coli Rosetta (LysS) cells, colonies were selected on LB medium containing 100 μg mL⁻¹ ampicillin (AMP) and 60 μg mL⁻¹ chloramphenicol (CM) plates. Individual colonies were grown overnight in 5 mL of LB-AMP-CM medium at 25°C, and 2.5 mL of the culture was used to inoculate 500 mL of LB-AMP-CM fresh medium. Cells were grown at 25°C until an A600 of 0.6 was reached, after which the culture was induced with 0.5 mM isopropylthio-β-D-galactoside and allowed to grow for 16 h at 25°C. The cells were harvested by centrifugation at 5,000g for 10 min and resuspended in 20 mL of phosphate-buffered saline. Resuspended cells were sonicated with a microtip probe in ice until the viscosity disappeared. After sonication, the samples were centrifuged at 10,000g for 25 min. The supernatant and pellet were tested by SDS-PAGE for solubility of the fusion protein by Coomassie blue stain. The soluble proteins were applied to a glutathione-Sepharose column for purification following the manufacturer's instructions (GE Healthcare). Protein concentration was determined by Bradford assay using serum albumin as a standard.

Enzyme Assays and Analysis of Reaction Products

The affinity-purified enzyme was used to determine substrate specificity and enzymatic parameters. In a final assay volume of 100 μL, various buffers and conditions were tested: buffer 1, 50 mM Tris-HCl, pH 8.0; buffer 2, 50 mM KH2PO4, pH 7.5; and buffer 3, 50 mM citrate buffer, pH 5.5. All contained 14 mM 2-mercaptoethanol, 2.5 mM UDP-Glc/UDP-Gal, the recombinant enzyme (7.0 μg), and the following corresponding substrates: 100 μM queretin, 100 μM kaempferol, 100 μM kaempferol-3-O-glucoside, 100 μM dihydrokaempferol, 100 μM thymetin, 100 μM isorhamnetin, 100 μM lutein, 100 μM apigenin, 100 μM luteolin, 100 μM quercetin, 100 μM rutin, 100 μM delphinidin, 100 μM caffeic acid, 100 μM coumaric acid, 100 μM genistein, 100 μM esculetin, 100 μM daidzein, 100 μM transcinaminic acid, 1 μM sinapic acid, 1 μM indole acetic acid, and 1 μM abscisic acid. The glucosyltransferase activity assays were carried out at 30°C for 30 to 90 min. The reactions were terminated and the proteins precipitated by the addition of 20 μL of TCA (240 mg mL⁻¹). Subsequently, samples were centrifuged at 15,000g for 5 min to collect the supernatant, and aliquots were analyzed by reverse-phase HPLC as described previously (Moraga et al., 2009) using a C18 Ascentis, 25 × 4.6 mm, particle size 5 μm column (Supelco, Sigma-Aldrich).

Whole-Cell Biotransformation and Flavonoid Glucoside Purification and Identification

Whole-cell biotransformations were performed as described previously (Cartwright et al., 2008) with minor modifications. In brief, E. coli Rosetta (LysS) cells harboring UGT707B1 in a pGEX-5X-3 vector were grown at 28°C, 150 rpm, in 200 mL of M9 minimal medium containing 0.4% glycerol, 100 μg mL⁻¹ AMP, and 60 μg mL⁻¹ CM to an A600 = 0.75 and then induced with 0.1 mM isopropylthio-β-D-galactoside and incubated for 16 h. After the induction period, each substrate, kaempferol, quercetin, apigenin, naringenin, caffeic acid, and luteolin, was added to the bacterial cultures with a final concentration of 50 μM. Additional glycerol was added to the cultures, with a final concentration of 0.15% (v/v), at 4, 7, 25, and 31 h. E. coli cultures harboring the empty vector
were used as a control. The E. coli cultures were collected and centrifuged 36 h after the addition of substrates. The supernatant was extracted twice with ethyl acetate, and the collected ethyl acetate phase was evaporated in a Rotavapor. Residues were dissolved in 500 μL of methanol and subjected to HPLC analysis.

Data Collection and Statistical Analysis

Data presented are means ± SE of at least three independent experiments, with at least 10 plant samples per transgenic line or the wild type. The experimental data were analyzed by one-way ANOVA (Tukey’s honestly significant difference test); the comparisons between the mean values of each transgenic line and the wild type were evaluated by the least significant difference test at P < 0.01.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession number HE90682.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Chastal alignment of monocotyledonous sequences from the UGT707 family.

Supplemental Figure S2. Western-blot analysis of UGT707B1 in tepals and stigmas.

Supplemental Figure S3. SDS-PAGE analysis of recombinant UGT707B1.

Supplemental Figure S4. Late flowering phenotype of UGT707B1 transgenic plants.

Supplemental Figure S5. MS spectra of fx.

Supplemental Figure S6. New metabolites levels in UGT707B1-expressing lines.

ACKNOWLEDGMENTS

We thank J. Argandoña (Instituto Botánico, Universidad de Castilla-La Mancha) for excellent technical support, Dr. M. Orozco-Cardenas (Plant Transformation Research Center, University of California) for Arabidopsis transformation, and K.A. Walsh (Escuela Técnica Superior de Ingenieros Agrónomos, Universidad de Castilla-La Mancha) for language revision.

Received March 30, 2012; accepted May 29, 2012; published May 30, 2012.

LITERATURE CITED


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


Osmani SA, Bak S, Møller BL (2009) Substrate specificity of plant UDP-dependent glucosyltransferases predicted from crystal structures and homology modeling. Phytochemistry 70: 325–347

