PCP
PLANT & CELL PHYSIOLOGY

Two Ginseng UDP-Glycosyltransferases Synthesize Ginsenoside Rg_3 and Rd

Suk-Chae Jung[1], Woohyun Kim[1], Sung Chul Park[1], Jinkil Jeong[1], Myung Keun Park[1], Soohwan Lim[1], Yeon Lee[1], Wan-Taek Im[2], Jun Hyoung Lee[2], Giltsu Choi[1,*] and Sun Chang Kim[1,*]

[1]Department of Biological Sciences, KAIST, Daejeon 305-701, Republic of Korea
[2]KAIST Institute for Biocentury, KAIST, Daejeon, 305-701, Republic of Korea
*Corresponding authors: Giltsu Choi, E-mail, gchoi@kaist.ac.kr; Fax, +82-42-350-2610; Sun Chang Kim, E-mail, sunkim@kaist.ac.kr; Fax, +82-42-350-2610.
(Received July 7, 2014; Accepted October 9, 2014)

Ginseng is a medicinal herb that requires cultivation under shade conditions, typically for 4–6 years, before harvesting. The principal components of ginseng are ginsenosides, glycosylated tetracyclic terpenes. Dammarene-type ginsenosides are classified into two groups, protopanaxadiol (PPD) and protopanaxatriol (PPT), based on their hydroxylation patterns, and further diverge to diverse ginsenosides through differential glycosylation. Three early enzymes, dammarenediol-II synthase (DS) and two P450 enzymes, protopanaxadiol synthase (PPDS) and protopanaxatriol synthase (PPTS), have been reported, but glycosyltransferases that are necessary to synthesize specific ginsenosides have not yet been fully identified. To discover glycosyltransferases responsible for ginsenoside biosynthesis, we sequenced and assembled the ginseng transcriptome de novo and characterized two UDP-glycosyltransferases (PgUGTs): PgUGT74AE2 and PgUGT94Q2. PgUGT74AE2 transfers a glucose moiety from UDP-glucose (UDP-Glc) to the C3 hydroxyl groups of PPD and compound K to form Rh_2 and F2, respectively, whereas PgUGT94Q2 transfers a glucose moiety from UDP-Glc to Rh_2 and F2 to form Rg_3 and Rd, respectively. Introduction of the two *UGT* genes into yeast together with *PgDS* and *PgPPDS* resulted in the de novo production of Rg_3. Our results indicate that these two UGTs are key enzymes for the synthesis of ginsenosides and provide a method for producing specific ginsenosides through yeast fermentation.

Keywords: Ginsenoside biosynthesis • *Panax ginseng* • UDP-glycosyltransferase • UGT74 • UGT94 • Yeast.

Abbreviations: *ATR*, Arabidopsis cytochrome P450 reductase gene; CK, compound K; DEC, differentially expressed contig; DS, dammarenediol-II synthase; *ERG7*, encoding lanosterol synthase; FKPM, fragments per kilobase of exon per million fragments mapped; GO, gene ontology; *HMGR*, 3-hydroxy-3-methylglutaryl coenzyme A reductase; LC-MS/MS, liquid chromatography–tandem mass spectrometry; MeJA, methyl jasmonate; MRM, multiple reaction monitoring; PPD, protopanaxadiol; PPDS, protopanaxadiol synthase; PPT, protopanaxatriol; PPTS, protopanaxatriol synthase; TLC, thin-layer chromatography; *tHMGR*, truncated HMG-CoA reductase; UDP-Glc, UDP-glucose; UGT, UDP-glycosyltransferase.

Introduction

Ginseng (*Panax ginseng* C.A Meyer) is a popular medicinal herb widely used for its health benefits. The beneficial effects of ginseng are attributed to ginsenosides—glycosylated dammarene-type tetracyclic triterpene compounds (Choi et al. 2013). More than 40 different ginsenosides producing various pharmacological effects have been identified in ginseng roots (Fuzzati 2004). They are classified into two groups: the protopanaxadiol (PPD) group and the protopanaxatriol (PPT) group. PPD and PPT have different hydroxylation patterns (**Supplementary Table S1**); PPD has hydroxyl groups at C3, C12 and C20, whereas PPT has hydroxyl groups at C3, C6, C12 and C20. PPD and PPT are further glycosylated into various ginsenosides by glucose and other sugars. Representative PPD group ginsenosides include Rh_2, Rg_3, compound K (CK), F2, Rd and Rb_1, whereas PPT group ginsenosides include F1, Rg_1, Re, Rh_1 and Rg_2. Different ginsenosides are reported to have different pharmacological effects (Choi et al. 2013).

The ginsenoside biosynthetic pathway has only been partially elucidated. The pathway is shared with other triterpene pathways up to the synthesis of 2,3-oxidosqualene, which is produced by a series of condensation reactions of isopentenyl diphosphate and dimethylallyl diphosphate through the actions of isopentenyl diphosphate isomerase, geranyl diphosphate synthase, farnesyl diphosphate synthase, squalene synthase and squalene epoxidase (Vranova et al. 2013, Thimmappa et al. 2014). 2,3-Oxidosqualene is subsequently cyclized by dammarenediol-II synthase (DS), a triterpene cyclase, to form dammarenediol-II (Tansakul et al. 2006). Dammarenediol-II, which has hydroxyl groups at C3 and C20, is hydroxylated by protopanaxadiol synthase (PgPPDS, CYP716A47), a P450 enzyme, at its C12 position to form PPD (Han et al. 2011). PPD is further hydroxylated by protopanaxatriol synthase (PgPPTS, CYP716A53v2), another P450 enzyme, at its C6 position to form PPT (Han et al. 2012). PPD is glycosylated at its C3 and/or C20 position, yielding PPD-type ginsenosides, whereas PPT is glycosylated at its C6 and/or C20

Regular Paper

Plant Cell Physiol. 55(12): 2177–2188 (2014) doi:10.1093/pcp/pcu147, Advance Access publication on 14 October 2014, available online at www.pcp.oxfordjournals.org

position, yielding PPT-type ginsenosides (**Supplementary Table S1**). UDP-glycosyltransferases (UGTs) are predicted to be responsible for the synthesis of various ginsenosides by forming *O*-β1,2- or *O*-β1,6-glycosidic linkages. Among ginsenoside biosynthetic enzymes, DS, PPDS and PPTS have been reported and their abilities to synthesize PPD or PPT have been demonstrated in yeast (Dai et al. 2013, Dai et al. 2014). However, no UGTs, except one that converts PPD to CK, have been identified for ginsenoside biosynthesis (Yan et al. 2014).

UGTs are enzymes that transfer a sugar from UDP-sugar to various metabolites, including hormones and secondary metabolites (Hou et al. 2004, Bowles et al. 2006). UGTs typically act at the final steps of biosynthetic pathways to increase solubility, stability, storability, bioactivity or bioavailability of metabolites. Each plant genome contains hundreds of different *UGT* genes, a reflection of the enormous diversity of plant metabolites. For example, the model plant *Arabidopsis thaliana*, though not famous for its diverse secondary metabolites, nevertheless contains 107 UGTs, which are classified into 14 groups (Group A to Group N) based on their amino acid sequences (Li et al. 2001, Ross et al. 2001). Different UGTs display substrate specificity for both sugar donors and acceptors. For example, UGT78D2 transfers a glucose molecule from UDP-glucose (UDP-Glc) to flavonols (kaempferol, quercetin) and anthocyanidin (cyanidin) at their C3 positions to produce flavonol 3-*O*-glucosides and cyanidin 3-*O*-glucoside, respectively (Lee et al. 2005, Tohge et al. 2005). It has been shown that such glycosylation is necessary for the stability and storability of UGTs in vivo. In contrast, UGT89C1 transfers rhamnose from UDP-rhamnose to flavonol 3-*O*-glucosides at their C7 positions to produce flavonol 3-*O*-glucoside-7-*O*-rhamnoside (Yonekura-Sakakibara et al. 2007). UGT89C1 uses neither UDP-Glc nor anthocyanidin 3-*O*-glucoside as a substrate, demonstrating different specificities for both UDP-sugar and acceptor compared with UGT78D2. UGTs that have shown to glycosylate saponins also display sugar and donor specificities. For example, UGT71G1 transfers glucose from UDP-Glc to medicagenic acid at the C28 position (Achnine et al. 2005); UGT73P2 transfers galactose from UDP-galactose to soyasapogenol B monoglucurionide at the C3 position to produce soyasaponin III; UGT91H4 transfers rhamnose from UDP-rhamnose to soyasaponin III at the C3 position of soyasaponin III to produce soyasaponin I (Shibuya et al. 2010); UGT73C11 transfers glucose to oleanolic acid and hederagenin at C3 positions to produce 3-*O*-β-D-Glc oleanolic aicd and 3-*O*-β-D-Glc hederagenin, respectively (Augustin et al. 2012); SvUGT74M1 transfers glucose to gypsogenic acid and quillaic acid at the C28 position. Given the substrate specificities of different UGTs, it is imperative that specific ginseng UGTs be identified to clarify and engineer the ginsenoside biosynthetic pathway for use in producing specific ginsenosides in other organisms.

In the present study, we isolated and characterized two ginseng UGTs that participate in ginsenoside biosynthesis. One of these, termed PgUGT74AE2, converts PPD to Rh_2 and CK to F2 by forming a 3-*O*-glycosidic linkage at the C3 position. The other, termed PgUGT94Q2, converts Rh_2 to Rg_3 and F2 to Rd by forming a β1,2-glycosidic linkage. When the two *UGT* genes were introduced into yeast together with DS and PPDS, the yeast were able to produce Rg_3. These two UGTs represent specific ginsenoside biosynthetic UGTs that can be used to produce specific ginsenoside in heterologous hosts.

Results

Transcriptomic analysis identifies MeJA-inducible genes in *P. ginseng*

Since MeJA (methyl jasmonate) is known to increase ginsenoside biosynthesis (Kim et al. 2004, Kim et al. 2007), we assumed that unidentified ginsenoside biosynthetic genes would be up-regulated by MeJA. We performed RNA-seq analysis of ginseng roots and leaves, with and without MeJA treatment, and assembled gene contigs de novo. A total of 254,858 contigs spanning >202 million base pairs were obtained with an N50 value of 1,325 bp (**Supplementary Table S2**).

To determine which genes were up-regulated by MeJA, we first defined differentially expressed contigs (DECs) in both leaves and roots using a criterion of a >2-fold change in expression with MeJA treatment. Contigs that were expressed at very low levels [<2 fragments per kilobase of exon per million fragments mapped (FKPM)] were excluded and the remaining 36,860 contigs were used for the DEC analysis. This analysis identified 7,366 leaf and 5,993 root DECs (**Fig. 1A**). A comparison between leaves and roots revealed 2,567 shared DECs, which is highly significant ($P = e^{-1084}$, hypergeometric test). A scatter plot analysis further showed that 2,093 of the 2,567 shared DECs (82%) were either up-regulated in both leaves and roots or down-regulated in both leaves and roots by MeJA (**Fig. 1B**), indicating that MeJA regulates the expression of both overlapping and distinct contigs in leaves and roots. To examine which genes are differentially expressed in ginseng in response to MeJA, we mapped shared DECs to Arabidopsis genes by BlastX (e-value $\geq 10^{-5}$). If more than one Arabidopsis gene had an e-value $> 10^{-5}$, we mapped it to the gene with the highest e-value. The 2,567 shared DECs were mapped to 1,191 Arabidopsis genes (**Supplementary Table S3**), of which 449 genes corresponded to MeJA-inducible DECs and 747 genes corresponded to MeJA-repressible DECs (**Fig. 1C**). A gene ontology (GO) analysis indicated that MeJA-inducible genes were enriched for secondary metabolic processes, whereas MeJA-repressible genes were enriched for photosynthesis (**Fig. 1D**). The enrichment of these GO terms is consistent with the biological effects of MeJA in other plants (De Geyter et al. 2012, Lenka et al. 2012). Taken together, our results indicate that MeJA induces sets of genes in the ginseng plant whose products are involved in secondary metabolism.

MeJA induced the expression of isoprenoid pathway genes in ginseng. In this pathway, seven early genes, including *HMGR* (encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase), are responsible for the synthesis of isopentenyl pyrophosphate and dimethylallyl pyrophosphate from acetyl-CoA (**Supplementary Fig. S1**). Among these genes, all except that encoding isopentenyl diphosphate isomerase were induced at least 2-fold by MeJA in leaves, roots or both. *HMGR*, encoding

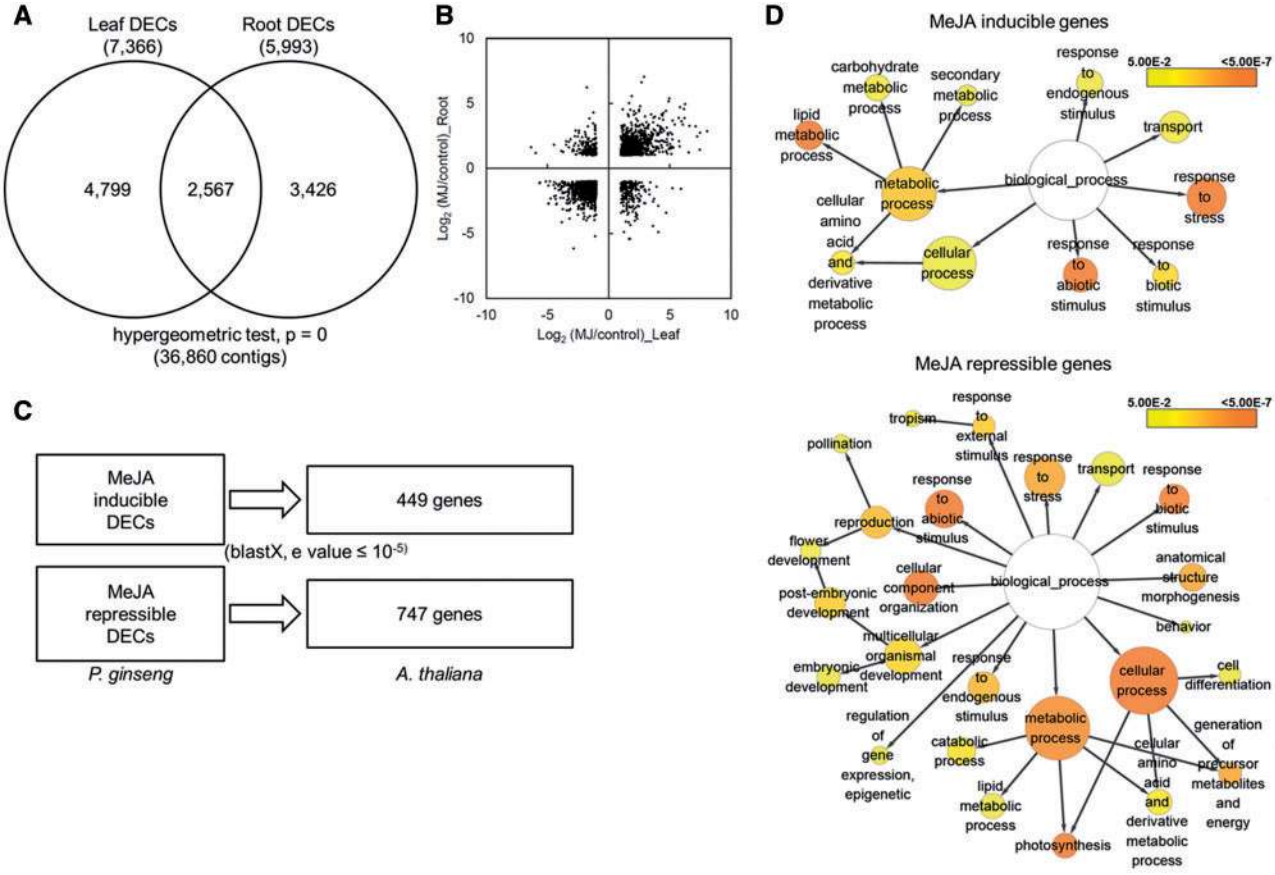

Fig. 1 Transcriptomic analysis of MeJA-inducible genes in *P. ginseng*. (A) Venn diagram showing overlapping DECs in ginseng leaves and roots. Contigs with FKPM values <2 were not used for the DEC analysis. Numbers indicate contig numbers in each section. (B) Scatter plot analysis showing that the shared DECs are regulated in the same direction by MeJA in the leaf and root. (C) Mapping of ginseng DECs to Arabidopsis genes by BlastX. (D) BiNGO analysis of mapped genes showing enrichment for metabolic process terms, including secondary metabolic processes, among MeJA-inducible genes. Upper panel: MeJA-inducible genes. Lower panel: MeJA-repressible genes. The color bar at the bottom indicates *P*-values.

an enzyme that catalyzes the rate-limiting step in this pathway, was induced 178-fold by MeJA in leaves and 6-fold in roots. Isopentenyl diphosphate and dimethylallyl pyrophosphate are subsequently condensed and converted to 2,3-oxidosqualene by the action of farnesyl diphosphate synthase, squalene synthase and squalene epoxidase. Similar to early genes, all three late genes were also induced >2-fold by MeJA in both leaves and roots. The produced 2,3-oxidosqualene is cyclized by PgDS and further hydroxylated by PgPPDS to produce protopanaxadiol, a precursor of PPD-type ginsenosides. Among these genes, *PgDS*, encoding an enzyme catalyzing the first committed step of the ginsenoside biosynthesis, was induced 13-fold in leaves and 8-fold in roots, whereas *PgPPDS* was induced 7-fold in leaves and 5-fold in roots. Overall, our transcriptome data indicate that MeJA induces the expression of isoprenoid pathway genes and further suggest that MeJA probably activates the expression of other unidentified ginsenoside biosynthetic genes in ginseng.

Two ginseng UGTs convert PPD to Rg_3 and CK to Rd

Our ginseng transcriptome included 19 UGT contig groups that are induced by MeJA >2-fold both in leaf and in root. Among them, we were able to clone and purify six full-length UGTs (see Accession/AGI numbers in the Materials and Methods). Three UGTs showed biochemical activities toward ginsenosides. One of them was identical to the reported UGTPg1 (PgUGT71A27) (Yan et al. 2014); thus we focused on two other UGTs. A sequence analysis indicated that the two UGTs were similar to UGTs from other plants including those from Arabidopsis. One UGT clustered together with UGT74 family members from various plants including Arabidopsis. We named this *P. ginseng* (Pg) glycosyltransferase, PgUGT74AE2 (**Fig. 2**) after consulting with the UGT Nomenclature Committee (Mackenzie et al. 1997, Mackenzie et al. 2005). Among UGT74 family members, UGT74F1 has been shown to glucosylate salicylate by forming a 2-*O*-glycosidic linkage, whereas UGT74F1 and UGT74F2 glucosylate both salicylate and anthranilate by forming an ester linkage (Quiel and Bender 2003). SvUGT74M1, another UGT74 family member from *Saponaria vacaria*, has been shown to glucosylate gypsogenic acid by forming an ester linkage at their C28 position (Meesapyodsuk et al. 2007). The other UGT was clustered together with UGT94 family members, including BpUGT94B1, from *Bellis perennis* (Bp), and SiUGT94D1, from *Sesamum indicum* (Si). We named this

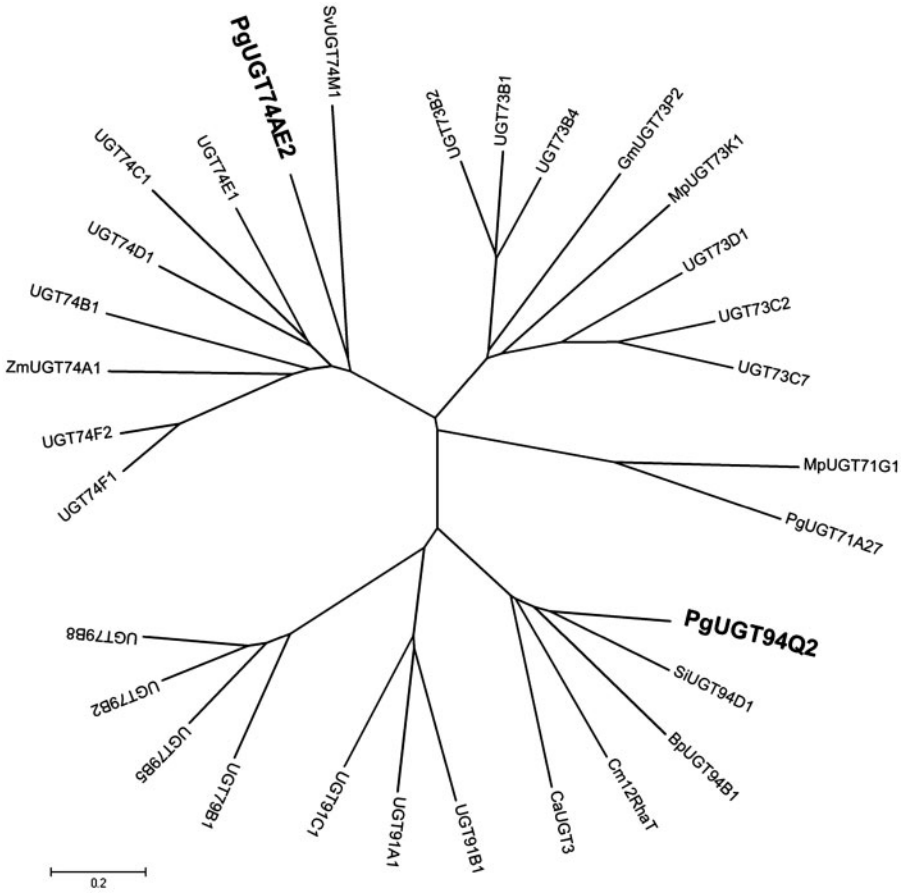

Fig. 2 The two ginseng UGTs are clustered with UGT74 and UGT94 family members. PgUGT74AE2 and PgUGT94Q2 are clustered with UGT74 and UGT94, respectively. MEGA5 software was used to align amino acid sequences of UGTs and draw the Neighbor–Joining tree (Tamura et al. 2011). Accession numbers or AGI codes are indicated in the Materials and Methods.

glycosyltransferase PgUGT94Q2 after consulting with the UGT Nomenclature Committee (Mackenzie et al. 1997, Mackenzie et al. 2005). It has been shown that BpUGT94B1 adds a glucuronosyl moiety to 3-*O*-glucosides of cyanidin by forming a β1,2 linkage (Sawada et al. 2005), whereas SiUGT94D1 adds a glucose to 2-*O*-glucoside of sesaminol by forming a β1,6 linkage (Noguchi et al. 2008). These observations suggest that PgUGT74AE2 glycosylates ginsenosides by forming either *O*-glycosidic or glycosyl ester linkages, whereas glycosylation by PgUGT94Q2 adds the second sugar molecule to previously glycosylated ginsenosides by forming either a β1,2 linkage or a 1,6 linkage. Among ginsenosides, PPD is glucosylated through *O*-linkage at its C3 and/or C20 positions to form ginsenosides, including Rh_2 and CK, whereas PPT is glucosylated through *O*-linkage at its C6 and/or C20 positions to form ginsenosides, including Rh_1 and F1 (**Supplementary Table S1**). Many ginsenosides are formed by additional glycosylation through β1,2-glucosidic linkage (e.g. Rg_3, Rd and Rf) or through β1,6-glucosidic linkage (e.g. Rb_1).

We next used recombinant proteins to determine whether the newly identified PgUGT74AE2 and PgUGT94Q2 could glucosylate any ginsenosides. For these enzyme assays, recombinant PgUGTs were incubated with 10 different ginsenosides (PPD, CK, Rh_2, F2, Rg_3, Rd, Rg_2, Rh_1, F1 and PPT) in the presence of UDP-Glc. PgUGT74AE2 produced slowly migrating spots on thin-layer chromatography (TLC) when incubated with PPD and CK, but not with other ginsenosides (**Fig. 3A**); these newly appearing spots co-migrated with Rh_2 and F2, respectively. The reaction products were also analyzed by HPLC. Consistent with the TLC data, PgUGT74AE2 produced ginsenosides that co-migrated with Rh_2 and F2 when incubated with PPD and CK, respectively (**Fig. 3B**). To show further that PgUGT74AE2 produced Rh_2 and F2, we used liquid chromatography–tandem mass spectrometry (LC-MS/MS) with detection by multiple reaction monitoring (MRM). The observed retention time of the standard analytes were 20.12 min for PPD (transition at m/z 461.1→425.5), 12.32 min for Rh_2 (transition at m/z 645.3→23.2), 11.22 min for CK (transition at m/z 645.4→23.2) and 7.35 min for F2 (transition at m/z 807.5→627.5), respectively. The reaction products matched with Rh_2 and F2 (**Fig. 3C**), confirming that PgUGT74AE2 converted PPD to Rh_2 and CK to F2. K_m values for PPD and CK were 25 and 40 μM, respectively (**Table 1**). These results indicate that PgUGT74AE2 glucosylates the C3 position of both PPD and CK, but not C20 or C12 positions. PgUGT74AE2 also did not glucosylate hydroxyl residues at C3, C6, C12 or C20 positions in any PPT-type ginsenoside, displaying both strong acceptor specificity and regioselectivity.

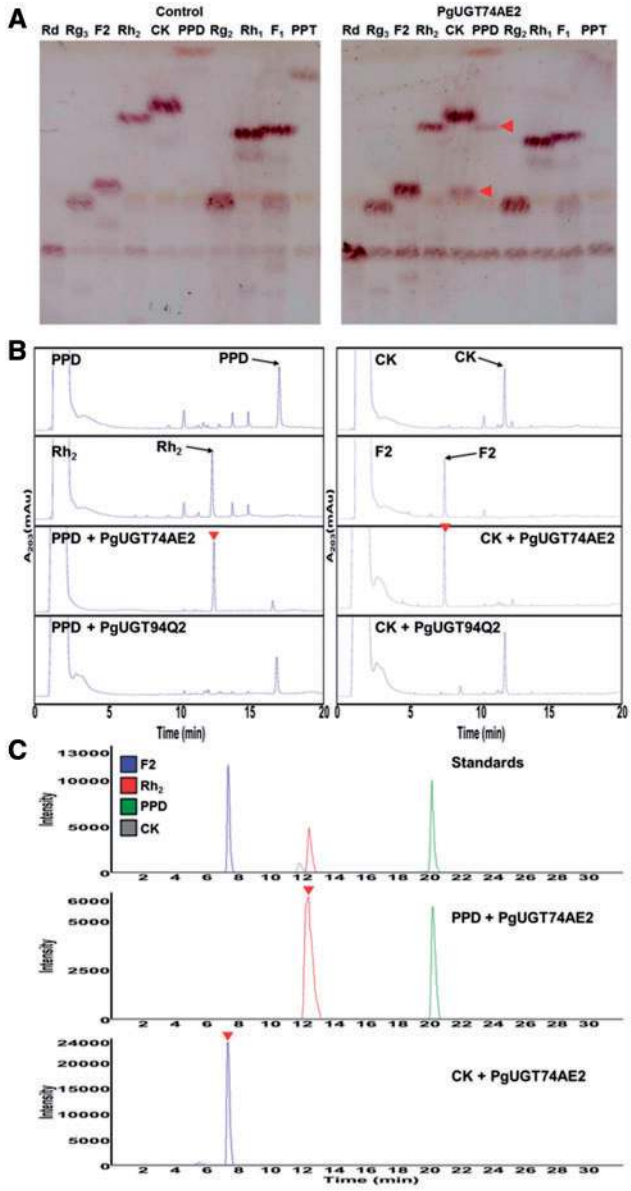

Fig. 3 PgUGT74AE2 converts PPD to Rh_2 and CK to F2. (A) A TLC analysis of 10 different ginsenosides incubated in the presence of UDP-Glc with (right panel) or without (left panel) PgUGT74AE2. PgUGT74AE2 glucosylated only PPD and CK to Rh_2 and F2, respectively (indicated by red triangles). (B) HPLC analysis of PgUGT74AE2 reaction products with PPD (left panel) and CK (right panel). Newly produced Rh_2 and F2 are indicated by inverted red triangles. (C) LC-MS/MS analysis of PgUGT74AE2 reaction products with PPD (middle panel) and CK (bottom panel). The top panel shows retention times of F2, Rh_2, PPD and CK. Newly produced Rh_2 and F2 are indicated by inverted red triangles.

PgUGT94Q2 generated slowly migrating spots on TLC when incubated with Rh_2 and F2, but not with other ginsenosides (**Fig. 4A**); these newly appearing spots co-migrated with Rg_3 and Rd, respectively. The reaction products were also analyzed by HPLC. Consistent with TLC data, PgUGT94Q2 produced ginsenosides that co-migrated with Rg_3 when incubated with Rh_2 and with Rd when incubated with F2 (**Fig. 4B**). The observed retention times of products were 8.03 min (transition at m/z 807.4→365.3) when incubated with Rh_2 and 6.13 min (transition at m/z 969.3→789.4) when incubated with F2 (**Fig. 4C**). These retention times matched with the observed retention times for Rg_3 (8.03 min, transition at m/z 807.4→365.3) and for Rd (6.13 min, transition at m/z 969.3→789.4), respectively, confirming that PgUGT94Q2 produced Rg_3 and Rd. K_m values for Rh_2 and F2 were 311 and 188 μM, respectively (**Table 1**). These results indicate that PgUGT94Q2 glucosylated the C3-*O*-glucoside of both Rh_2 and F2, but not the C20-*O*-glucoside of CK and F2, also displaying strong substrate specificity and regioselectivity. Taken together, these results indicate that PgUGT74AE2 is a glycosoyltransferase that converts PPD to Rh_2 and CK to F2, whereas PgUGT94Q2 is a glycosyltransferase that converts Rh_2 to Rg_3 and F2 to Rd.

The biochemical activities of the two PgUGTs suggested that these enzymes catalyze successive glucosylation reactions to convert PPD to Rg_3 or CK to Rd in ginseng. We tested whether we could recapitulate successive glucosylations of PPD to Rg_3 or CK to Rd in a single reaction. For this enzyme assay, both PgUGT74AE2 and PgUGT94Q2 were incubated with either PPD or CK in the presence of UDP-glc, and the products were analyzed by HPLC. PPD was successfully converted to Rg_3 in a single reaction tube, whereas CK was converted to Rd (**Fig. 5**). These results indicate that PgUGT74AE2 and PgUGT94Q2 could successively glucosylate PPD and CK to produce Rg_3 and Rd, respectively.

PgUGT74AE2 and *PgUGT94Q2* mRNAs are expressed at higher levels in roots than in leaves

Two *PgUGT* genes were highly induced by MeJA both in both roots and leaves (**Fig. 6A**). Since ginseng root has been traditionally used for medicinal purposes, we further examined whether the two *PgUGT* genes were expressed mainly in the root. The organ-specific expression patterns of the two *PgUGT* genes were determined together with those of other known ginsenoside biosynthetic genes in our transcriptome data. Many isoprenoid biosynthetic genes were expressed in both roots and leaves (**Fig. 6B**), but relative expression levels were different between the two organs (**Fig. 6C**). Among seven early

Table 1 K_m values of PgUGT74AE2 for PPD and CK, and PgUGT94Q2 for Rh_2 and F2

Substrate	PgUGT74AE2		PgUGT94Q2	
	PPD	CK	Rh_2	F2
K_m (μM)	25 ± 4.3	40 ± 2.1	311 ± 14	188 ± 42
k_{cat} (s^{-1})	$5.56 \times 10^{-5} \pm 2.09 \times 10^{-5}$	$13.47 \times 10^{-5} \pm 1.67 \times 10^{-5}$	8.6 ± 0.19	0.41 ± 0.03
k_{cat}/K_m ($M^{-1}s^{-1}$)	2.24 ± 0.41	3.37 ± 0.8	27,683.3 ± 1,404.77	2,183.06 ± 714.29

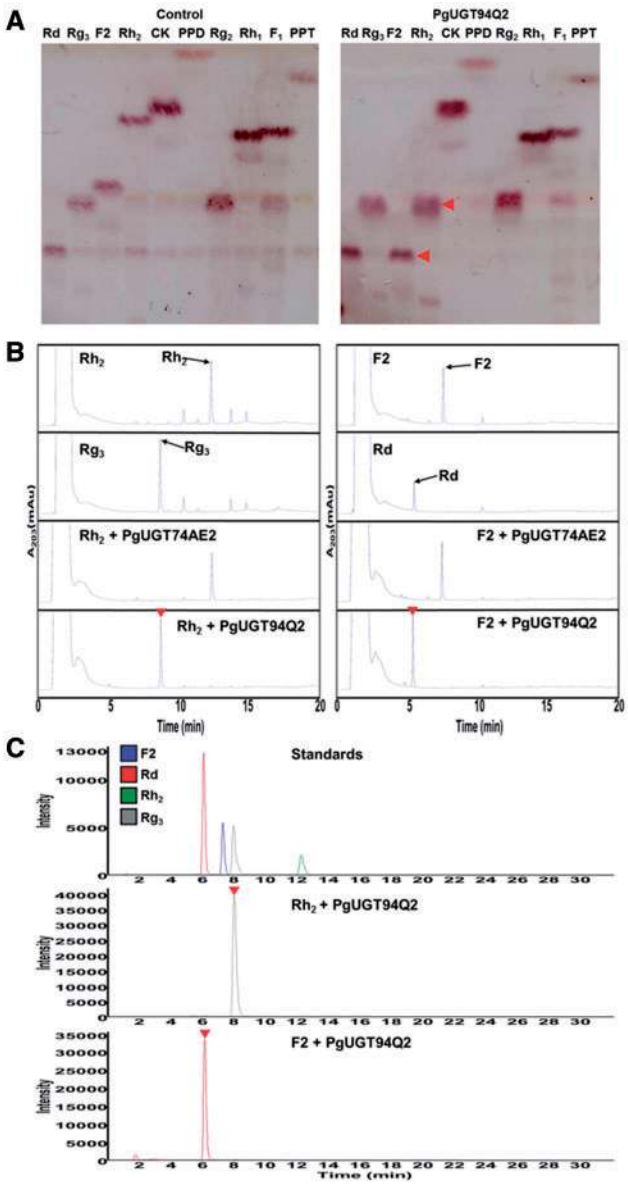

Fig. 4 PgUGT94Q2 converts Rh_2 to Rg_3 and F2 to Rd. (A) TLC analysis of 10 different ginsenosides incubated in the presence of UDP-Glc with (right panel) or without (left panel) PgUGT94Q2. PgUGT94Q2 glucosylated only Rh_2 and F2 to Rg_3 and Rd, respectively (indicated by red triangles). (B) HPLC analysis of PgUGT94Q2 reaction products with Rh_2 (left panel) and F2 (right panel). Newly produced Rg_3 and Rd are indicated by red triangles. (C) LC-MS/MS analysis of PgUGT94Q2 reaction products with Rh_2 (middle panel) and F2 (bottom panel). The top panel shows retention times of Rd, F2, Rg_3 and Rh_2. Newly produced Rg_3 and Rd are indicated by inverted red triangles.

isoprenoid pathway genes that lead to the synthesis of isopentenyl diphosphate and dimethylallyl pyrophosphate from acetyl-CoA, five, including the rate-limiting *HMGR*, were expressed at higher levels in the root than in the leaf, suggesting that acetyl-CoA is more actively converted to isoprene units in the root. Isopentenyl diphosphate and dimethylallyl pyrophosphate are subsequently converted to 2,3-oxidosqualene by the

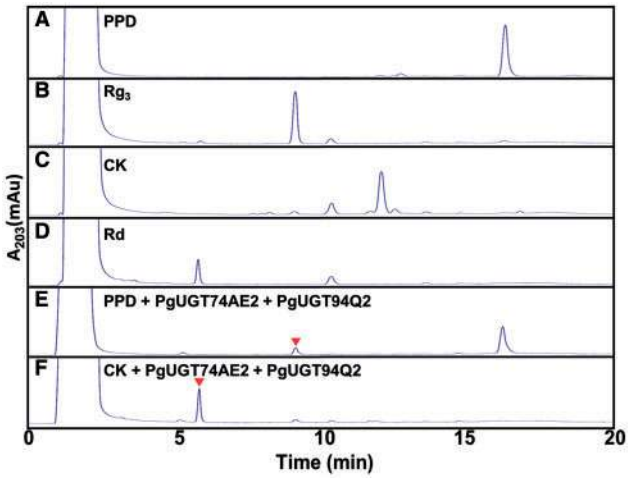

Fig. 5 PgUGT74AE2 and PgUGT94Q2 directly convert PPD to Rg_3 and CK to Rd. (A, B, C and D) PPD, Rg_3, CK and Rd standard, respectively. PPD (E) or CK (F) was incubated with both PgUGT74AE2 and PgUGT94Q2 in the presence of UDP-Glc. Reaction mixtures were analyzed by HPLC; newly produced Rg_3 and Rd are indicated by inverted triangles with names in parentheses.

action of three gene products (farnesyl diphosphate synthase, squalene synthase and squalene epoxidase). Among these, the farnesyl diphosphate synthase gene was expressed similarly in the root and the leaf, whereas squalene synthase and squalene epoxidase genes were expressed at lower levels in the root. 2,3-Oxidosqualene is further cyclized by 2,3-oxidosqualene cyclases to various triterpenoids, such as lanosterol and β-amyrin. Among 2,3-oxidosqualene cyclases, PgDS cyclizes 2,3-oxidosqualene to dammarenediol-II, which in turn is hydroxylated by PgPPDS to form PPD and subsequently hydroxylated by PgPPTS to form PPT. The transcriptome data indicated that *PgDS* was expressed similarly in the root and the leaf, whereas *PgPPDS* was more highly expressed in the root. Both *PgUGT74AE2* and *PgUGT94Q2* were also much more highly expressed in roots than in leaves (>4-fold). Taken together, these gene expression patterns suggest that ginsenosides are synthesized in both leaves and roots, but are glucosylated more actively in roots.

PgUGT74AE2 and PgUGT94Q2 enable the de novo synthesis of ginsenosides in yeast

Panax ginseng is a slow-growing, obligate shade plant that is typically cultivated in the shade for 4–6 years before harvesting. This limitation of ginseng cultivation adds substantially to the cost to mass-produce the plant. Thus, we investigated whether it was possible to produce ginsenosides de novo in a more amenable heterologous system using PgUGT74AE2 and PgUGT94Q2. We chose yeast (*Saccharomyces cerevisiae* CEN.PK 2-1D) as a heterologous production system for its long history of fermentation for human consumption.

To produce ginsenosides in yeast, we first replaced the promoter of the yeast lanosterol synthase gene (*ERG7*) with a methionine-repressible *MET3* promoter (**Supplementary Fig. S2A**). Lanosterol synthase converts the majority of

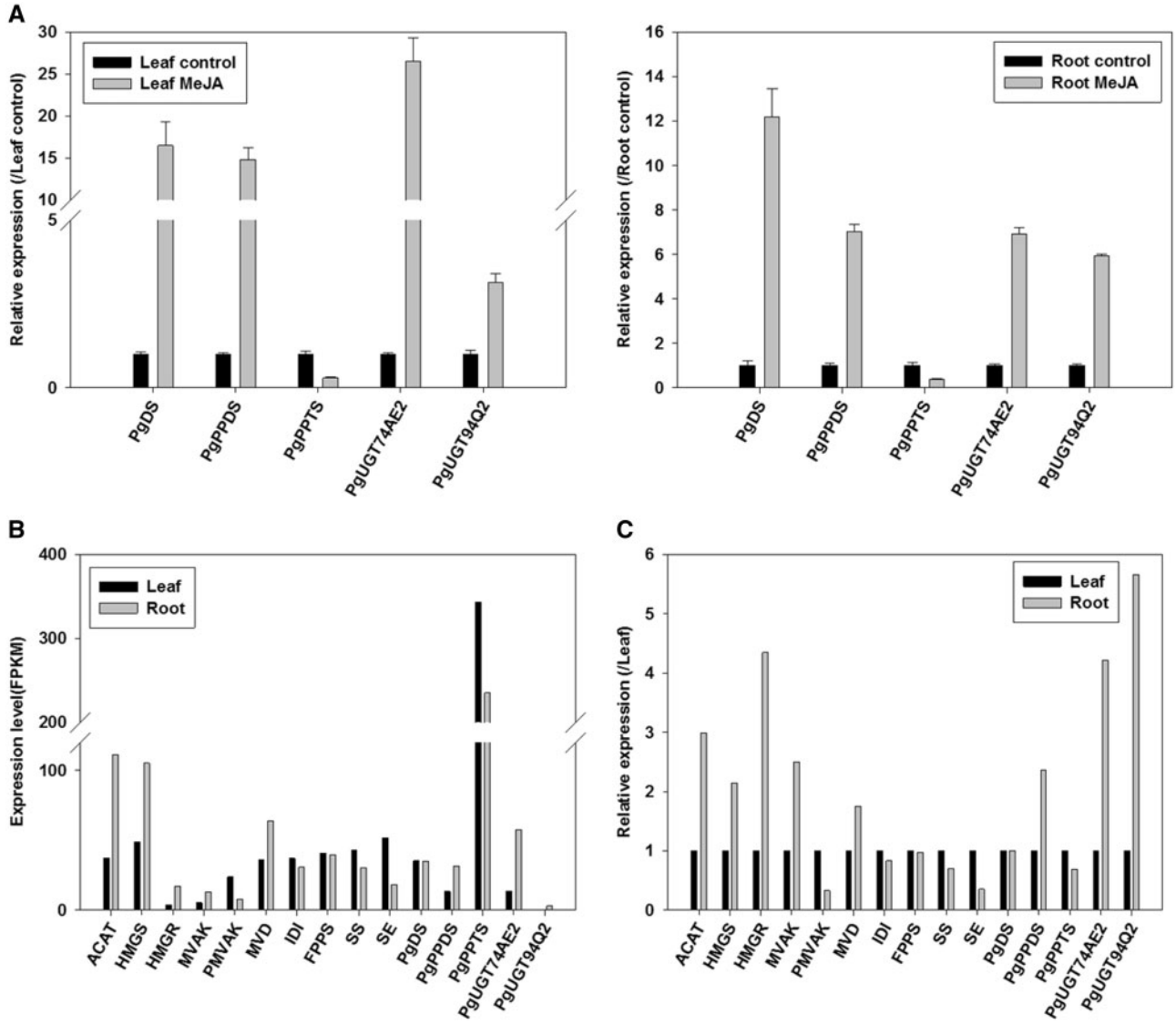

Fig. 6 *PgUGT74AE2* and *PgUGT94Q2* are expressed at higher levels in roots than in leaves. (A) Induction of the two *UGT* genes (*PgUGT74AE2* and *PgUGT94Q2*) by MeJA. Expression levels of ginsenoside biosynthetic genes were determined by quantitative real-time PCR using ginseng leaf and root samples treated with MeJA. Expression levels were normalized to tubulin mRNA levels; the relative expression levels of MeJA-treated samples are indicated (SD, $n = 3$). (B) Expression levels of *PgUGT74AE2*, *PgUGT94Q2* and other isoprenoid pathway genes. The expression levels were extracted from transcriptome data as FKPM values. (C) Levels of *PgUGT74AE2*, *PgUGT94Q2* and other isoprenoid pathway genes, expressed relative to levels in leaves.

2,3-oxidosqualene to lanosterol, a precursor of ergosterol in yeast. Thus, this maneuver allows expression of the endogenous lanosterol synthase gene to be repressed by adding methionine, which in turn results in the accumulation of 2,3-oxidosqualene for use by other pathways. We confirmed that addition of methionine to the engineered yeast line did indeed induce the accumulation of 2,3-oxidosqualene (**Supplementary Fig. S2B**). Next, we introduced expression vectors for *PgDS*, *PgPPDS* and an Arabidopsis Cyt P450 reductase gene (*ATR2*, AT4G30210). PgDS cyclizes 2,3-oxidosqualene to form dammarenediol-II and PgPPDS hydroxylates dammarenediol-II to form PPD. *ATR2* was included because plant P450 reductase is more efficient than endogenous yeast P450 reductase in reducing plant P450 (Ro et al. 2002), in this case PgPPDS, which is a prerequisite for P450 enzymatic activity. Finally, we introduced *PgUGT74AE2* and *PgUGT94Q2* together.

To determine if ginsenosides could be produced de novo in yeast, we grew each yeast line using batch-fed fermentation for 10 d (OD_{600} ~80) in the presence of methionine. Yeast cells were harvested by centrifugation, and ginsenosides were analyzed by HPLC and LC-MS/MS. Yeast harboring both *PgUGT74AE2* and *PgUGT94Q2* produced a new peak at 9.161 min which is identical to the retention time of Rg_3 in HPLC (**Fig. 7A**). The observed retention time of products in LC-MS/MS detected by MRM was 7.98 min with the transition at 807.4→365.3, which is identical to Rg_3 (**Fig. 7B**). We also analyzed the product with negative ion electrospray mass spectrometry (ES-MS) at 390°C. The products produced ions at *m/z*

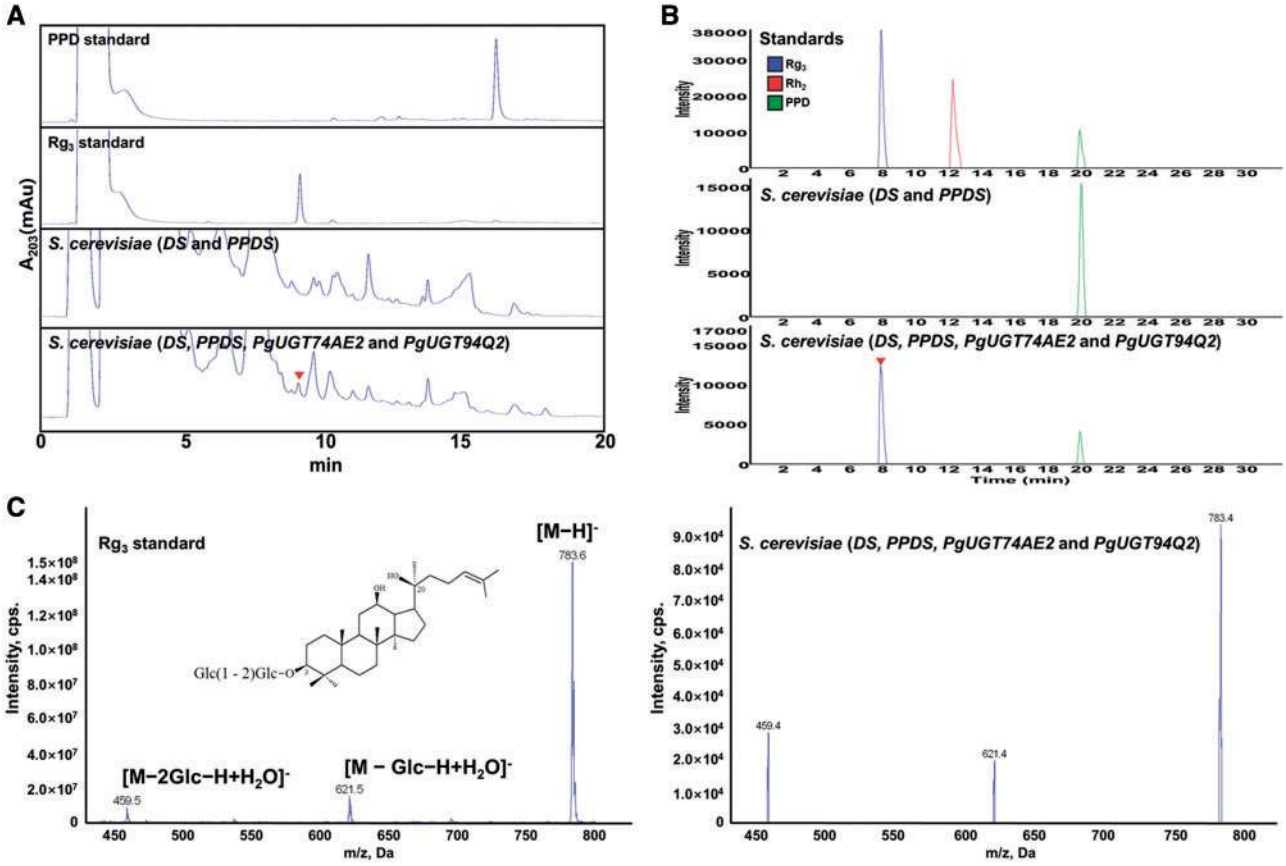

Fig. 7 Yeast harboring *PgUGT74AE2* and *PgUGT94Q2* synthesize Rg_3 de novo. (A) HPLC analysis of yeast harboring *PgUGT74AE2* and *PgUGT94Q2*. In addition to the two *UGT* genes, yeast also contains methionine-repressible *ERG7*, N-terminally truncated HMGR (*tHMGR*), *PgDS*, *PgPPDS* and *ATR2*. The Rg_3 produced is indicated by an inverted red triangle. (B) LC-MS/MS analysis of yeast-produced Rg_3. The upper panel indicates retention times of control compounds, and the lower panel indicates Rg_3 and PPD produced in yeast. The produced Rg_3 is indicated by an inverted red triangle. (C) Negative ion electrospray mass spectrometry (ES-MS) analysis of yeast-produced Rg_3. The left panel shows ES-MS data for standard Rg_3 and the right panel shows those for yeast-produced Rg_3.

621 and 459 that matched with ions produced by Rg_3 in ES-MS (**Fig. 7C**), confirming the production of Rg_3 in yeast. Taken together, our results indicate that the two *UGT* genes identified here enable the de novo production of ginsenoside through yeast fermentation (1.3 mg l^{-1}).

Discussion

UGTs are expected to play central roles in producing various ginsenosides since most of them are primarily derived through differential glycosylation of two dammarene-type tetracyclic terpenes, PPD and PPT. However, with the exception of one UGT that converts PPD to CK, no other UGTs involved in the pathway have been cloned. In this report, we cloned two ginseng UGTs—PgUGT74AE2 and PgUGT94Q2—that catalyze specific glucosylations of PPD-type ginsenosides. PgUGT74AE2 glucosylated the C3-OH position of either PPD or CK to produce Rh_2 and F2, respectively, by forming an *O*-glycosidic bond, whereas PgUGT94Q2 glucosylated 3-*O*-glucosides of the PPD-type ginsenosides, Rh_2 and F2, to produce Rg_3 and Rd, respectively, by forming a β1,2-glycosidic bond (**Fig. 8**). In the presence of the two UGTs, PPD and CK were directly converted to Rg_3 and Rd, respectively, in vitro through successive glucosylations of PPD and CK (**Fig. 5**). Consistent with this, introduction of the two *UGT* genes into yeast together with *PgDS*, *PgPPDS* and *ATR2* enabled the de novo production of Rg_3 (**Fig. 7**).

Similar to other UGTs, both PgUGT74AE2 and PgUGT94Q2 displayed strong acceptor specificity and regioselectivity. PgUGT74AE2 glucosylated only PPD and CK out of 10 different ginsenosides, six of which are PPD-type and four of which are PPT-type. Among them, two PPD-type ginsenosides (PPD and CK) and four PPT-type ginsenosides (PPT, Rg_2, Rh_1 and F1) have a hydroxyl group at the C3 position, yet the C3-OH positions of only PPD-type ginsenosides served as acceptors for PgUGT74AE2. The only difference between PPD and PPT is the presence of a hydroxyl group at the C6 position, further indicating that PgUGT74AE2 can distinguish acceptors with C6-H and C6-OH. Similarly, PgUGT94Q2 glucosylated only Rh_2 and F2 out of 10 ginsenosides. PgUGT74AE2 and PgUGT94Q2 also display regioselectivity, as evidenced by the observations that PgUGT74AE2 glucosylated only C3-OH, and not C12-OH or C20-OH in PPD-type ginsenosides, and PgUGT94Q2 added a glucose molecule to the 3-*O*-glucoside

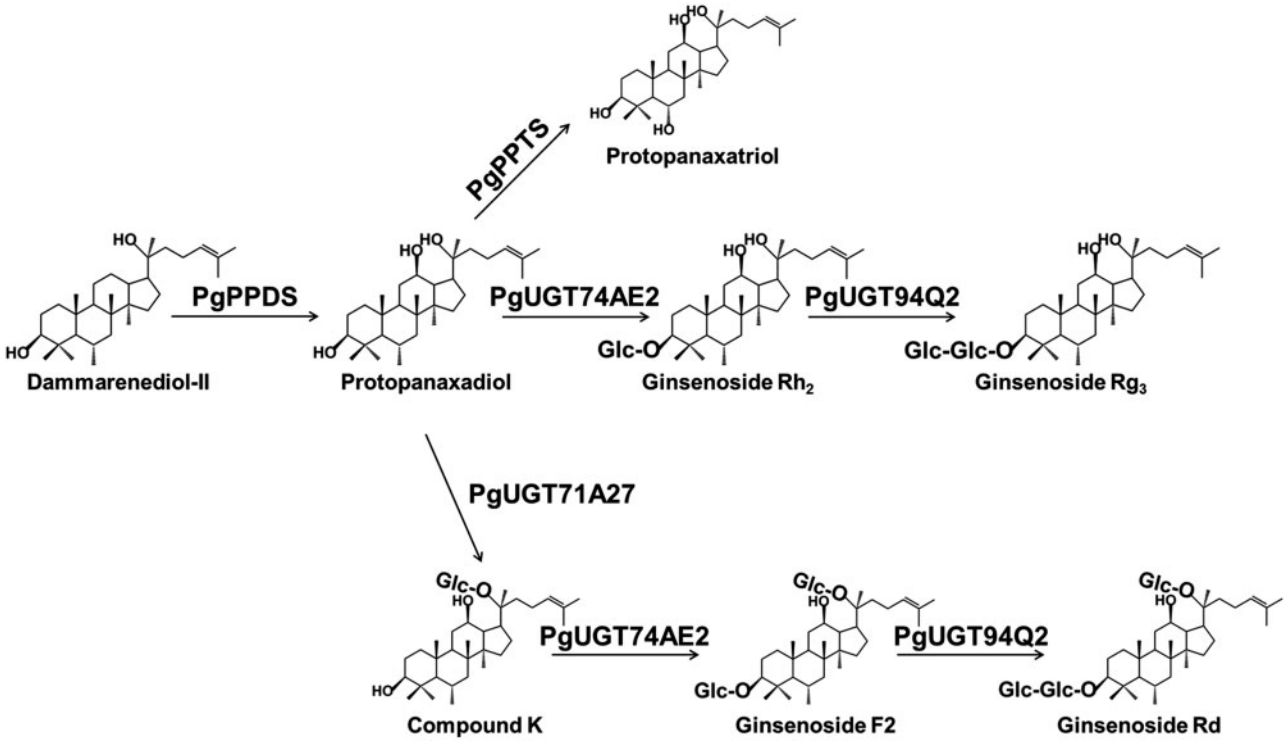

Fig. 8 A diagram showing the ginsenoside biosynthetic pathway.

of PPD-type ginsenosides, but not to the 20-*O*-glucoside of PPD-type ginsenosides.

Though PgUGT74AE2 and PgUGT94Q2 display substrate specificity and regioselectivity, they could glucosylate more than one ginsenoside by recognizing a common structural feature of acceptors or by disregarding different features of acceptors. As we have shown, PgUGT74AE2 glucosylated both PPD and CK, which have a C3-OH in common, but have either a C20-OH (PPD) or C20-*O*-glucoside (CK). Similarly, PgUGT94Q2 attached a glucose moiety to the 3-*O*-glucosyl group of Rh_2 and F2, which have a common C3-*O*-glucoside, but have either C20-OH (Rh_2) or C20-*O*-glucoside (F2). Their ability to glucosylate specific yet different acceptors implies that the organization of their acceptor binding pockets is specific enough to distinguish features such as C6-H and C6-OH, but flexible enough to accommodate both C20-OH and C20-*O*-glucoside. UGTs in the same family also display wide acceptor specificities. Both PgUGT74AE2 and SvUGT74M1 belong to the same UGT74 family and glucosylate triterpenoid saponins. However, two UGTs bind to different saponins and glucosylate different positions: PgUGT74AE2 glucosylates the -OH at the C3 position of PPD, whereas SvUGT74M1 glucosylates the -COOH at the C28 position of gypsogenic acid. The glucosylation of almost opposite ends of triterpene skeletons by these two UGTs indicates that the acceptor binding sites of two UGTs are diverged enough to accommodate the binding of saponins in different orientations. Similarly, UGT74 family members in Arabidopsis have been shown to glucosylate salicylic acids and anthranilate rather than saponins, further supporting that UGTs even in the same family display wide variations in their acceptor specificities. Studies have shown that both a sugar donor and a sugar acceptor are bound to a cleft formed by N- and C-terminal domains of UGT, and mutations in even a single amino acid in these regions affect the specificity of UGTs (Osmani et al. 2008, Modolo et al. 2009, Noguchi et al. 2009, Wang 2009). A detailed biochemical analysis, including an X-ray crystallographic analysis, would be helpful in understanding the acceptor specificity of the two ginseng UGTs.

Kinetic data suggest that the catalytic efficiency of PgUGT74AE2 is relatively low (**Table 1**), whereas that of PgUGT94Q2 is comparable with that of other UGTs (Achnine et al. 2005, Shibuya et al. 2010, Augustin et al. 2012). The low catalytic efficiency of PgUGT74AE2 is due to its low turnover numbers rather than to K_m values. For example, the K_m value of PgUGT74AE2 for CK is 40 μM, which is comparable with K_m values of SvUGT74M1 for gypsogenic acid (170 μM) and UGT71G1 for hederagenin (166 μM). The k_{cat} value of PgUGT74AE2 for CK is lower (0.00013 s^{-1}) than those of SvUGT74M1 (1.13 s^{-1}) and UGT71G1 (0.025 s^{-1}), resulting in low catalytic efficiency (3.37 $M^{-1}s^{-1}$ compared with 6,500 $M^{-1}s^{-1}$ for SvUGT74M1 and 153 $M^{-1}s^{-1}$ for UGT71G1). Unlike PgUGT74AE2, the catalytic efficiency of PgUGT94Q2 is 2,183 $M^{-1}s^{-1}$, which is comparable with those of SvUGT74M1 and UGT71G1. Consistent with the low catalytic efficiency of PgUGT74AE2, the products of PgUGT74AE2 are virtually absent when PgUGT74AE2 and PgUGT94Q2 are incubated together with PPD or CK in vitro (**Fig. 5**). Similarly, yeast harboring both PgUGT74AE2 and PgUGT94Q2 accumulates mainly Rg_3 but not Rh_2.

The two UGTs identified here are not mere enzymatic curiosities. In addition to their roles in the ginsenoside biosynthetic pathway, they provide new tools for engineering ginsenoside biosynthesis to produce specific ginsenosides, both in vitro and de novo in heterologous systems. As we have shown, the two recombinant UGTs could be used to convert PPD to Rh_2 and Rg_3 or CK to F2 and Rd in vitro. Since PPD could be obtained en masse by deglycosylating total ginsenosides, one could specifically produce the above-mentioned ginsenosides in vitro. The two UGT genes combined with other biosynthetic genes could also be used to produce specific ginsenoside de novo in heterologous systems such as microorganisms, plants and other organisms. As an example, we showed that the two *UGT* genes together with *PgDS*, *PgPPDS* and *ATR2* could be used to produce Rg_3. Rg_3 rarely exists naturally in freshly harvested ginseng root, but is present in so-called red ginseng, which is produced by steaming and drying harvested ginseng root. The amount of Rg_3 we produced was about 1.3 mg l^{-1}, which is roughly equivalent to that in two 6-year-old red ginseng roots (Kim et al. 2010). Since we have not optimized the fermentation condition for mass production, the Rg_3 level could presumably be increased through further optimization. Similar to Rg_3, other ginsenosides, such as Rh_2, F2, CK and Rd, might be produced using combinations of UGTs. Alternatively, the activities of the two UGTs could be altered in the ginseng plant itself, through either overexpression or repression, to channel the metabolic flux into specific ginsenosides.

Materials and Methods

Transcriptome analysis

Ginseng leaves and roots were incubated in 1/2 MS (half-strength Murashige and Skoog) liquid medium for 12 h with shaking in the absence or presence of 200 μM MeJA. Total RNA was extracted using an RNA prep kit (Sigma), and 100 nucleotide, paired-end sequencing was done using an Illumina HiSeq2000 system (Illumina Inc.). Contigs were assembled de novo using the Trinity program (version 2011-11-26) (Grabherr et al. 2011). The expression level of the assembled contigs (transcripts) was examined by Bowtie (Langmead et al. 2009) (version 0.12.8), and FPKM (Trapnell et al. 2010) values were calculated using a program containing the RSEM (Li and Dewey 2011) algorithm, distributed by the Trinity group.

For DECs and GO analyses, contigs with FPKM values ≤ 2 were first removed prior to further analysis. Contigs that were induced or repressed >2-fold by MeJA were defined as DECs. For GO analyses, contigs were mapped to Arabidopsis genes using BLAST (e-values $\leq 10^{-5}$), and the resulting 1,191 genes were used for BiNGO analysis; *P*-values were adjusted for multiple testing using the Benjamini and Hochberg's method to control for false discovery rate (Benjamini 1995).

In vitro enzymatic assay of PgUGTs

The two UGT genes were cloned into pGEX4T-1 using the primer pairs PgUGT74AE2-*Bam*HI/PgUGT74AE2-*Eco*RI and PgUGT94Q2-*Bam*HI/PgUGT94Q2-*Eco*RI (**Supplementary Table S4**). Recombinant PgUGT74AE2 and PgUGT94Q2 proteins were purified from *Escherichia coli* BL21-CodonPlus (DE3)-RIL by glutathione–Sepharose affinity chromatography (GE Healthcare) according to the manufacturer's instructions.

Enzyme assays were performed in a reaction buffer (10 mM PBS buffer, pH 7) containing purified PgUGT74AE2 or PgUGT94Q2, ginsenoside (5 mM) and UDP-Glc (50 mM). The 10 different ginsenosides used in the assay were PPD, PPT, CK, Rg_3, Rh_2, F2, Rd, Rg2, Rh_1 and F1 (see **Supplementary Table S1** for structures). Reaction mixtures were incubated at 35°C for 12 h, and the products were analyzed by either TLC or HPLC. TLC was performed using $60F2_{54}$ silica gel plates (Merck) with a mobile phase of acetone : methanol : double-distilled water (65 : 35 : 10, v/v). Products resolved on TLC plates were detected by spraying with 10% (v/v) H_2SO_4, followed by heating at 110°C for 5 min (Wang et al. 2011). HPLC analyses were performed using an ODS(2) C18 column (Phenomenex) at a flow rate of 1 ml min^{-1} as follows: 0 min, 68% water and 32% acetonitrile; 8 min, 35% water and 65% acetonitrile; 12 min, 0% water and 100% acetonitrile; 20 min, 0% water and 100% acetonitrile; 20.1 min, 68% water and 32% acetonitrile; and 28 min, 68% water and 32% acetonitrile. Ginsenosides were monitored at a wavelength of 203 nm with a UV-detector (Agilent Technologies).

LC-ESI-MS/MS analysis

Identification of each ginsenoside was performed using an HPLC-MS/MS system composed of an HPLC System (HP1100; Agilent Technologies), a triple-quadrupole tandem mass spectrometer (API-2000; Applied Biosystem) equipped with an autosampler, a heated electrospray ionization source (H-ESI), a triple-stage quadrupole mass analyzer, and Analyst 1.4 software for data acquisition. A reversed-phase column (Fortis H2o C18, 2.1×100 mm, 3 μm pore size; Fortis Technologies Ltd.) was used for sample separation. The mobile phase for chromatographic separation consisted of 0.01% acetic acid aqueous water (A) and 0.01% acetic acid aqueous acetonitrile (B). The gradient elution program, with a constant flow rate of 250 μl min^{-1}, was as follows: 0 min, 68% A and 32% B; 3 min, 45% A and 55% B; 8 min, 40% A and 60% B; 13 min, 20% A and 80% B; 18 min, 0% A and 100% B; 22 min, 0% A and 100% B; 22.1 min, 68% A and 32% B; 30 min, 68% A and 32% B. The column temperature was set at 25°C. To detect ginsenoside compounds (PPD, Rh_2, Rg_3, CK, F2 and Rd), the MRM method was used. The transitions were set at *m/z* 461.1→425.5 for PPD, at *m/z* 645.3→23.2 for Rh2, at *m/z* 807.4→365.3 for Rg3, at *m/z* 645.4→23.2 for CK, *m/z* 807.5→627.5 for F2 and at *m/z* 969.3→789.4 for Rd, respectively. For full-scan MS analyses, spectra were recorded in the *m/z* range from 400 to 1,000 according to the transition pattern (Kim et al. 2012). The MS/MS conditions were optimized by introducing a standard solution of analyte via a syringe pump at 10 μl min^{-1}. The ES-MS parameters were as follows: ion spray voltage, −4,200 V; ion source gas 1, 20; curtain gas, 20; collision gas, 2.

De novo synthesis of ginsenoside in *S. cerevisiae*

To repress the endogenous yeast lanosterol synthase gene (*ERG7*) by methionine, we replaced the *ERG7* promoter with the *MET3* promoter by homologous recombination, as described previously (Mao et al. 2002, Ro et al. 2006). After the replacement, yeast cells were selected on 5-fluoroorotic acid plates to recover the *URA3* marker, and the resulting yeast cells were used for further analysis.

To clone ginsenoside biosynthetic genes, we amplified *PgDS*, *PgPPDS*, *PgUGT74AE2* and *PgUGT94Q2* from cDNA using specific gene primer sets (**Supplementary Table S4**). *ATR2* was amplified from Arabidopsis cDNA, and the N-terminally truncated HMG-CoA reductase (*tHMGR*) gene was amplified from yeast genomic DNA using specific primer sets (**Supplementary Table S4**). After each gene was linked to the *GPD* promoter, *PgUGT74AE2* and *PgUGT94Q2* were cloned into the same pRS423 vector; *PgDS* was cloned into pRS424; *PgPPDS* and *ATR2* were cloned into pRS425; and *tHMGR* was cloned into pRS426. The three resulting vectors were introduced into the *ERG7*-repressible yeast line.

De novo synthesis of ginsenoside was examined by culturing *S. cerevisiae* harboring ginsenoside biosynthetic genes in a bioreactor. Defined medium, described by van Hoek et al., containing methionine was used for the fermentation (van Hoek et al. 2000, Lenihan et al. 2008, Westfall et al. 2012). The batch culture medium contained 20 g l^{-1} glucose, 15 g l^{-1} $(NH_4)_2SO_4$, 8 g l^{-1} KH_2PO_4, 0.72 g l^{-1} $ZnSO_4 \cdot 7H_2O$, 6.15 g l^{-1} $MgSO_4 \cdot 7H_2O$, 12 ml l^{-1} vitamin solution, 0.3 g l^{-1} methionine and 10 ml l^{-1} trace metal solution. Trace metal solution for the medium contained 15 g l^{-1} EDTA, 10.2 g l^{-1} $ZnSO_4 \cdot 7H_2O$, 0.50 g l^{-1} $MnCl_2 \cdot 4H_2O$, 0.5 g l^{-1} anhydrous $CuSO_4$, 0.86 g l^{-1} $CoCl_2 \cdot 6H_2O$, 0.56 g l^{-1} $Na_2MoO_4 \cdot 2H_2O$, 3.84 g l^{-1} $CaCl_2 \cdot 2H_2O$ and 5.12 g l^{-1} $FeSO_4 \cdot 7H_2O$. Vitamin solution contained 0.05 g l^{-1} biotin, 1 g l^{-1} calcium pantothenate, 1 g l^{-1} nicotinic acid, 25 g l^{-1} myoinositol, 1 g l^{-1} thiamine HCl, 1 g l^{-1} pyridoxol HCl and

0.2 g l^{-1} 4-aminobenzoic acid. Feeding solution contained 578 g l^{-1} glucose, 9 g l^{-1} KH_2PO_4, 3.5 g l^{-1} K_2SO_4, 0.28 g l^{-1} Na_2SO_4, 5.12 g l^{-1} $MgSO_4 \cdot 7H_2O$ and 1 g l^{-1} methionine. For batch-fed fermentation, seed cultures were grown for 2 d at 30°C, and 20 ml of seed cultures were inoculated into 1 liter of batch medium. Batch-fed fermentation was performed in a 1.5 liter Bioreactor (Biotron) at 30°C with airflow of 1 liter min^{-1} and 400 r.p.m. agitation. The pH-stat feeding strategy was used for batch-fed fermentation by feeding glucose (>pH 5.51) and by adding 5% NH_4OH (<pH 5.4).

For the analysis of de novo synthesized ginsenosides, 50 ml of cells were harvested by centrifugation (10 min, 2,898 × *g*), resuspended in 20 ml of double-distilled water, and disrupted by sonication (Vibra-cell; Sonics & Materials). Metabolites were extracted with 50% methanol (v/v), purified on a SEP-PAK C18 cartridge, and analyzed using HPLC and LC-MS/MS.

Accession/AGI numbers

UGT74E1 (AT1G05675), UGT74C1 (AT2G31790), UGT74D1 (AT2G31750), UGT74B1 (AT1G24100), UGT74F2 (AT2G43820), UGT74F1 (AT2G43820), UGT73C2 (AT2G36760), UGT73C7 (AT3G53160), UGT73D1 (AT3G53150), UGT73B4 (AT2G15490), UGT73B2 (AT4G34135), UGT73B1 (AT4G34138), UGT91B1 (AT5G65550), UGT91A1 (AT2G22590), UGT91C1 (AT5G49690), UGT79B1 (AT5G54060), UGT79B2 (AT4G27560), UGT79B5 (AT1G50580), UGT79B8 (AT2G22930), SiUGT94D1 (BAF99027), BpUGT94B1 (Q5NTH0), CaUGT3 (BAH80312), Cm1,2RhaT (AFB73772), GmUGT73P2 (BAI99584), MtUGT71G1 (AAW56092), MtUGT73K1 (AAW56091), SvUGT74M1 (ABK76266), ZmUGT74A1 (NP_001105326), PgUGT4 (KM491306), PgUGT19 (KM491308), PgUGT31 (KM491307), PgUGT71A27 (KM491309), PgUGT74AE2 (JX898529) and PgUGT94Q2 (JX898530).

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the Intelligent Synthetic Biology Center of Global Frontier Project [grant No. 2011-0031955 to S.K.]; the BioGreen 21 Program [SSAC, PJ008110 to J.L.]; the Basic Science Research Program [grant No. NRF-2012R1A2A1A01003133 to G.C.].

Disclosures

The authors have no conflicts of interest to declare.

References

Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W. and Dixon, R.A. (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *Plant J.* 41: 875–887.

Augustin, J.M., Drok, S., Shinoda, T., Sanmiya, K., Nielsen, J.K., Khakimov, B. et al. (2012) UDP-glycosyltransferases from the UGT73C subfamily in *Barbarea vulgaris* catalyze sapogenin 3-O-glucosylation in saponin-mediated insect resistance. *Plant Physiol.* 160: 1881–1895.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. Ser. B. Methodol.* 57: 289–300.

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

Choi, J., Kim, T.H., Choi, T.Y. and Lee, M.S. (2013) Ginseng for health care: a systematic review of randomized controlled trials in Korean literature. *PLoS One* 8: e59978.

Dai, Z., Liu, Y., Zhang, X., Shi, M., Wang, B., Wang, D. et al. (2013) Metabolic engineering of *Saccharomyces cerevisiae* for production of ginsenosides. *Metab. Eng.* 20: 146–156.

Dai, Z., Wang, B., Liu, Y., Shi, M., Wang, D., Zhang, X. et al. (2014) Producing aglycons of ginsenosides in bakers' yeast. *Sci. Rep.* 4: 3698.

De Geyter, N., Gholami, A., Goormachtig, S. and Goossens, A. (2012) Transcriptional machineries in jasmonate-elicited plant secondary metabolism. *Trends Plant Sci.* 17: 349–359.

Fuzzati, N. (2004) Analysis methods of ginsenosides. *J. Chromatogr. B* 812: 119–133.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I. et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29: 644–652.

Han, J.Y., Hwang, H.S., Choi, S.W., Kim, H.J. and Choi, Y.E. (2012) Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol.* 53: 1535–1545.

Han, J.Y., Kim, H.J., Kwon, Y.S. and Choi, Y.E. (2011) The Cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol.* 52: 2062–2073.

Hou, B., Lim, E.K., Higgins, G.S. and Bowles, D.J. (2004) N-glucosylation of cytokinins by glycosyltransferases of *Arabidopsis thaliana*. *J. Biol. Chem.* 279: 47822–47832.

Kim, J.K., Cui, C.H., Yoon, M.H., Kim, S.C. and Im, W.T. (2012) Bioconversion of major ginsenosides Rg_1 to minor ginsenoside F1 using novel recombinant ginsenoside hydrolyzing glycosidase cloned from *Sanguibacter keddieii* and enzyme characterization. *J. Biotechnol.* 161: 294–301.

Kim, M.H., Hong, H.D., Kim, Y.C., Rhee, Y.K., Kim, K.T. and Rho, J. (2010) Ginsenoside changes in red ginseng manufactured by acid impregnation treatment. *J. Ginseng Res.* 34: 93–97.

Kim, Y.S., Hahn, E.J., Murthy, H.N. and Paek, K.Y. (2004) Adventitious root growth and ginsenoside accumulation in *Panax ginseng* cultures as affected by methyl jasmonate. *Biotechnol. Lett.* 26: 1619–1622.

Kim, Y.S., Yeung, E.C., Hahn, E.J. and Paek, K.Y. (2007) Combined effects of phytohormone, indole-3-butyric acid, and methyl jasmonate on root growth and ginsenoside production in adventitious root cultures of *Panax ginseng* C.A. Meyer. *Biotechnol. Lett.* 29: 1789–1792.

Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10: R25.

Lee, Y., Yoon, H.R., Paik, Y.S., Liu, J.R., Chung, W.I. and Choi, G. (2005) Reciprocal regulation of *Arabidopsis* UGT78D2 and BANYULS Is critical for regulation of the metabolic flux of anthocyanidins to condensed tannins in developing seed coats. *J. Plant Biol.* 48: 356–370.

Lenihan, J.R., Tsuruta, H., Diola, D., Renninger, N.S. and Regentin, R. (2008) Developing an industrial artemisinic acid fermentation process to support the cost-effective production of antimalarial artemisinin-based combination therapies. *Biotechnol. Prog.* 24: 1026–1032.

Lenka, S.K., Boutaoui, N., Paulose, B., Vongpaseuth, K., Normanly, J., Roberts, S.C. et al. (2012) Identification and expression analysis of methyl jasmonate responsive ESTs in paclitaxel producing *Taxus cuspidata* suspension culture cells. *BMC Genomics* 13: 148.

Li, B. and Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12: 323.

Li, Y., Baldauf, S., Lim, E.K. and Bowles, D.J. (2001) Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. *J. Biol. Chem.* 276: 4338–4343.

Mackenzie, P.I., Bock, K.W., Burchell, B., Guillemette, C., Ikushiro, S., Iyanagi, T. et al. (2005) Nomenclature update for the mammalian

UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet. Genomics* 15: 677–685.

Mackenzie, P.I., Owens, I.S., Burchell, B., Bock, K.W., Bairoch, A., Belanger, A. et al. (1997) The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics* 7: 255–269.

Mao, X., Hu, Y., Liang, C. and Lu, C. (2002) MET3 promoter: a tightly regulated promoter and its application in construction of conditional lethal strain. *Curr. Microbiol.* 45: 37–40.

Meesapyodsuk, D., Balsevich, J., Reed, D.W. and Covello, P.S. (2007) Saponin biosynthesis in *Saponaria vaccaria*. cDNAs encoding beta-amyrin synthase and a triterpene carboxylic acid glucosyltransferase. *Plant Physiol.* 143: 959–969.

Modolo, L.V., Escamilla-Trevino, L.L., Dixon, R.A. and Wang, X. (2009) Single amino acid mutations of *Medicago* glycosyltransferase UGT85H2 enhance activity and impart reversibility. *FEBS Lett.* 583: 2131–2135.

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.

Noguchi, A., Horikawa, M., Fukui, Y., Fukuchi-Mizutani, M., Iuchi-Okada, A., Ishiguro, M. et al. (2009) Local differentiation of sugar donor specificity of flavonoid glycosyltransferase in Lamiales. *Plant Cell* 21: 1556–1572.

Osmani, S.A., Bak, S., Imberty, A., Olsen, C.E. and Moller, B.L. (2008) Catalytic key amino acids and UDP-sugar donor specificity of a plant glucuronosyltransferase, UGT94B1: molecular modeling substantiated by site-specific mutagenesis and biochemical analyses. *Plant Physiol.* 148: 1295–1308.

Quiel, J.A. and Bender, J. (2003) Glucose conjugation of anthranilate by the *Arabidopsis* UGT74F2 glucosyltransferase is required for tryptophan mutant blue fluorescence. *J. Biol. Chem.* 278: 6275–6281.

Ro, D.K., Ehlting, J. and Douglas, C.J. (2002) Cloning, functional expression, and subcellular localization of multiple NADPH-cytochrome P450 reductases from hybrid poplar. *Plant Physiol.* 130: 1837–1851.

Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M. et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440: 940–943.

Ross, J., Li, Y., Lim, E. and Bowles, D.J. (2001) Higher plant glycosyltransferases. *Genome Biol.* 2 REVIEWS3004.

Sawada, S., Suzuki, H., Ichimaida, F., Yamaguchi, M.A., Iwashita, T., Fukui, Y. et al. (2005) 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.

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.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731–2739.

Tansakul, P., Shibuya, M., Kushiro, T. and Ebizuka, Y. (2006) Dammarenediol-II synthase, the first dedicated enzyme for ginsenoside biosynthesis, in *Panax ginseng*. *FEBS Lett.* 580: 5143–5149.

Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P. and Osbourn, A. (2014) Triterpene biosynthesis in plants. *Annu. Rev. Plant Biol.* 65: 225–257.

Tohge, T., Nishiyama, Y., Hirai, M.Y., Yano, M., Nakajima, J., Awazuhara, M. et al. (2005) Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. *Plant J.* 42: 218–235.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J. et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28: 511–515.

van Hoek, P., de Hulster, E., van Dijken, J.P. and Pronk, J.T. (2000) Fermentative capacity in high-cell-density fed-batch cultures of baker's yeast. *Biotechnol. Bioeng.* 68: 517–523.

Vranova, E., Coman, D. and Gruissem, W. (2013) Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* 64: 665–700.

Wang, L., Liu, Q.-M., Sung, B.-H., An, D.-S., Lee, H.-G., Kim, S.-G. et al. (2011) Bioconversion of ginsenosides Rb_1, Rb_2, Rc and Rd by novel β-glucosidase hydrolyzing outer 3-O glycoside from *Sphingomonas* sp. 2F2: cloning, expression, and enzyme characterization. *J. Biotechnol.* 156: 125–133.

Wang, X. (2009) Structure, mechanism and engineering of plant natural product glycosyltransferases. *FEBS Lett.* 583: 3303–3309.

Westfall, P.J., Pitera, D.J., Lenihan, J.R., Eng, D., Woolard, F.X., Regentin, R. et al. (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl Acad. Sci. USA* 109: E111–E118.

Yan, X., Fan, Y., Wei, W., Wang, P., Liu, Q., Wei, Y. et al. (2014) Production of bioactive ginsenoside compound K in metabolically engineered yeast. *Cell Res.* 24: 770–773.

Yonekura-Sakakibara, K., Tohge, T., Niida, R. and Saito, K. (2007) Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in *Arabidopsis* by transcriptome coexpression analysis and reverse genetics. *J. Biol. Chem.* 282: 14932–14941.