PCP
PLANT & CELL PHYSIOLOGY

Regular Paper

# Structural and Functional Analysis of UGT92G6 Suggests an Evolutionary Link Between Mono- and Disaccharide Glycoside-Forming Transferases

Fong-Chin Huang[1], Ashok Giri[1,2], Melina Daniilidis[1], Guangxin Sun[1], Katja Härtl[1], Thomas Hoffmann[1] and Wilfried Schwab[1,*]

[1]Biotechnology of Natural Products, Technische Universität München, Liesel-Beckmann-Str. 1, D-85354 Freising, Germany
[2]Plant Molecular Biology Unit, Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune, MS 411 008, India
*Corresponding author: E-mail, wilfried.schwab@tum.de; Fax, +49-8161-71-2950.
(Received September 22, 2017; Accepted January 30, 2018)

**Glycosylation mediated by UDP-dependent glycosyltransferase (UGT) is one of the most common reactions for the biosynthesis of small molecule glycosides. As glycosides have various biological roles, we characterized *UGT* genes from grapevine (*Vitis vinifera*). In silico analysis of *VvUGT* genes that were highly expressed in leaves identified *UGT92G6* which showed sequence similarity to both monosaccharide and disaccharide glucoside-forming transferases. The recombinant UGT92G6 glucosylated phenolics, among them caffeic acid, carvacrol, eugenol and raspberry ketone, and also accepted geranyl glucoside and citronellyl glucoside. Thus, UGT92G6 formed mono- and diglucosides in vitro from distinct compounds. The enzyme specificity constant $V_{max}/K_m$ ratios indicated that UGT92G6 exhibited the highest specificity towards caffeic acid, producing almost equal amounts of the 3- and 4-*O*-glucoside. Transient overexpression of *UGT92G6* in *Nicotiana benthamiana* leaves confirmed the production of caffeoyl glucoside; however, the level of geranyl diglucoside was not elevated upon overexpression of *UGT92G6*, even after co-expression of genes encoding geraniol synthase and geraniol UGT to provide sufficient precursor. Comparative sequence and 3-D structure analysis identified a sequence motif characteristic for monoglucoside-forming UGTs in UGT92G6, suggesting an evolutionary link between mono- and disaccharide glycoside UGTs. Thus, UGT92G6 functions as a mono- and diglucosyltransferase in vitro, but acts as a caffeoyl glucoside UGT *in N. benthamiana*.**

**Keywords:** Agroinfiltration • Caffeoyl glucoside • Diglucoside • Geranyl glucoside • Glycosyltransferase • *Nicotiana benthamiana* • *Vitis vinifera*.

**Abbreviations:** GES, geraniol synthase; GPP, geranyl diphosphate; GST, glutathione *S*-transferase; LC, liquid chromatography; LC-UV-MS, liquid chromatography coupled with UV and mass spectrometry detection; ORF, open reading frame; PSPG, plant secondary product glycosyltransferase; UGT, UDP-dependent glycosyltransferase; VvGT14a, *Vitis vinifera* geranyl glucosyltransferase.

## Introduction

UDP-dependent glycosyltransferases (UGTs; EC 2.4.x.y) constitute a large family of enzymes that are involved, among other roles, in the biosynthesis of small molecule glycosides (Breton et al. 2006). UGTs transfer a sugar molecule from an activated carbohydrate donor, usually UDP-glucose, to a wide range of acceptors, ranging from hormones and secondary metabolites, including phenylpropanoids, flavonoids, anthocyanidins and alkaloids, to xenobiotics and toxins. The transfer of a sugar onto a lipophilic acceptor changes its chemical properties, alters its bioactivity and enables access to membrane transporter systems. Thus, glycosides find use in various industrial applications. Terpene and phenolic glycosides are found to have antifungal and antimicrobial activity (Sahari and Asgari 2013), and have attracted much attention in the food industry (Mastelic et al. 2004). Glycosides of long chain alcohols or acids have surfactant and emulsifying properties, and are, therefore, contained in detergents and cosmetics (De Roode et al. 2003). Glycosides of flavors and fragrances are used as water-soluble, storage-stable and odorless pro-aroma molecules that can break down to the desired aroma compounds under controlled conditions (De Roode et al. 2001, Herrmann 2007, Schwab et al. 2015).

Plant genomes contain a large number (on average 150) of genes encoding UGTs of small lipophilic molecules (Yonekura-Sakakibara and Hanada 2011, Caputi et al. 2012). In vitro studies have shown that a single gene product can glycosylate multiple substrates of diverse origins, and multiple enzymes can also glycosylate the same substrate (Bönisch et al. 2014a, Bönisch et al. 2014b). In addition to monoglucosides, most of the glycosidically bound small molecules occur as disaccharide, trisaccharide or even oligosaccharide glycosides (Brahmachari et al. 2011, Yu et al. 2011). In grapevine, monoterpenols are chemically bound to disaccharides in the form of 6-*O*-α-L-rhamnopyranosyl-β-D-glucopyranosides (rutinosides), 6-*O*-α-L-arabinofuranosyl-β-D-glucopyranosides and 6-*O*-β-D-apiofuranosyl-β-D-glucopyranosides (Williams et al. 1982). A literature survey revealed at least 20 biochemically characterized di-/trisaccharide glycoside-forming plant UGTs. Two have been characterized from *Citrus* species, which catalyzed the transfer

*Plant Cell Physiol.* 59(4): 862–875 (2018) doi:10.1093/pcp/pcy028, Advance Access publication on 10 February 2018, available online at www.pcp.oxfordjournals.org

of L-rhamnose to D-glucose to form flavanone (and flavones) rutinosides and neohesperidinosides (Frydman et al. 2004, Frydman et al. 2013). A UDP-rhamnose:anthocyanidin 3-*O*-glucoside 6″-*O*-rhamnosyltransferase extracted from *Petunia hybrida* also produced rutinosides (Nakajima et al. 2005), whereas 1,6-glucosyltransferases from *Catharanthus roseus* and *Sesamum indicum* formed flavonoid gentiobiosides (Masada et al. 2009) and (+)-sesaminol 2-*O*-gentiobioside (Noguchi et al. 2008), respectively. Sophoroside-forming glycosyltransferases were isolated from *Bellis perennis* (Sawada et al. 2008), *Crocus sativus* (Trapero et al. 2012) (the enzymatic activity has not been biochemically confirmed), *Ipomoea purpurea* (Morita et al. 2005), *Panax ginseng* (Wang et al. 2015) and *Arabidopsis thaliana* (Yonekura-Sakakibara et al. 2014). A 2″-*O*-xylosyltransferase was also characterized from *A. thaliana* (Yonekura-Sakakibara et al. 2012). Three disaccharide-forming UGTs from *Glycine max* produced flavonoid rutinosides (Rojas Rodas et al. 2014), sophorosides (Di et al. 2015) and gentiobiosides (Rojas Rodas et al. 2016). Four additional *G. max* UGTs were described to catalyze the formation of soyasaponin I and III, and soyasapogenol Aa and Ab, producing di- and trisaccharides (Shibuya et al. 2010, Sayama et al. 2012), whereas UGTs implicated in the 2′-*O*-xylosylation of cyanidin 3-*O*-galactoside, the 2′-*O*-glucosylation of eugenyl *O*-β-D-xylopyranosyl-(1→6)-*O*-β-D-glucopyranososide and the 6′-*O*-xylosylation of geranyl β-D-glucoside were isolated from *Actinidia chinensis*, *Solanum lycopersicum* and *Camellia sinensis*, respectively (Montefiori et al. 2011, Tikunov et al. 2013, Ohgami et al. 2015).

Comparative genomics, which exploits similarities and differences in proteins, RNA and DNA of various organisms, is useful for the detection of related genes in different genomes and yielded the first sequences of monoterpenol UGTs from *Vitis vinifera* (Bönisch et al. 2014a, Bönisch et al. 2014b). This research also produced a large collection of *UGT* genes which were analyzed in follow-up studies. Only recently, *UGT* genes from *V. vinifera* supposedly active in the glucosylation of aroma-active compounds have been isolated from this collection and partially characterized (Härtl et al. 2017). The cloning efforts coupled with gene expression analysis yielded, among other *UGT* genes, *UGT92G6* (VIT_00031678001), which has been shown to be transcribed in grape berries, inflorescences and leaves, but not in roots (Härtl et al. 2017; Supplementary Fig. S1). However, while other candidate genes of the *UGT* collection have been successfully characterized, the natural function of UGT92G6 remained elusive. Since the encoded protein of *UGT92G6* showed sequence similarity to both monosaccharide and disaccharide glycoside-producing UGTs, we chose this gene for detailed analysis.

In this study, we describe the functional characterization of *UGT92G6* from *V. vinifera*. *Escherichia coli* was chosen a production host for the recombinant protein, and the substrate specificities and kinetic parameters of the recombinant UGT92G6 were determined. Homology modeling and 3-D structure analysis were performed to identify essential amino acids. Additionally, *UGT92G6* was transiently overexpressed in *N. benthamiana* leaves using an established viral vector system to investigate the glycoside products produced by UGT92G6 in planta.

## Results

### Phylogenetic analysis positions UGT92G6 in between monosaccharide and disaccharide UGTs

Phylogenetic analysis of the amino acid sequences of UGT92G6, biochemically characterized UGT proteins from grapevine and other functional di-/trisaccharide-forming UGTs revealed that UGT92G6 is obviously a member of a separate cluster (**Fig. 1**). Interestingly, it shows sequence similarity to both monosaccharide and disaccharide glycoside-producing UGTs (Supplementary Fig. S3). UGT92G6 carries a methionine at position 25 (consensus sequence) and lacks proline (replaced with methionine) at position 29 (consensus sequence); these two amino acids are highly conserved in disaccharide UGTs and monosaccharide UGTs, respectively. Three disaccharide glycoside-producing UGTs from *Glycine max* (GmSap3Glu2″Gat, GmSap22A3″Xt and GmSap22A3″Xt) also occupy an intermediate position similar to UGT92G6 as they carry a phenylalanine and leucine at position 25 and 29, respectively. They lack the conserved Met25 but show the conserved Leu29 of disaccharide UGTs. Furthermore, the grapevine protein UGT92G6 lacks proline (replaced with cysteine) at position 90 (consensus sequence), which is conserved in disaccharide UGTs but contains a GSS motif at position 555–557 (consensus sequence), a special feature of monoglucoside UGTs and the three UGTs from *G. max* (Supplementary Fig. S3).

### Substrate specificity of recombinant UGT92G6 indicates preference towards specific phenolics and monoterpenyl glucosides

To determine whether *UGT92G6* encodes a functional glycosyltransferase, the open reading frame (ORF) of the gene was cloned into a glutathione *S*-transferase (GST) fusion vector for expression in *E. coli*. The recombinant protein was assayed for glucosylation activity with a variety of substrates. They can be grouped into aglycone and glucoside substrates, and have been found in grapes and wines or are important flavor molecules. The aglycone substrates examined were carvacrol, thymol, menthol, vanillin, eugenol, caffeic acid, *p*-coumaric acid, raspberry ketone, 2-phenylethanol, resveratrol, geraniol, citronellol, linalool, terpineol, farnesol, maltol, furaneol, hexanol, octanol and decanol, whereas glucoside substrates were geranyl glucoside, citronellyl glucoside, carvacryl glucoside, thymyl glucoside, vanillyl glucoside, octyl glucoside and decyl glucoside. Products were analyzed by LC-UV-MS (liquid chromatography coupled with UV and mass spectrometry detection). All substrates were also incubated with negative (empty vector) controls. No glucoside was formed in the control experiments. The results showed that UGT92G6 exhibited glucosylation activity towards specific phenolic compounds (**Figs. 2A, 3A**), namely caffeic acid, *p*-coumaric acid, carvacrol, thymol, eugenol, vanillin, raspberry ketone and 2-phenylethanol. However, it did not show enzyme activity towards the phenolic compound resveratrol. In addition, monoterpenols, maltol, furaneol, farnesol and alkanols were not glucosylated. Remarkably, UGT92G6 showed activity towards monoterpenyl

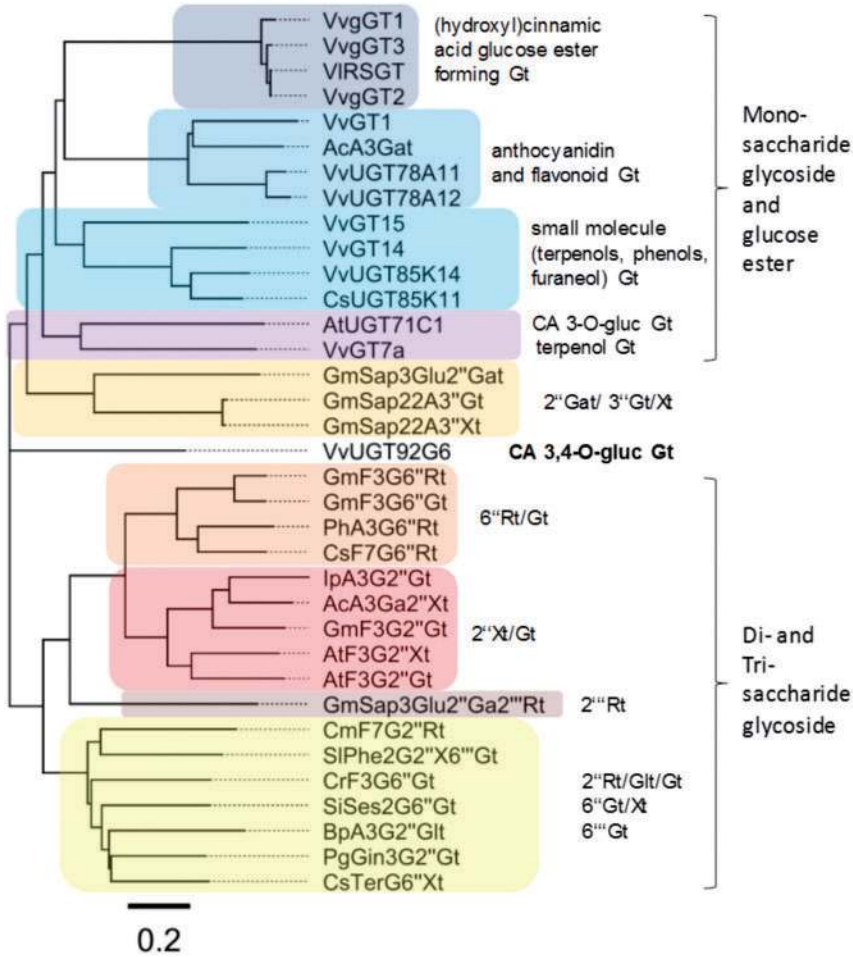


**Fig. 1** Phylogenetic analysis of functionally characterized UGTs. The bar represents 0.2 amino acid substitutions per site. The GenBank accession numbers for the sequences are shown in parentheses: Monosaccharide glycoside-forming glycosyltransferases: AcA3Gat (ADC34700); CsUGT85K11 (BAO51834); VvGt1 (gi|261260083); VvGT7a (XP_002276546); VvGT14 (XP_002285770); VvGT15 (XP_010650963); VvUGT85K14 (gi|225468660); VvUGT78A11 (CAN74919); VvUGT78A12 (BAI22847); VvgGT1 (gi|363805186); VvgGT2 (gi|363805188); VvgGT3 (gi|363805190); VlRSGT (gi|110932098); di-/trisaccharide glycoside-forming glycosyltransferases: AcA3Ga2"Xt (FG404013); AtF3GT2"Xt (gi|75311632); AtF3G2"GT (XP_002866013); BpA3G2"Glt (Q5NTH0); CmF7G2"Rt (gi|378405177); GmSap3Glu2"Gat (D4Q9Z4); GmSap3Glu2"Ga2'"Rt (D4Q9Z5); GmSap22A3'"Gt (BAM29362); GmSap22A3"Xt (BAM29363); CrF3G6"Gt (gi|242345159); CsiF7G6"Rt (gi|75265643); CsTerG6"Xt (BAO51835); GmF3G6"Rt (BAN91401); GmF3G6"Gt (BAV56172); GmF3G2"Gt (BAR88077); IpA3G2"Gt (gi|62857206); PgGin3G2"Gt (AKA44579); PhA3G6"Rt (gi|397567); SiSes2G6"Gt (BAF99027); SlPhe2G2"X6'"Gt (AGO03777). Gt, glucosyl-; Gat, galactosyl-; Rt, rhamnosyl-; Xt, xylosyl-; Glt, glucuronosyltransferase. Ac, *Actinidia chinensis*; At, *Arabidopsis thaliana*; Bp, *Bellis perennis*; Cm, *Citrus maxima*; Cr, *Catharanthus roseus*; Cs, *Camellia sinensis*; Csi, *Citrus sinensis*; Gm, *Glycine max*, Ip, *Ipomoea purpurea*; Pg, *Panax ginseng*; Ph, *Petunia hybrida*; Si, *Sesamum indicum*; Sl, *Solanum lycopersicum*; Vv, *Vitis vinifera*.

glucosides, namely geranyl glucoside and citronellyl glucoside (**Fig. 2**), but was inactive towards glucosides derived from phenolic compounds and alkanols.

## UGT92G6 displays both mono- and diglucosyltranferase activities

Full-scan and tandem MS (MS2) mass spectral data were used to characterize the reaction products of the UGT92G6 enzyme further. Carvacryl glucoside and thymyl glucoside were detected predominantly as the corresponding formate adduct, *m/z* 357 $[M + HCOO]^-$, in full-scan mode (**Fig. 2B**). UGT92G6 displayed high activity towards carvacrol but showed much lower enzyme activity towards the isomeric thymol (**Fig. 2A**). Eugenyl glucoside was also detected as the corresponding formate adduct (*m/z* 371) in full-scan mode (**Fig. 2B**).

Two isomeric glucosides, namely caffeoyl 3-*O*-glucoside and caffeoyl 4-*O*-glucoside, were formed from caffeic acid in almost equal amounts (**Fig. 3A**). Caffeic acid glucose ester could be excluded as a potential product because the products were not hydrolyzed under alkaline conditions. Both caffeoyl 3-*O*-glucoside and caffeoyl 4-*O*-glucoside were detected at *m/z* 341 $[M-H]^-$ in full-scan mode (**Fig. 3B, C**). MS2 identified the expected product ion at *m/z* 179 $[M\text{-}C_6H_{10}O_5\text{-}H]^-$ corresponding to caffeic acid. An authentic standard demonstrated that product 2 (**Fig. 3A, C**) is caffeoyl 3-*O*-glucoside and, consequently, product 1 is caffeoyl 4-*O*-glucoside.

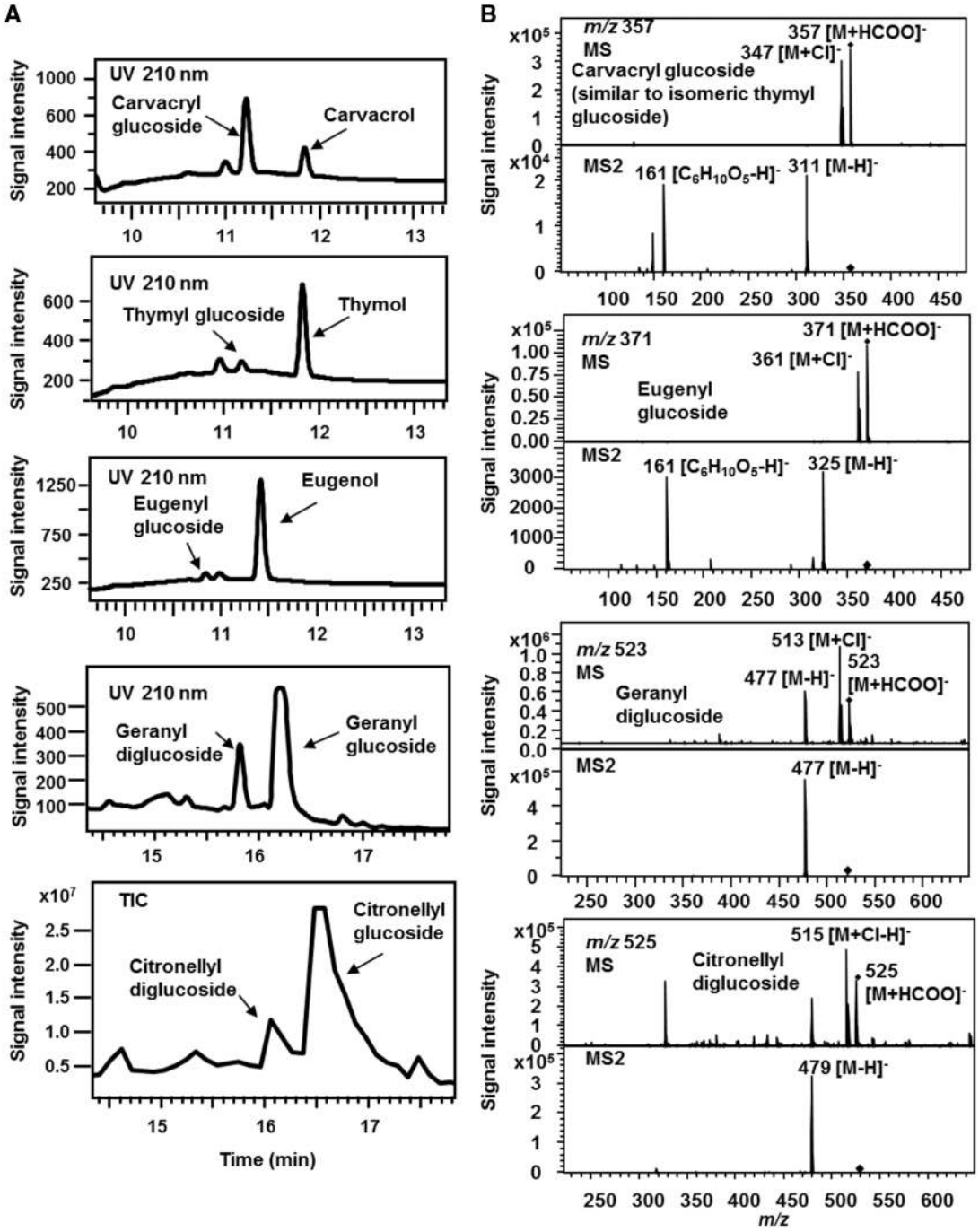


**Fig. 2** LC-UV-MS analysis of glucosides produced by recombinant UGT92G6. Chromatograms (A) and mass spectra (MS and MS2) of the products (B) are shown. Substrates are carvacrol, thymol, eugenol, geranyl glucoside and citronellyl glucoside. TIC, total ion chromatogram; UV, ultraviolet detection.

UGT92G6 also displayed diglucosyltransferase activity towards monoterpene glucosides, such as geranyl glucoside and citronellyl glucoside (**Fig. 2A**). Geranyl diglucoside and citronellyl diglucoside were detected as the corresponding formate adducts at *m/z* 523 $[M + HCOO]^-$ and *m/z* 525 $[M + HCOO]^-$ in full-scan mode (**Fig. 2B**), respectively. UGT92G6 displayed higher enzyme activity towards geranyl glucoside than towards citronellyl glucoside. The regioselectivity of this biocatalytic reaction was not further analyzed.

## Kinetic parameters of recombinant UGT92G6 displayed high specificity towards caffeic acid followed by carvacrol and geranyl glucoside

The enzymatic activity of UGT92G6 was analyzed at different pH values (pH 5.5–9) and temperatures (4–60°C). UGT92G6 enzyme worked best at pH 8 and 30°C. These values were used to determine the kinetics of UGT92G6 towards caffeic acid, carvacrol and geranyl glucoside. (**Table 1**). The $K_m$ values for

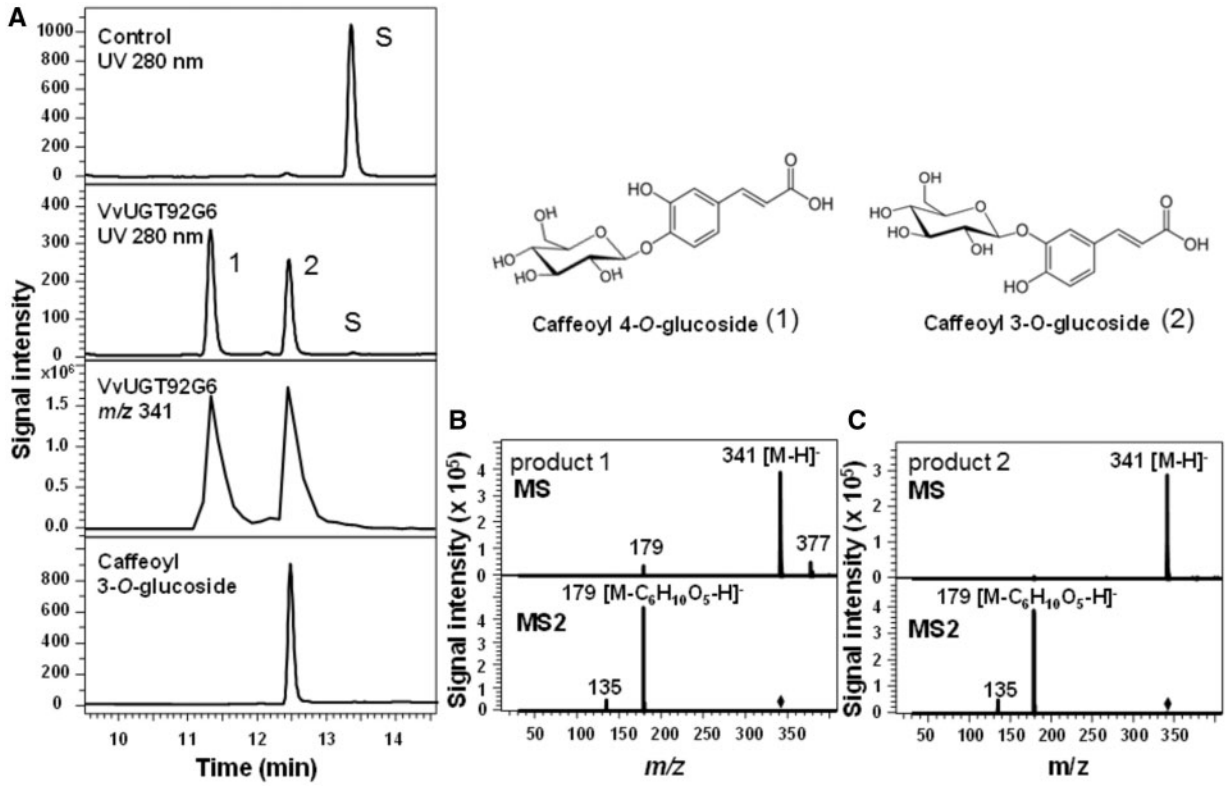


**Fig. 3** Identification of enzymatically formed products catalyzed by recombinant UGT92G6 from caffeic acid (substrate S) using LC-UV-MS. (A) Ultraviolet detection (UV) at 280 nm of the substrate and the products, MS trace at *m/z* 341 of the products and UV (280 nm) detection of the reference caffeoyl 3-*O*-glucoside. (B) MS and MS2 spectra of product 1. (C) MS and MS2 spectra of product 2. Product 1, caffeoyl 4-*O*-glucoside; product 2, caffeoyl 3-*O*-glucoside.

**Table 1** Kinetic properties of purified recombinant UGT92G6 enzyme

| Substrates | $K_m$ (μM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($M^{-1}$ $s^{-1}$) |
|---|---|---|---|
| Caffeic acid | 375.50 ± 33.12 | 1.19 ± 0.10 | 3169 |
| Carvacrol | 215.81 ± 33.27 | 0.09 ± 0.00 | 417 |
| Geranyl glucoside | 331.58 ± 63.15 | 0.01 ± 0.00 | 30 |

Substrate concentrations were varied from 10 to 3,000 μM. Values are the means of three separate determinations (see Supplementary Fig. S2)

caffeic acid, carvacrol and geranyl glucoside were within a comparable range of 376, 216 and 332 μM, respectively (**Table 1**). UGT92G6 exhibited the highest specificity towards caffeic acid ($k_{cat}/K_m$ 3,169 $M^{-1}$ $s^{-1}$), followed by carvacrol ($k_{cat}/K_m$ 417 $M^{-1}$ $s^{-1}$) and geranyl glucoside ($k_{cat}/K_m$ 30 $M^{-1}$ $s^{-1}$).

## Conversion of caffeic acid by UGT92G6 in *N. benthamiana* leaves in planta and leaf extracts by enzymatic assays

In order to investigate the caffeoyl glucoside products produced by UGT92G6 in planta, *UGT92G6* was transiently overexpressed in *N. benthamiana* leaves using an established viral vector system (Marillonnet et al. 2004, Marillonnet et al. 2005). As controls, untreated wild-type leaves were used, and an empty vector was infiltrated. The leaves were harvested 7 d after agroinfiltration. First, leaf protein extracts were prepared and the resulting crude protein suspension subjected to an in vitro assay offering caffeic acid as the sugar acceptor substrate. *UGT92G6*-infiltrated tobacco leaf extract was able to convert caffeic acid to two isomeric products, namely caffeoyl 3-*O*-glucoside and caffeoyl 4-*O*-glucoside. As the glucosides were not produced in the reaction mixtures containing wild-type leaf and control leaf extracts (**Fig. 4A**), the successful expression of *UGT92G6* in tobacco was confirmed.

Secondly, the wild-type leaves as well as the infiltrated leaves were directly subjected to qualitative and quantitative metabolite analysis by LC-UV-MS. In contrast to the in vitro leaf extract assay, only a high level of caffeoyl-3-*O*-glucoside, but not of caffeoyl-4-*O*-glucoside, was detected in the leaves infiltrated with *UGT92G6* (**Fig. 4B**). The concentration was about 8- and 24-fold higher in the agroinfiltrated leaves than in leaves infiltrated with the empty vector control and in untreated leaves, respectively (**Fig. 4C**). Since carvacrol, eugenol and

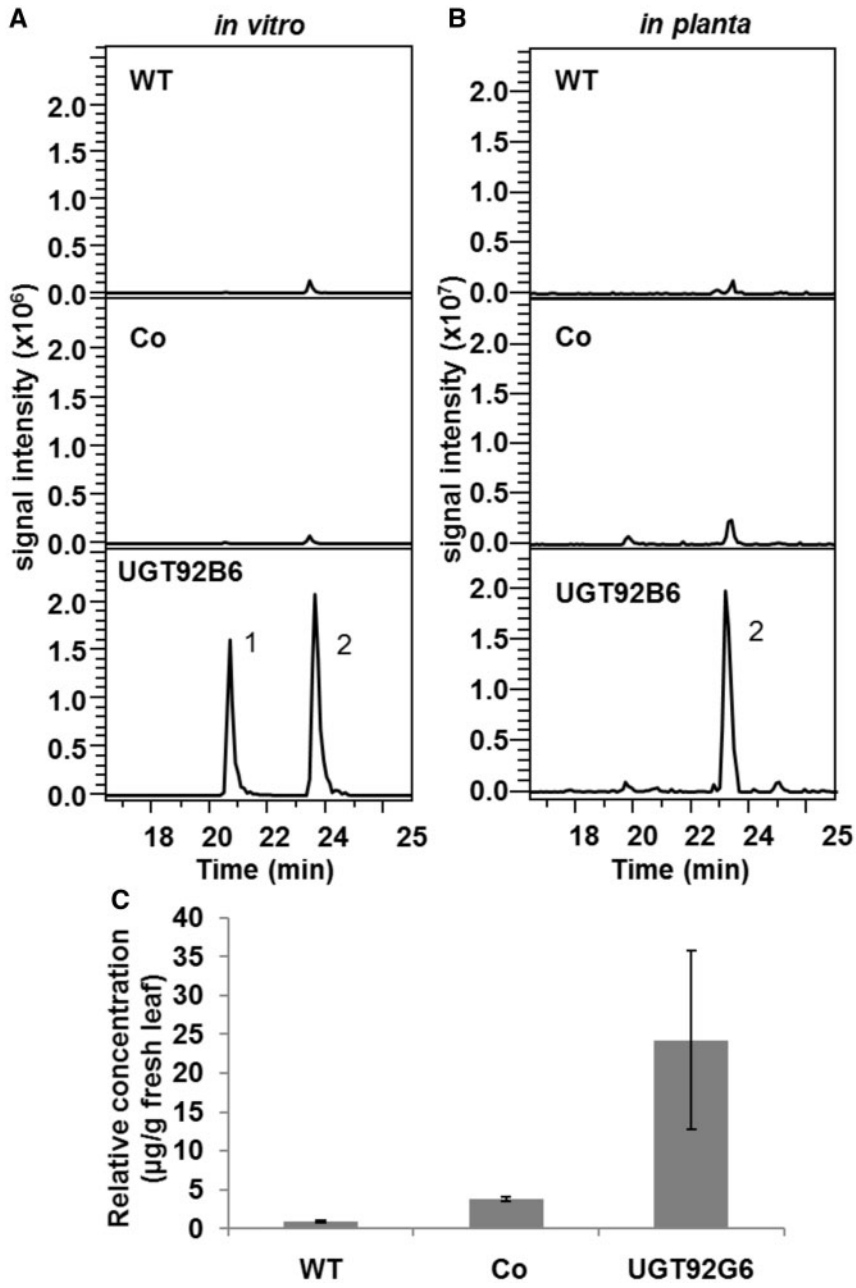


**Fig. 4** Transient overexpression of *UGT92G6* in *N. benthamiana*. Metabolites produced in enzyme activity assays and glycosides present in infiltrated leaves were analyzed by LC-UV-MS. (A) Protein extracts of untreated leaves (WT), and leaves infiltrated with an empty vector control (Co) and *UGT92G6* were incubated with caffeic acid in an in vitro assay. (B) Qualitative metabolite analysis of caffeoyl glucoside formation in untreated leaves (WT) and leaves infiltrated with an empty vector control (Co) and *UGT92G6*. Caffeoyl glucoside products were detected at *m*/*z* 341. Product 1, caffeoyl 4-*O*-glucoside; product 2, caffeoyl 3-*O*-glucoside. (C) Quantitative metabolite analysis of caffeoyl-3-*O*-glucoside formation in untreated leaves (WT), and leaves infiltrated with an empty vector control (Co) and *UGT92G6*. Levels of caffeoyl-3-*O*-glucoside are presented as the mean ± SEM. All leaves were analyzed 7 d after agroinfiltration. Quantification of caffeoyl-3-*O*-glucoside was performed by LC-UV-MS analysis as microgram per gram of the fresh weight with the help of the internal standard, 4-methylumbelliferyl β-D-glucuronide (relative concentration).

2-phenylethanol were transformed by *UGT92G6* in vitro, we also looked for carvacryl, eugenyl and 2-phenylethyl glucoside in the *UGT92G6*-infiltrated plants, but no such glucosides were observed. Carvacrol, eugenol and 2-phenylethanol were also not detected in *N. benthamiana* leaves by LC-UV-MS analysis.

## Conversion of geranyl glucoside by UGT92G6 in *N. benthamiana* leaves in planta and leaf extracts by enzymatic assays

Next, the products formed by UGT92G6 from geranyl glucoside were investigated using infiltrated leaf extracts as the enzyme

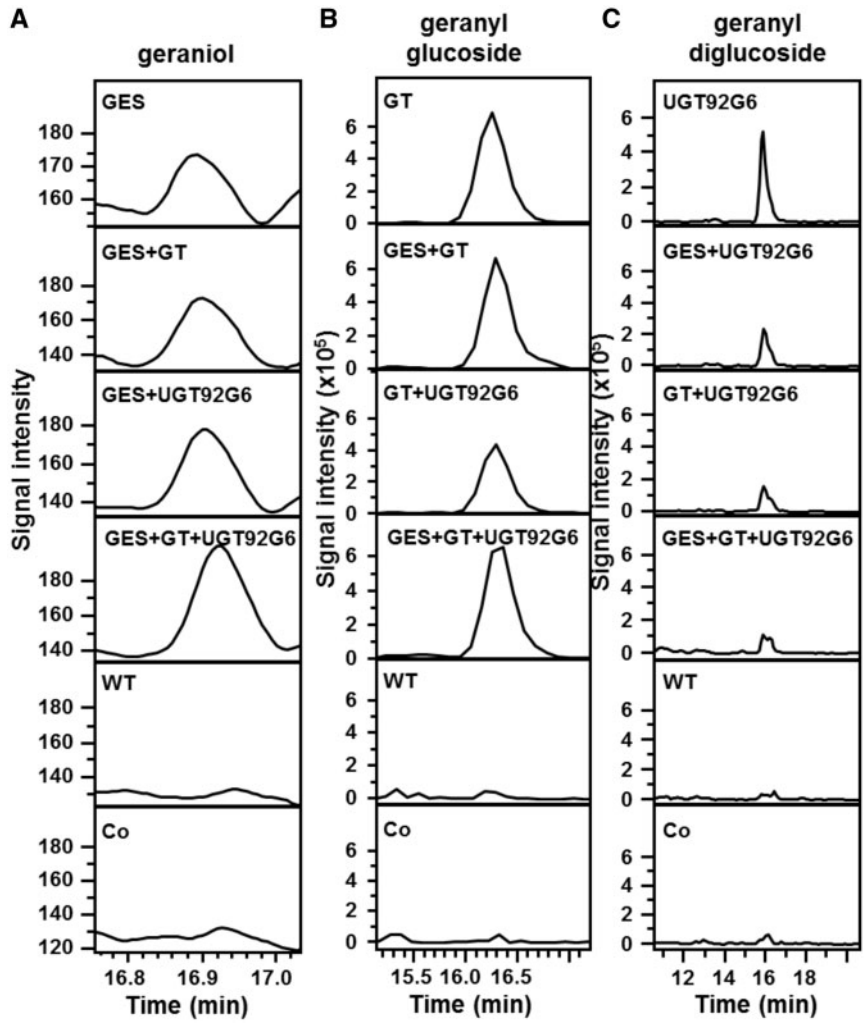


**Fig. 5** In vitro enzyme activity assay containing extracts of untreated leaves (WT), and leaves agroinfiltrated with *GES*, *VvGT14a* (GT) and *UGT92G6* for the detection of geraniol, geranyl monoglucoside and geranyl diglucoside. Analysis of (A) geraniol synthase activity in the extracts using geranyl diphosphate as a substrate; geraniol was detected at UV 210 nm; (B) geranyl monoglucosyltransferase activity using geraniol and UDP-glucose as substrates; geranyl glucoside was detected at *m*/*z* 361 and (C) geranyl diglucosyltransferase activity using geranyl glucoside and UDP-glucose as substrates; geranyl diglucoside was detected at *m*/*z* 523. Crude proteins were extracted from the leaves 7 d after agroinfiltration.

source (in vitro) (**Fig. 5**) as well as in leaves in planta (**Fig. 6**). Again, wild-type leaves and leaves infiltrated with an empty vector served as controls. The samples were harvested 7 d after agroinfiltration. As recombinant UGT92G6 readily produced geranyl diglucoside in vitro, this product was qualitatively and quantitatively analyzed in *N. benthamiana*. To provide a sufficient level of geranyl glucoside in the tobacco leaves, its synthetic pathway was reconstituted from geranyl diphosphate (GPP) in planta by transient co-expression of genes for geraniol synthase (*GES*) and a known geraniol glucosyltransferase gene from *V. vinifera VvGT14a* (*GT*). GES catalyzes the formation of geraniol from geranyl diphosphate (Dong et al. 2013), and VvGT14a produces geranyl glucoside from geraniol (Bönisch et al. 2014a).

To test whether the infiltrated genes were successfully expressed in *N. benthamiana*, leaf protein extracts were prepared from all samples, and the enzyme activities of UGT92G6, GES and VvGT14a were evaluated in vitro. The leaves infiltrated with *GES* and empty vector (control), as well as the wild type were tested for GES activity (**Fig. 5A**). As a result, geraniol was detected in the reaction mixtures containing *GES*-treated leaf extracts (**Fig. 5A**) but not in the reactions containing control and wild-type leaf extracts. The activity of geraniol glucosyltransferase and diglucosyltransferase was also examined in the leaves infiltrated with *VvGT14a* and *UGT92G6*, respectively. The result showed that geraniol was converted to geranyl monoglucoside by leaves infiltrated with *VvGT14a* (**Fig. 5B**), and geranyl glucoside was transformed to geranyl diglucoside by leaves infiltrated with *UGT92G6* (**Fig. 5C**). In contrast, no terpenoids were detected in the reactions containing control and wild-type leaf extracts. Taken together, *GES*, *VvGT14a* and *UGT92G6* were successfully expressed in *N. benthamiana*. Consequently, the biosynthetic pathways of geraniol and geranyl glucoside were successfully reconstituted in *N. benthamiana*.

Co-expression of *UGT92G6* with the other two genes led to a decreased level of geranyl diglucoside (**Fig. 5C**). In general, co-

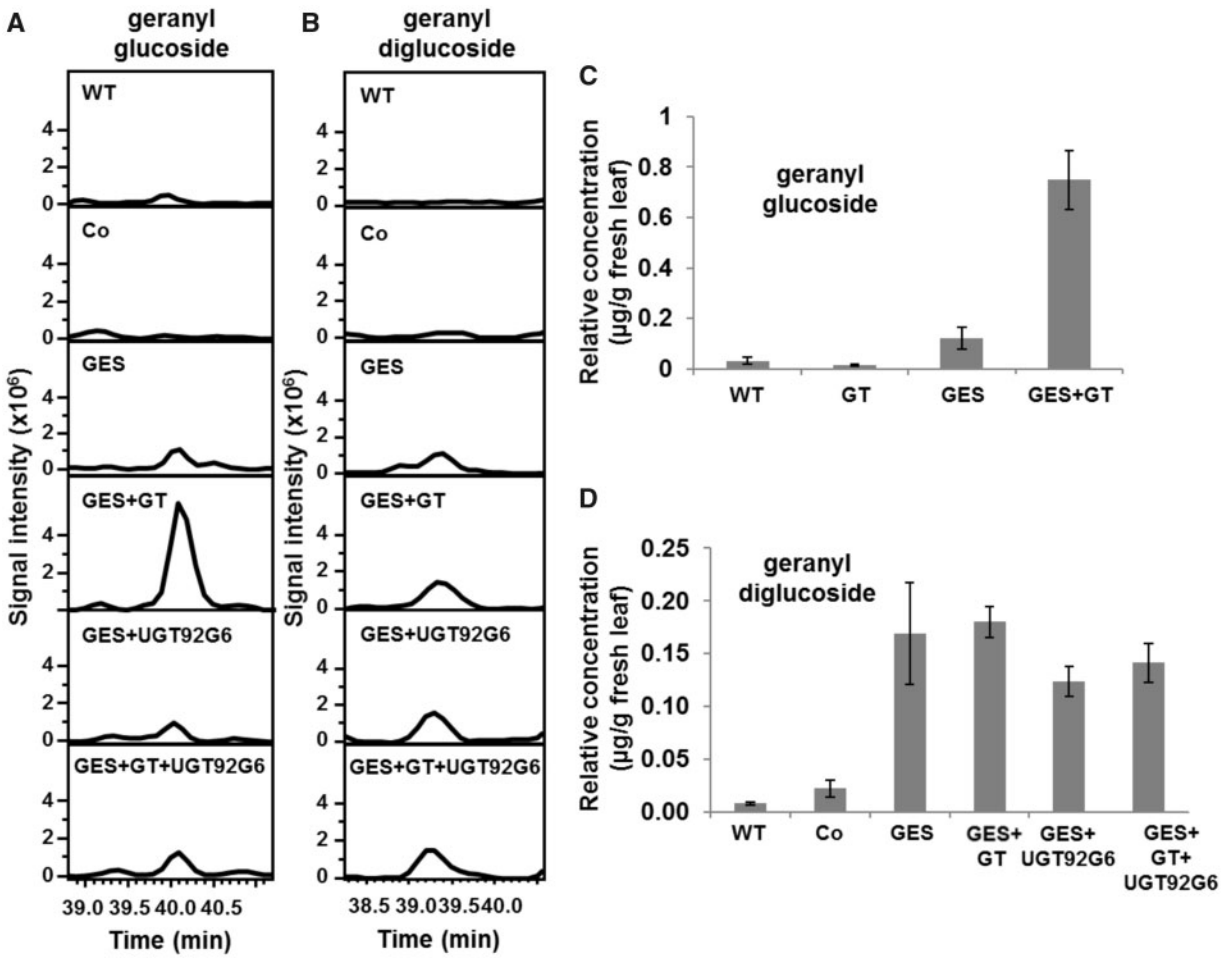


**Fig. 6** Metabolite analysis and quantification of geranyl mono- and diglucoside in *N. benthamiana* leaves in planta by LC-UV-MS. (A) Qualitative analysis of geranyl glucoside formation in untreated leaves (WT), and leaves agroinfiltrated with an empty vector control (Co), with geraniol synthase (GES), with *GES* and *VvGT14a* (GES + GT), with *GES* and *UGT92G6* and with *GES*, *GT* and *UGT92G6*. Geranyl glucoside was monitored at *m*/*z* 361. (B) Qualitative analysis of geranyl diglucoside formation. Geranyl diglucoside was monitored at *m*/*z* 523. (C) Quantitative analysis of geranyl glucoside formation. (D) Quantitative analysis of geranyl diglucoside formation. All leaves were analyzed 7 d after agroinfiltration. Data (C and D) are presented as the mean ± SEM. The quantity of geranyl glucoside and geranyl diglucoside is calculated as microgram per gram of the fresh weight, with the help of the internal standard 4-methylumbelliferyl β-D-glucuronide (relative concentration).

infiltration of several genes in tobacco using the viral vector system reduces the expressed protein level of each gene. As the enzyme activity of UGT92G6 against geranyl glucoside is very low already, the negative impact of the co-infiltration with other genes on its activity becomes more visible.

Finally, the levels of geranyl mono- and diglucoside were assessed in planta by extraction of the metabolites from infiltrated and control leaves (**Fig. 6**). Qualitative analysis showed that co-expression of *GES* with the confirmed geraniol UGT *VvGT14a* led to an increased level of geranyl glucoside (**Fig 6A**). This translates to a 6- and 54-fold higher level of geranyl glucoside in leaves treated with both *GES* and *VvGT14a* than with only *GES* or *VvGT14a*, respectively (**Fig. 6C**). The result verified the functional expression of *VvGT14a* in *N. benthamiana* and the biotransformation of geraniol into geranyl glucoside, confirming the monoglucosyltransferase activity of VvGT14a in planta.

Geranyl diglucoside was detected in all leaves infiltrated with *GES*, in *GES* single infiltration, *GES* in combination with *VvGT14a*, *GES* in combination with *UGT92G6* and *GES* in combination with *VvGT14a* and *UGT92G6*, but not in the control leaves (**Fig. 6B**). Geranyl diglucoside was also not detected in leaves treated with *UGT92G6* alone or in leaves treated with both *VvGT14a* and *UGT92G6* (data not shown). This result pointed to the absence of geraniol in native tobacco leaves, which prevented the formation of geranyl glucoside and subsequently geranyl diglucoside. It also indicated the presence of endogenous tobacco diglucosyltransfease enzymes capable of glycosylating geranyl glucoside in planta. Co-infiltration of *GES* with *VvGT14a* and/or *UGT92G6* did not further increase the level of geranyl diglucoside of 0.12–0.17 µg $g^{-1}$ fresh leaf that was produced when only *GES* was overexpressed (**Fig. 6D**). Thus, UGT92G6 did not significantly contribute to the formation of geranyl diglucoside in *N. benthamiana*.

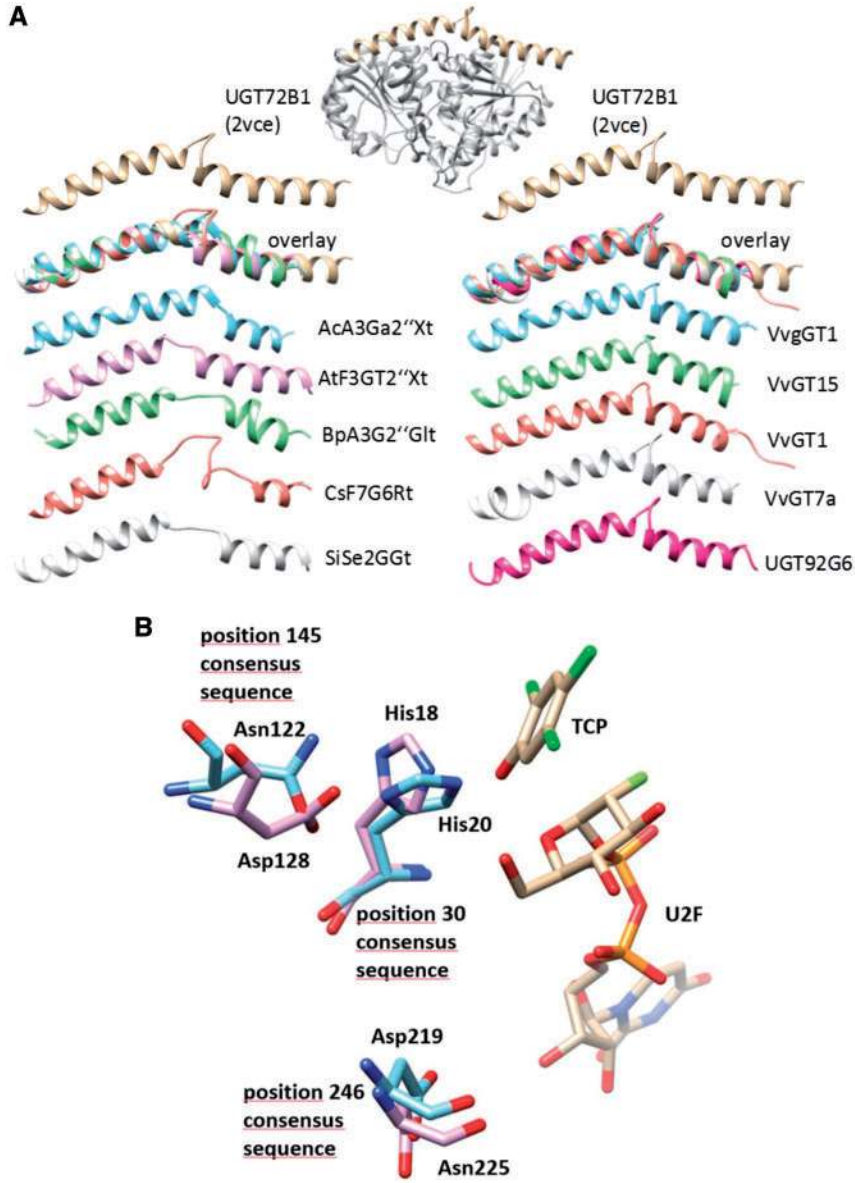


**Fig. 7** 3D protein structure (ribbon diagram) of UGT72B1 (top; 2vce from https://www.rcsb.org/pdb/home/home.do) (A). Both backbone helices are highlighted in gold. Excised backbone helices of the calculated structures (SWISS-MODEL; https://swissmodel.expasy.org/) of diglucosyltransferases (left; AcA3Ga''Xt, AtF3GT2''Xt, BpA3G2''Glt, CsF7G6Rt, SiSe2GGT) and monoglucosyltransferases (right; VvgGT1, VvGT15, VvGT1, VvGT7a) including the candidate UGT92G6. (B) Amino acids in the calculated active sites of UGT72B1 (pink) and VvgGT1 (light blue). U2F, UDP-2-deoxy-2-fluoroglucose; TCP, trichlorophenol.

## 3-D structural analysis of UGT92G6 corroborates its distinct enzymatic activities

Finally, the 3-D protein structure of UGT92G6 was generated by homology modeling and compared with the tertiary structure of known glucoside- and diglucoside-forming UGTs to investigate the significance of the observed sequence differences. The analysis revealed that the GSS motif at position 555–557 (consensus sequence) is apparently a striking feature of the monoglucosyltransferases (Supplementary Fig. S3) and separates the two backbone helices (**Fig. 7A;** highlighted in gold). This motif folds into a well-defined loop structure, whereas sequences lacking the GSS motif (amino acid sequences of diglucosyltransferases) form various 3-D geometries (**Fig. 7A**). It should be mentioned that three disaccharide glycoside-producing UGTs from *G. max* (GmSap3Glu2''Gat, GmSap22A3''Xt and GmSap22A3''Xt) also showed the GSS motif but their monoglycoside-forming activity has not been tested yet.

Furthermore, protein sequence comparison of biochemically characterized caffeoyl 3-*O*- and 4-*O*-glucoside UGTs (AtUGT71C1 and UGT92G6) with phenylpropanoid glucose ester-forming UGTs (Supplementary Fig. S4) confirmed the functional significance of the GSS motif (consensus sequence position 503–505) and revealed amino acid exchanges, such as at position 145 (aspartate/asparagine, consensus sequence) and position 246 (asparagine/aspartate) (**Fig. 7B**). Histidine (consensus sequence

position 30) is considered crucial for catalysis as it deprotonates the acceptor substrate enabling it to undergo nucleophilic attack at the donor substrate (UDP-glucose). The deprotonation of histidine is subsequently stabilized through interaction with aspartate (consensus sequence 145) in *O*-glucoside-forming UGTs (Offen et al. 2006). In phenylpropanoid ester-forming UGTs, stabilization by aspartate seems to be less important. Interestingly, at a distance of 10 Å from the catalytically active histidine, a reverse amino acid exchange (asparagine/aspartate position 246) was observed (Supplementary Fig. S4; **Fig 7B**). This exchange might be important for the reverse orientation of the phenylpropanoid acid substrate in the active site.

## Discussion

Glycosyltransferases have many functions in cells and constitute a large family of >100 genes in plant genomes each containing a consensus sequence enabling their easy identification. With more and more plant genome sequences becoming available, the substrate specificities of the encoded enzymes are now beginning to be established. Although only a few crystal structures of plant UGTs have been published, the high 3-D structural conservation of the plant UGTs render comparative sequence analysis and homology modeling promising tools to explain and predict substrate preference of unknown UGTs.

### Comparative sequence and 3-D structure analysis: GSS motif and asparagine/aspartate exchange corroborate the biochemical properties of UGT92G6

The calculated 3-D structure of UGT92G6 shows the typical GT-B fold consisting of two separate Rossmann domains connected by a linker region (**Fig. 7**; Breton et al. 2006, Breton et al. 2012). The catalytic site is located between the domains. Variations in UGT sequences are pronounced in the N-terminal domains, which accommodate very different acceptors, except for the amino acids around the catalytically active histidine (Supplementary Fig. S4). The C-terminal domain is highly conserved and interacts with the donor substrate, the nucleotide-sugar (Offen et al. 2006). A peptide motif characteristic of the plant GT-B fold, the PSPG (plant secondary product glycosyltransferase) box (position 433–476 in the consensus sequence; see Supplementary Fig. S3), is responsible for nucleotide-sugar binding (Offen et al. 2006, Jánváry et al. 2009). As with other GT-B structures, the final two C-terminal helices cross from the C-terminal domain to complete the N-terminal domain fold (**Fig. 7**). In the calculated 3-D structures of monoglucosyltransferases, the helices are separated by a well-defined loop structure, whereas diglucosyltransferases showed heterogenous geometries in the linker region connecting both helices. Amino acid sequence analysis revealed that the well-defined loop structure can be attributed to a conserved GSS motif in monoglucosyltransferases (consensus sequence positions 555–557 and 503–505 in Supplementary Figs. S3 and S4, respectively). The significance of the sequence following the PSPG box for enzymatic activity has already been emphasized, as truncation at the C-terminus rendered an anthocyanin 5-*O*-glucosyltransferase inactive (Jánváry et al. 2009). Since the C-terminus is situated at the opposite side of the cleft forming the active site, it was proposed that the two helices are important for stabilizing the spatial structure of the deep cleft by fixing the acceptor and donor domains.

Orthologous proteins of UGT92G6 were identified in 30 plant species (model and crop plants; Supplementary Figs. S5, S6) and a phylogenetic analysis was performed to understand the relationship between them. Sequence identity ranged from 41.4% (*A. thaliana*) to 69.6% (*Prunus persica*). Amino acids are highly conserved around the catalytically active histidine and aspartate (positions 62 and 187, respectively; for the consensus sequence, see Supplementary Fig. S6) and in the PSPG box (position 438–481; for the consensus sequence, see Supplementary Fig. S6). In addition, all sequences show a $GSSXXXX(D/E)_2$ motif at position 556 (for the consensus sequence, see Supplementary Fig. S6). The high sequence identity of UGT92G6 orthologs indicates that their function is highly conserved during evolution. The most structurally similar UGT functionally characterized is AtUGT71C1 (26% amino acid identity; Supplementary Fig. S6).

In the calculated structure of UGT92G6, the attacking hydroxyl of the acceptor substrate [exemplified by trichlorophenol (TCP)] lies 2.7 Å from His18 (**Fig. 7**) identical to the distance determined in the crystal structure of VvGT1 (Offen et al. 2006). His18 (see Supplementary Fig. S4) has been shown to play the role of a catalytic base for the deprotonation of the hydroxyl to allow nucleophilic attack at the anomeric C1 center of the sugar donor (U2F). Protonation of His18 is subsequently stabilized through interaction with Asp128. However, in glucose ester-forming UGTs such as VvgGT1 (see Supplementary Fig. S4), Asp128 is replaced by Asn122 whereby stabilization of protonated histidine is reduced. As most of the phenylpropanoid acid molecules already exist in deprotonated form (p*Ka* 4.62) at physiological pH, an additional proton acceptor (aspartate) is probably no longer required. The replacement aspartate/asparagine was found in all biochemically characterized UGTs, which produce (hydroxyl)cinnmanic acid glucose ester (see Supplementary Fig. S4). Interestingly, the reverse substitution (asparagine/aspartate) was observed at position 246 (consensus sequence) deeper in the active site (**Fig. 7B**). Asp219 might determine the orientation of the phenylpropanoid acid substrate in the catalytic center through hydrogen bonds with the hydroxyl groups while exposing the carboxyl to the catalytic His20 and the anomeric C1 center of the UDP-sugar.

### Caffeic acid is the natural substrate of UGT92G6

Phenolic acids are one of the most important quality parameters of wine, and contribute to characteristics such as astringency and bitterness (Mendoza et al. 2011). They are hydroxylated derivatives of benzoic acids (e.g. gallic, syringic, vanillic and protocatechuic acids) and cinnamic acids (e.g. caffeic and *p*-coumaric acids) (Herrmann 1995), and accumulate as glycosylated conjugates. In this study, UGT92G6 was able to glucosylate caffeic acid to form two isomeric glucosides, namely

caffeoyl 3-*O*-glucoside and caffeoyl 4-*O*-glucoside, in vitro. Glucosylation of caffeic acid is known to lead to the production of three structurally different conjugates. Transfer of glucose to the carboxyl group yields 1-*O*-caffeoylglucose (Winter and Herrmann 1984), an energy-rich glucose ester, functioning as a transient intermediate in the formation of other metabolites (Mock and Strack 1993), and regioselective transfer to the 3- and 4-hydroxy group results in two glucosides, caffeoyl 3-*O*-glucoside and caffeoyl 4-*O*-glucoside, respectively (Lim et al. 2003). Although the enzymatic formation of caffeic acid glucose ester has been demonstrated by several authors (Lim et al. 2001, Lunkenbein et al. 2006, Khater et al. 2012, Schulenburg et al. 2016), enzymes and their coding genes involved in the production of caffeoyl 3-*O*- and and 4-*O*-glucoside have rarely been reported (Lim et al. 2003).

In planta, caffeoyl 3-*O*-glucoside was produced in the *UGT92G6*-infiltrated tobacco leaves, indicating that the substrate is available in the plant. However, caffeoyl 4-*O*-glucoside was not accumulated, although the latter was formed in vitro by UGT92G6. Metabolism of caffeoyl 4-*O*-glucoside, e.g. by a regiospecific β-glucosidase, might explain this observation. UGT92G6 was also found to glucosylate other phenolic compounds, such as carvacrol, eugenol and 2-phenylethanol in vitro, but the corresponding glucosides were not detected in *UGT92G6*-infiltrated tobacco leaves. Lack of phenols in the plant might be attributed to this outcome, which was confirmed by LC-MS-UV analysis of tobacco leaves. Overall, it can be concluded that caffeic acid is the natural substrate of UGT92G6.

Hydroxycinnamates such as caffeic acid are involved in the response of plants to pathogens and interaction with symbionts (Babst et al. 2014). However, there are still ambiguities about the role of glucoconjugated hydroxycinnamic acids for lignin construction (Le Roy et al. 2016). Down-regulation of the hydroxycinnamoyl-CoA *O*-methyltransferase gene in poplar caused an accumulation of caffeoyl 3-*O*-glucoside due to a redirection of the metabolic flux away from monolignol biosynthesis and thus lignin formation (Meyermans et al. 2000). Since *O*-glucosides of caffeic acid also accumulated when wild-type plants were fed with the dihydroxycinnamic acid, accumulation of the glucosides is generated by access of the free acid. Thus, it seems that *O*-glucosylation of caffeic acid is essential for sequestration rather than for lignin biosynthesis.

## UGT92G6 displays promiscuous activity for terpenyl glucosides in vitro

As geranyl disaccharide glucosides have been isolated and characterized in grapes (Williams et al. 1982), the presence of diglycosyltransferase genes in grapevine was expected. UGT92G6 glucosylated geranyl and citronellyl glucosides in vitro, albeit with low activity (**Table 1**). Extensive agroinfiltration experiments demonstrated that the minor enzymatic activity of UGT92G6 for terpenyl glucosides could not contribute to the formation of geranyl diglucoside in planta, even when a sufficient amount of geraniol and geranyl glucoside was provided. Thus, the diglucosyltransferase activity is a promiscuous function of UGT92G6, which does not result in the production of terpenyl diglucosides in *N. benthamiana*. However, the result emphasizes the validity of the comparative sequence approach as UGT92G6 differs from mono- and diglucoside-forming UGTs (see Supplementary Fig. S3) but combines both activities. Jensen (1976) already hypothesized that primordial proteins had to be highly promiscuous in order to build metabolic networks. However, enzyme promiscuity is not only a primordial trait, it is a very widespread characteristic of modern proteins including UGTs (Song et al. 2015, Dewitte et al. 2016). Unlike enzymes involved in primary metabolism, proteins of secondary metabolism are catalytically less efficient but have a larger mechanistic elasticity and show a broader substrate tolerance, which enables them to produce a large number of secondary metabolites (Sonawane et al. 2017). The fitness gain provided by one of the products and low selective pressure may allow the enzymes to sustain other activities even though they are not in use. In the case of UGTs, it may be advantageous to have enzymes available that might glucosylate toxic substrates.

Comparative sequence analysis is now a common approach to identifying and characterizing functional genes. However, this approach can have many pitfalls, and thus confirmation of results by wet-lab experiments is required. We have predicted the enzymatic activity of a UGT and were able to confirm the respective substrate preference. Determination of the natural role of the UGT finally led to the finding of amino acids crucial for substrate recognition and product formation. Our results are useful to improve the predictive power of comparative sequence analyses for UGTs and to speed up the functional assignment of new *UGT* genes.

# Materials and Methods

## In silico analysis

Nucleotide and amino acid sequence analyses were performed using Geneious (http://www.geneious.com/). The orthologous sequences of UGT92G6 from different model and crop plants were aligned using Geneious Alignment for obtaining a multiple sequence alignment file. A phylogenetic tree was constructed using the Neighbor–Joining method available in the Geneious program. Default values were applied. Visualization of protein models was accomplished with Chimera 1.11.2 (https://www.cgl.ucsf.edu/chimera/), and protein 3-D models were generated with SWISS-MODEL (https://swissmodel.expasy.org/) applying default values and using the crystallized structure of UGT72B1 (PDB 2VCE) as a template (Brazier-Hicks et al. 2007).

## Chemicals

Commercial chemicals geraniol, eugenol, thymol, carvacrol, caffeic acid and GPP were purchased in analytical grade from Sigma-Aldrich; caffeoyl 3-*O*-gucoside was purchased from Carbosynth; and geranyl glucoside and citronellyl glucoside (>95% purity by LC) were synthesized from geraniol and citronellol, respectively, using VvGT15c from *V. vinifera* (Bönisch et al. 2014a).

## Cloning and heterologous expression of *UGT92G6*

Total RNA was extracted from berry skin cDNA of *V. vinifera* cv. White Riesling 239-34 Gm following the cetyltrimethylammonium bromide method as previously described by Zeng and Yang (2002), adapted by Reid et al. (2006). The first-strand cDNAs were synthesized from 1 μg of total RNA using Superscript III RTase (Invitrogen) and a GeneRacer oligo(dT) primer [5′-GCT GTC AAC GAT ACG CTA CGT AAC GGC ATG ACA GTG $T_{(18)}$-3′]. The UGT92G6 (GSVIVT01031678001)

full-length ORF was amplified by PCR with the cDNA template using primers 5′-CGC GGA TCC ATG GGT TCT CAG CAG GAA CAT-3′ (forward) and 3′-ATA GTT TAG CGG CCG CTC AGT AAC CTT GTC TCT TGG A-5′ (reverse). The PCR products were digested and then ligated with the pGEX-4T1 vector (Novagen) to yield pGEX-4T1-*UGT92G6*. The recombinant genes were subjected to sequencing to confirm the sequence of the inserts. Furthermore, this construct was transformed into the BL21(DE3)pLysS cell strain for expression of recombinant protein. A 4 ml overnight culture was used to inoculate a 400 ml culture in LB medium containing appropriate antibiotics. Cultures were grown at 37°C until an $OD_{600}$ of 0.6 was reached. Expression of the protein was induced by the addition of 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and the cultures were grown at 18°C for an additional 20 h. Cells were harvested by centrifugation (5,000×*g*, 20 min, 4°C) and re-suspended in 10 ml of binding buffer which was suited for subsequent resin purification. Cells were lysed by sonication on ice with an MS 73 sonotrode (Bandelin electronic) 10 times for 30 s at 10% of maximal power. Cell debris was removed by centrifugation (12,000×*g*, 30 min, 4°C). Purification of the expression products was performed using a GST–glutathione resin (Novagen) according to the manufacturer's instructions. Proteins were analyzed by SDS–PAGE with 12% polyacrylamide gels and stained with Coomassie brilliant blue R-250. The protein concentration was determined by the Bradford assay (Bradford 1976).

## Enzyme assays and product identification using LC-UV-MS/MS

For the initial screening, the standard assay (200 µl) contained 50 µl of crude protein extract, 100 mM Tris–HCl buffer (pH 7.5), 500 µM UDP-glucose and 600 µM substrate. The reaction was initiated by the addition of UDP-glucose, incubated at 30°C with constant shaking for 60 min, and was stopped by heating at 75°C for 10 min. After 5 min of centrifugation at 13,400×*g* twice, the clear supernatant was used for LC-UV-MS analysis. Products were monitored using the diagnostic ions listed in Supplementary Table S1. Authentic reference glucosides were used for identification.

The LC system consisted of a quaternary pump and a variable wavelength detector, all from Agilent 1100 (Bruker Daltonics). The column was a LUNA C18. 100A 150 × 2 mm (Phenomenex). LC was performed with the following binary gradient system: solvent A, water with 0.1% formic acid; and solvent B, 100% methanol with 0.1% formic acid. The gradient program used was as follows: 0–7 min, 90% A/10% B to 50% A/50% B; 7–10 min, 50% A/50% B to 100% B, hold for 5 min; 15–20 min, 100% B to 90% A/10% B, hold for 10 min. The flow rate was 0.2 ml $min^{-1}$. Attached to the LC was a Bruker esquire 3000plus mass spectrometer with an ESI interface that was used to record the mass spectra. The ionization voltage of the capillary was 4,000 V and the end plate was set to −500 V.

## Determination of kinetic parameters

The pH optimum of UGT92G6 was determined using different buffers for various pH ranges: 50 mM citric acid (pH 5.5 and 6.0), 50 mM sodium phosphate (pH 6–8) and 50 mM Tris–HCl (pH 8–9) in 0.5 unit intervals. The optimal temperature was evaluated from 4 to 60°C at pH 7.5. The kinetic parameters were determined under optimum conditions and were calculated by three determinations using Microsoft Excel Solver (see Supplementary Fig. S2).

## Agroinfiltration of *GES*, *VvGT14a* and *UGT92G6* into *N. benthamiana*

The full-length ORFs of *GES*, *VvGT14a* and *UGT92G6* were amplified by PCR from plasmid DNA pMiS1-optRBS-*GES* (a generous gift from Professor Jens Schrader, DECHEMA, Frankfurt, Germany), *VvGT14a*-pGEX-4T1 and *UGT92G6*-pGEX-4T1, respectively. The PCR products were double digested with restriction enzymes, and then ligated with pICH11599 vector (Marillonnet et al. 2004, Marillonnet et al. 2005), which were digested with the same restriction enzymes as for the inserts to yield *GES*-pICH11599, *VvGT14a*-pIC11599 and *UGT92G6*-pICH11599. The recombinant genes were subjected to sequencing to confirm the sequence of the inserts.

*Agrobacterium* was used to deliver the various modules into plant cells (Marillonnet et al. 2004, Marillonnet et al. 2005). pICH17388, pICH14011, *GES*-pICH11599, *VvGT14a*-pIC11599 and *UGT92G6*-pICH11599 were separately transformed into the *Agrobacterium tumefaciens* strain AGL0 using the freeze–thaw technique, and integrity was confirmed by PCR. *Agrobacterium* strains carrying each pro-vector module were mixed and infiltrated into *N. benthamiana* using a syringe without a needle as described previously (Huang et al. 2010). *GES*, *VvGT14a* and *UGT92G6* were single or mixed infiltrated into *N. benthamiana*. Empty pICH11599 vector was also infiltrated into *N. benthamiana* as a negative control.

## Enzyme extraction from *N. benthamiana* leaves and assay

Aliquots of 100 mg FW samples of *N. benthamiana* leaves infiltrated with *Agrobacterium* were ground to fine powder in liquid nitrogen with a mortar and a pestle, followed by re-suspension in 300 µl of protein extraction buffer (50 mM sodium phosphate buffer, pH 7.5, 10 mM EDTA, 0.1% Triton X-100, 5 mM β-mercaptoethanol). The homogenate was centrifuged at 4°C, 16,000×*g* for 10 min to remove cell debris. Total protein content was determined by Bradford assay.

GES activity assay was carried out for functional characterization, using GPP as a substrate. Enzymatic assays were done in 0.3 ml of reaction buffer containing 50 mM Tris–HCl, 33 µM $MgCl_2$, 0.17 mM $MnCl_2$, 18 µM GPP and 100 µl of GES-infiltrated leaf extract. The reaction mixture was incubated at 30°C for 1 h. The reaction was stopped by adding 0.5 ml of hexane, mixing thoroughly by vortexing. The mixture was centrifuged at 13,400×*g* for 5 min, and the supernatant hexane phase was collected. The extraction was repeated with hexane (0.2 ml). Then the hexane phase was collected, evaporated to dryness, re-suspended in 100 µl of ethanol and analyzed by LC-UV-MS. An authentic geraniol standard was used for calibration, and detection was performed at UV 210 nm as geraniol is barely detectable by MS.

The standard assay (200 µl) for glycosyltransferase activity contained 50 µl of crude protein extract from *N. benthamiana*, 100 mM Tris–HCl buffer (pH 7.5), 500 µM UDP-glucose and 600 µM acceptor substrate. The reaction was initiated by the addition of UDP-glucose, incubated at 30°C with constant shaking for 1 h and was stopped by heating at 70°C for 15 min. After 10 min of centrifugation at 13,400×*g*, the clear supernatant was analyzed by LC-UV-MS.

## Metabolite analysis

Metabolites in individual *N. benthamiana* leaves were analyzed 7 d after agroinfiltration. The material was ground to fine powder in liquid nitrogen and kept at −80°C prior to analyses. For metabolite analysis, 400 mg of powder was used for each of the three biological replicates. For sample extraction, 4-methylumbelliferyl β-D-glucuronide (250 µl of a solution in methanol, 0.2 mg $ml^{-1}$) was added as an internal standard, yielding 50 µg of internal standard in each sample. After the addition of 250 µl of methanol, vortexing and sonication for 10 min, the sample was centrifuged at 13,400×*g* for 10 min. The supernatant was transferred to a new tube, and the residue was re-extracted with 500 µl of methanol. The supernatants were combined, dried in a vacuum concentrator and re-dissolved in 60 µl of 50% methanol. After 1 min of vortexing, 10 min of sonication and 10 min of centrifugation at 13,400×*g*, the clear supernatant was used for LC-UV-MS analysis. The gradient program used for LC was as follows: 0–30 min, 100% A to 50% A/50% B; 30–35 min, 50% A/50% B to 100% B, hold for 15 min; 50–55 min, 100% B to 100% A, hold for 10 min. The flow rate was 0.2 ml $min^{-1}$. In addition to qualitative identification based on fragmentation patterns and retention times in comparison with authentic reference substances, each compound was quantified by QuantAnalysis 2.0 (6300 Series Ion Trap LC/MS 6.2). All results were normalized against the internal standard 4-methylumbelliferyl β-D-glucuronide and expressed as µg equivalents $g^{-1}$ FW (relative concentration).

## Supplementary Data

Supplementary data are available at PCP online.

## Funding

This work was supported by the Alexander von Humboldt Foundation.

## Disclosures

The authors have no conflicts of interest to declare.

## References


Babst, B.A., Chen, H.Y., Wang, H.Q., Payyavula, R.S., Thomas, T.P. and Harding, S.A. (2014) Stress-responsive hydroxycinnamate glycosyltransferase modulates phenylpropanoid metabolism in *Populus*. *J. Exp. Bot.* 65: 4191–4200.

Bönisch, F., Frotscher, J., Stanitzek, S., Rühl, E., Wüst, M., Bitz, O., et al. (2014a) A UDP-glucose: monoterpenol glucosyltransferase adds to the chemical diversity of the grapevine metabolome. *Plant Physiol.* 165: 561–581.

Bönisch, F., Frotscher, J., Stanitzek, S., Rühl, E., Wüst, M., Bitz, O., et al. (2014b) Activity-based profiling of a physiologic aglycone library reveals sugar acceptor promiscuity of family 1 UDP-glucosyltransferases from grape. *Plant Physiol.* 166: 23–39.

Bradford, M.M. (1976) A rapid and sensitive for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal. Biochem.* 72: 248–254.

Brahmachari, G., Mandal, L.C., Roy, R., Mondal, S. and Brahmachari, A.K. (2011) Stevioside and related compounds—molecules of pharmaceutical promise: a critical overview. *Arch. Pharm. Chem. Life Sci.* 1: 5–19.

Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.K., Bowles, D.J., et al. (2007) Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc. Natl. Acad. Sci. USA* 104: 20238–20243.

Breton, C., Fournel-Gigleux, S. and Palcic, M.M. (2012) Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* 22: 540–549.

Breton, C., Snajdrová, L., Jeanneau, C., Koca, J. and Imberty, A. (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology* 16: 29R–37R.

Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S. and Martens, S. (2012) A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. *Plant J.* 69: 1030–1042.

De Roode, B.M., Franssen, M.C.R., van der Padt, A. and Boom, R.M. (2003) Perspectives for the industrial enzymatic production of glycosides. *Biotechnol. Prog.* 19: 1391–1402.

De Roode, B.M., Oliehoek, L., van der Padt, A., Franssen, M.C.R. and Boom, R.M. (2001) Downstream processing of enzymatically produced geranyl glucoside. *Biotechnol. Prog.* 17: 881–886.

Dewitte, G., Walmagh, M., Diricks, M., Lepak, A., Gutmann, A., Nidetzky, B., et al. (2016) Screening of recombinant glycosyltransferases reveals the broad acceptor specificity of stevia UGT-76G1. *J. Biotechnol.* 233: 49–55.

Di, S., Yan, F., Rodas, F.R., Rodriguez, T.O., Murai, Y., Iwashina, T., et al. (2015) Linkage mapping, molecular cloning and functional analysis of soybean gene Fg3 encoding flavonol 3-O-glucoside/galactoside (1→2) glucosyltransferase. *BMC Plant Biol.* 15: 126.

Dong, L., Miettinen, K., Goedbloed, M., Verstappen, F.W., Voster, A., Jongsma, M.A., et al. (2013) Characterization of two geraniol synthases from *Valeriana officinalis* and *Lippia dulcis*: similar activity but difference in subcellular localization. *Metab. Eng.* 20: 198–211.

Frydman, A., Liberman, R., Huhman, D.V., Carmeli-Weissberg, M., Sapir-Mir, M., Ophir, R.W., et al. (2013) The molecular and enzymatic basis of bitter/non-bitter flavor of citrus fruit: evolution of branch-forming rhamnosyltransferases under domestication. *Plant J.* 73: 166–178.

Frydman, A., Weisshaus, O., Bar-Peled, M., Huhman, D.V., Sumner, L.W., Marin, F.R., et al. (2004) Citrus fruit bitter flavors: isolation and functional characterization of the gene Cm1,2RhaT encoding a 1,2 rhamnosyltransferase, a key enzyme in the biosynthesis of the bitter flavonoids of citrus. *Plant J.* 40: 88–100.

Härtl, K., Huang, F.-C., Giri, A.P., Franz-Oberdorf, K., Frotscher, J., Shao, Y., et al. (2017) Glucosylation of smoke-derived volatiles in grapevine (*Vitis vinifera*) is catalyzed by a promiscuous resveratrol/guaiacol glucosyltransferase. *J. Agric. Food Chem.* 65: 5681–5689.

Herrmann, A. (2007) Controlled release of volatile compounds under mild reaction conditions: from nature to consumer products. *Angew. Chem. Int. Ed.* 46: 5938–5967.

Herrmann, K.M. (1995) The shikimate pathway: early steps in the biosynthesis of aromatic compounds. *Plant Cell* 7: 907–919.

Huang, F.-C., Studart-Witkowski, C. and Schwab, W. (2010) Overexpression of hydroperoxide lyase gene in *Nicotiana benthamiana* using a viral vector system. *Plant Biotechnol. J.* 8: 783–713.

Jánváry, L., Hoffmann, T., Pfeiffer, J., Hausmann, L., Töpfer, R., Fischer, T.C., et al. (2009) A double mutation in the anthocyanin 5-O-glucosyltransferase gene disrupts enzymatic activity in *Vitis vinifera* L. *J. Agric. Food Chem.* 57: 3512–3518.

Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* 30: 409–425.

Khater, F., Fournand, D., Vialet, S., Meudec, E., Cheynier, V. and Terrier, N.J. (2012) Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis. *J. Exp. Bot.* 63: 1201–1214.

Le Roy, J., Huss, B., Creach, A., Hawkins, S. and Neutelings, G. (2016) Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Front. Plant Sci.* 7: 735.

Lim, E.K., Higgins, G.S., Li, Y. and Bowles, D.J. (2003) Regioselectivity of glucosylation of caffeic acid by a UDP-glucose:glucosyltransferase is maintained *in planta*. *Biochem. J.* 373: 987–992.

Lim, E.K., Li, Y., Parr, A., Jackson, R., Ashford, D.A. and Bowles, D.J. (2001) Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in *Arabidopsis*. *J. Biol. Chem.* 276: 4344–4349.

Lunkenbein, S., Bellido, M., Aharoni, A., Salentijn, E.M., Kaldenhoff, R., Coiner, H.A., et al. (2006) Cinnamate metabolism in ripening fruit. Characterization of a UDP-glucose:cinnamate glucosyltransferase from strawberry. *Plant Physiol.* 140: 1047–1058.

Marillonnet, S., Giritch, A., Gils, M., Kandzia, R., Klimyuk, V. and Gleba, Y. (2004) In planta engineering of viral RNA replicons: efficient assembly by recombination of DNA modules delivered by *Agrobacterium*. *Proc. Natl. Acad. Sci. USA* 101: 6852–6857.

Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V. and Gleba, Y. (2005) Systemic *Agrobacterium tumefaciens*-mediated transfection of viral replicons for efficient transient expression in plants. *Nat. Biotechnol.* 23: 718–723.

Masada, S., Terasaka, K., Oguchi, Y., Okazaki, S., Mizushima, T. and Mizukami, H. (2009) Functional and structural characterization of a flavonoid glucoside 1,6-glucosyltransferase from *Catharanthus roseus*. *Plant Cell Physiol.* 50: 1401–1415.

Mastelic, J., Jerkovic, I., Vinkovic, M., Dzolic, Z. and Vikic-Topic, D. (2004) Synthesis of selected naturally occurring glucosides of volatile compounds. Their chromatographic and spectroscopic properties. *Croat. Chem. Acta* 77: 491–500.

Mendoza, L., Matsuhiro, B., Aguirre, M.J., Isaacs, M., Sotés, G., Cotoras, M., et al. (2011) Characterization of phenolic acids profile from chilean red wines by high-performance liquid chromatography. *J. Chil. Chem. Soc.* 56: 688–691.

Meyermans, H., Morreel, K., Lapierre, C., Pollet, B., De Bruyn, A., Busson, R., et al. (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A O-methyltransferase, an enzyme involved in lignin biosynthesis. *J. Biol. Chem.* 275: 36899–36909.

Mock, H.-P. and Strack, D. (1993) Energetics of the uridine 5-diphosphoglucosehydroxy-cinnamic acid acyl-glucosyltransferase reaction. *Phytochemistry* 32: 575–579.

Montefiori, M., Espley, R.V., Stevenson, D., Cooney, J., Datson, P.M., Saiz, A., et al. (2011) Identification and characterisation of F3GT1 and F3GGT1, two glycosyltransferases responsible for anthocyanin biosynthesis in red-fleshed kiwifruit (*Actinidia chinensis*). *Plant J.* 65: 106–118.

Morita, Y., Hoshino, A., Kikuchi, Y., Okuhara, H., Ono, E., Tanaka, Y., et al. (2005) Japanese morning glory dusky mutants displaying reddish-brown or purplish-gray flowers are deficient in a novel glycosylation enzyme for anthocyanin biosynthesis, UDP-glucose: anthocyanidin 3-O-glucoside-2″-O-glucosyltransferase, due to 4-bp insertions in the gene. *Plant J.* 42: 353–363.

Nakajima, T., Matsubara, K., Kodama, H., Kokubun, H., Watanabe, H. and Ando, T. (2005) Insertion and excision of a transposable element governs the red floral phenotype in commercial petunias. *Theor. Appl. Genet.* 110: 1038–1043.

Noguchi, A., Fukui, Y., Iuchi-Okada, A., Kakutani, S., Satake, H., Iwashita, T., et al. (2008) Sequential glucosylation of a furofuran lignan, (+)-sesaminol, by *Sesamum indicum* UGT71A9 and UGT94D1 glucosyltransferases. *Plant J.* 54: 415–427.

Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B.G., Tarling, C.A., et al. (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J.* 25: 1396–1405.

Ohgami, S., Ono, E., Horikawa, M., Murata, J., Totsuka, K., Toyonaga, H., et al. (2015) Volatile glycosylation in tea plants: sequential glycosylations for the biosynthesis of aroma β-primeverosides are catalyzed by two *Camellia sinensis* glycosyltransferases. *Plant Physiol.* 168: 464–477.

Reid, K.E., Olsson, N., Schlosser, J., Peng, F. and Lund, S.T. (2006) An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT–PCR during berry development. *BMC Plant Biol.* 6: 27.

Rojas Rodas, F., Di, S., Murai, Y., Iwashina, T., Sugawara, S., Mori, T., et al. (2016) Cloning and characterization of soybean gene Fg1 encoding flavonol 3-O-glucoside/galactoside (1→6) glucosyltransferase. *Plant Mol. Biol.* 92: 445–456.

Rojas Rodas, F., Rodriguez, T.O., Murai, Y., Iwashina, T., Sugawara, S., Suzuki, M., et al. (2014) Linkage mapping, molecular cloning and functional analysis of soybean gene Fg2 encoding flavonol 3-O-glucoside (1→6) rhamnosyltransferase. *Plant Mol. Biol.* 84: 287–300.

Sahari, M.A. and Asgari, S. (2013) Effects of plants bioactive compounds on foods microbial spoilage and lipid oxidation. *Food Sci. Technol.* 1: 52–61.

Sawada, S., Suzuki, H., Ichimaida, F., Yamaguchi, M.A., Iwashita, T., Fukui, Y., et al. (2008) UDP-glucuronic acid:anthocyanin glucuronosyltransferase from red daisy (*Bellis perennis*) flowers. Enzymology and phylogenetics of a novel glucuronosyltransferase involved in flower pigment biosynthesis. *J. Biol. Chem.* 280: 899–906.

Sayama, T., Ono, E., Takagi, K., Takada, Y., Horikawa, M., Nakamoto, Y., et al. (2012) The Sg-1 glycosyltransferase locus regulates structural diversity of triterpenoid saponins of soybean. *Plant Cell* 24: 2123–2138.

Schulenburg, K., Feller, A., Hoffmann, T., Schecker, J.H., Martens, S. and Schwab, W. (2016) Formation of β-glucogallin, the precursor of ellagic acid in strawberry and raspberry. *J. Exp. Bot.* 67: 2299–2308.

Schwab, W., Fischer, T.C., Giri, A. and Wüst, M. (2015) Potential applications of glucosyltransferases in terpene glucoside production: impacts on the use of aroma and fragance. *Appl. Microbiol. Biotechnol.* 99: 165–174.

Shibuya, M., Nishimura, K., Yasuyama, N. and Ebizuka, Y. (2010) Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max*. *FEBS Lett.* 584: 2258–2264.

Sonawane, P.D., Pollier, J., Panda, S., Szymanski, J., Massalha, H., Yona, M., et al. (2017) Plant cholesterogenesis is mediated by a multi-step pathway sharing enzymes with phytosterol metabolism. *Nat. Plants* 3: 17101.

Song, C., Gu, L., Liu, J., Zhao, S., Hong, X., Schulenburg, K., et al. (2015) Functional characterization and substrate promiscuity of UGT71 glycosyltransferases from strawberry (*Fragaria*×*ananassa*). *Plant Cell Physiol.* 56: 2478–2493.

Tikunov, Y.M., Molthoff, J., de Vos, R.C., Beekwilder, J., van Houwelingen, A., van der Hoofte, J.J., et al. (2013) Non-smoky glycosyltransferase1 prevents the release of smoky aroma from tomato fruit. *Plant Cell* 25: 3067–3078.

Trapero, A., Ahrazem, O., Rubio-Moraga, A., Jimeno, M.L., Gomez, M.D. and Gomez-Gomez, L. (2012) Characterization of a glucosyltransferase enzyme involved in the formation of kaempferol and quercetin sophorosides in *Crocus sativus*. *Plant Physiol.* 159: 1335–1354.

Wang, P., Wei, Y., Fan, Y., Liu, Q., Wei, W., Yang, C., et al. (2015) Production of bioactive ginsenosides Rh2 and Rg3 by metabolically engineered yeasts. *Metab. Eng.* 29: 97–105.

Williams, P.J., Strauss, C.R., Wilson, B. and Massy-Westropp, R.A. (1982) Novel monoterpene disaccharide glycosides of *Vitis vinifera* grapes and wines. *Phytochemistry* 21: 2013–2020.

Winter, M. and Herrmann, K. (1984) Analysis of hydroxycinnamic acid esters and their glucosides by reversed-phase high-performance liquid chromatography after polyamide separation. *J. Chromatogr.* 315: 243–251.

Yonekura-Sakakibara, K., Fukushima, A., Nakabayashi, R., Hanada, K., Matsuda, F., Sugawara, S., et al. (2012) Two glycosyltransferases involved in anthocyanin modification delineated by transcriptome independent component analysis in *Arabidopsis thaliana*. *Plant J.* 69: 154–167.

Yonekura-Sakakibara, K. and Hanada, K. (2011) An evolutionary view of functional diversity in family 1 glycosyltransferases. *Plant J.* 66: 182–193.

Yonekura-Sakakibara, K., Nakabayashi, R., Sugawara, S., Tohge, T., Ito, T., Koyanagi, M., et al. (2014) A flavonoid 3-O-glucoside: 2″-O-glucosyltransferase responsible for terminal modification of pollen-specific flavonols in *Arabidopsis thaliana*. *Plant J.* 79: 769–782.

Yu, B.W., Luo, J.G., Wang, J.S., Zhang, D.M., Yu, S.S. and Kong, L.Y. (2011) Pentasaccharide resin glycosides from *Ipomoea pes-caprae*. *J. Nat. Prod.* 74: 620–628.

Zeng, Y. and Yang, T. (2002) RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. *Plant Mol. Biol. Rep.* 20: 417a–417e.