Supplementary Information, Data S1 Materials and methods

Prediction of UDP-glycosyltransferases

To construct a *Panax* cDNA database, nine *Panax* sequencing datasets were collected from the NCBI database (Supplementary Table 2). The Roche 454 reads were assembled using the GS de novo Assembler (version 2.3), Sanger reads were assembled with CAP3 [1] and reads from the illumina platform were assembled with Trinity (version r2012-06-08) [2]. A sequence pool was generated by combining the primitive assembly results and the *Panax* sequences (mRNAs and ESTs) retrieved from NCBI. A secondary assembly for the sequence pool was implemented using CAP3 and GS de Novo Assembler. All the assemblies were performed with default parameters. Finally, a *Panax* cDNA database with non-redundant sequence was generated.

ORF of cDNA sequences in the database was predicted by BESTORF (http://linux1.softberry.com/berry.phtml?topic=bestorf&group=programs&subgroup=gfind). ORFs encoding UDP-glycosyltransferases were identified by conserved PSPG box [3]. Protein sequences of the selected ORFs were aligned with MAFFT (http://mafft.cbrc.jp/alignment/software/). After calculating pairwise distance between these aligned protein sequences, they were classified into operational taxonomic units using DOTUR program (OTUs, 95% cutoff).

P. ginseng tissue culture and cloning of UDP-glycosyltransferase encoding cDNAs

Hormone-autotrophic (H; habituated) *P. ginseng* callus was cultured on hormone-free 67V medium and its callus line was maintained at 24°C in dark and subcultured once for every 28 days [4]. *P. ginseng* callus cells were frozen immediately in liquid nitrogen and ground into powder for isolation of total RNA using Trizol (Invitrogen, Carlsbad, CA, USA) as the manufacturer's instruction and then converted into cDNA using PrimeScript® RT reagent Kit with gDNA eraser (Takara, Dalian, China). The *P. ginseng* cDNA was used as the template to amplify the cDNAs potentially encoding UDP-glycosyltransferase with long ORFs (likely to be full length) and the cloned ORFs were then inserted into pMD18T independently. These cloned genes were sequenced and were designated corresponding GenBank accession numbers ranging from KF377581 to KF377596. The primers to amply the encoding genes or ORFs of the corresponding UDP-glycosyltransferases were designed according to the sequences of KF377581 to KF377596.

Heterologous expression of the UGTPg1 in *E. coli* BL21 (DE3)

The ORF of UGTPg1 was cloned into pET28a (Merck KGaA, Darmstadt, Germany). The C-terminal 6×His-tagged expression vector was transformed into *E. coli* BL21 (DE3). The recombinant *E. coli* BL21 (DE3) strain was cultured in LB medium (plus 50 μg/mL kanamycin) at 37°C at 200 rpm until the OD_{600} reached 0.6-0.8. The culture was cooled on ice, and 50 μM IPTG was added. After incubation at 16°C at 110 rpm for 18h, the cells were harvested by centrifugation at 4°C and suspended in 100 mM phosphate buffer (pH8.0), 1 mM PMSF and disrupted by

French Press (25 kpsi). The cell debris was removed by centrifugation at 12000 *g* for 20 min, and the supernatant was used for enzymatic assays. The pET28a-transformed *E. coli* BL21 (DE3) cells were treated in parallel as a control.
The UGTPg1 in the crude cell extract was quantified by indirect enzyme-linked immunosorbent assay (ELISA). The supernatant of the *E. coli* cell extract was diluted ten times in the antigen coating buffer (15 mM carbonate buffer, pH9.6), and a serial dilution samples (0, 2, 4, 6, 8, 10 ng/μL) of a C-terminal 6×His-tagged xylanase (53 kD) was used to draw a standard curve for quantitative reference. Equivalent volumes of samples and the standards were added into the wells of a Nunc-Immuno Plate (Thermo Fisher Scientific Inc., Waltham, MA, USA) and incubated overnight at 4°C. Each well was blocked with blocking buffer (5% non-fat milk in antigen coating buffer) at 37°C for 2 h. Then, the plate was washed with PBST buffer (20 mM phosphate, 150 mM NaCl, 0.05% Tween-20, pH7.4). The anti-6×His primary antibody was added into the wells and incubated at 37°C for 2 h. The plate was washed with PBST buffer followed by the addition of secondary antibody (Horseradish Peroxidase-conjugated) and incubated at 37°C for 1 h. The plate was washed with PBST buffer and the substrate 3,3',5,5'-Tetramethylbenzidine (TMB) was added into each well and the color detection and quantification was followed by the kit manual (Beyotime, Haimen, China).

The enzymatic assays for the UGTPgs

The reaction was carried out in a total volume of 100 μL containing 100 mM phosphate buffer (pH 8.0), 1% Tween-20, 5 mM UDP-glucose, 0.5 mM acceptor substrate and the recombinant *E. coli* extract for 2 h at 35°C and terminated by adding 100 μL *n*-butanol. The product was extracted and evaporated. The residue was dissolved in methanol for TLC and HPLC analysis. For kinetic study, the reaction mixture contained 100 mM phosphate buffer (pH8.0), 1% Tween-20, 5 mM UDP-glucose, 125 ng of the indirect ELISA-quantified UGTPg1 and acceptor substrate (40-500 μM) in a final volume of 100 μL. Reaction was incubated at 40°C for 15 min. HPLC analysis was used to quantify target product in each reaction. The Michaelis-Menten parameters were calculated by least square fitting of the kinetic model using Prism 5 (GraphPad, San Diego, CA, USA). All data are presented as mean ± SD of three independent repeats.

Construction of vectors (Supplementary Information, Table S6) and yeast strains (Supplementary Information, Table S7)

All the primers, vectors and strains used in this study were listed in the Supplementary Information, Table S6, Table S7 and Table S8, respectively.
pKS-ATR21: NADPH-cytochrome P450 reductase *ATR2-1*[5] was PCR amplified with primers CPRF and CPRR using *Arabidopsis thaliana* cDNA as the template and then ligated into plasmid pBlu2KSp.
pKS-SQE1: *Panax ginseng* squalene epoxidase gene *SQE1* was PCR amplified with primers SQE1-F and SQE1-R using *Panax ginseng* cDNA as the template and then ligated into plasmid pBlu2KSp.

BA21: Fragment TrpU-URA-GAL1p was PCR amplification as follows. PCR amplification of TrpU fragment was performed with primers P-TRPUF and TRPUR using BY4742 chromosome DNA as the template. PCR amplification of *URA3* marker was performed with primers URAF-TPR1 and OURAR using plasmid pYES2.1 as the template. PCR amplification of *GAL1* promoter was performed with primers OGAL1F and OGAL1R using plasmid pYES2.1 as the template. For the secondary PCR, 25 ng each of the purified PCR products mentioned above were used as the template and PCR amplified with primers P-TRPUF and OGAL1R to give TrpU-URA-GAL1p.
Fragment GAL1p-ATR21-TrpD was PCR amplification as follows. PCR amplification of *GAL1* promoter was performed as before. PCR amplification of *ATR21* ORF was performed with primers OA21F and OA21R-TRP1using plasmid pKS-ATR21 as the template. PCR amplification of TrpD fragment was performed with primers TRPDF and P-TRPDR using BY4742 chromosome DNA as the template. For the secondary PCR, 25 ng each of the purified PCR products mentioned above were used as the template and PCR amplified with primers OGAL1F and P-TRPDR to give GAL1p-ATR21-TrpD.
The fragment TrpU-URA-GAL1p was co-transformed with the fragment GAL1p-ATR21-TrpD into BY4742 to construct TrpU-URA-GAL1p-ATR21-TrpD and inserted in the *trp1* locus of the strain via homologous recombination, the result strain BA21 was verified with PCR.
pHCD: Plasmid pHCD was constructed as described in previous study (Han et al. 2011) [6] with some modification, the ORF of *CYP716A47* gene was amplified with CYPC-F (*Bam*HI site in 5' end) and CYPC-R(*Xho*I site in 5' end) using *Panax ginseng* cDNA as the template, the fragment was digested with *Bam*HI and *Xho*I and inserted into *Bam*HI and *Xho*I sites of vector pESC-HIS, the result plasmid was named pESC-HIS-CYP. Similarly, the ORF of *Panax ginseng* dammarenediol synthase gene *DDS* was amplified with primers ESC-PgDDSF (*Not*I site in 5' end) and ESC-PgDDSR(*Sac*I site in 5' end) and subcloned into pESC-HIS-CYP.
pMD18T-UGTPgs: the ORF of UGTPgx (x represent 1-16) was amplified from the cDNA of *Panax ginseng* and was cloned into pMD18T for sequencing.
pYES2.1-UGTPgs: C-terminally 6xHis-tagged UGTPgx (x represent 1-16) was amplified from pMD18T-UGTPgx and was subcloned into pYES2.1 with specific restriction sites.
pHCD-UGTPg1: PCR of the flanking sequences of *UGTPg1*was carried out using pYES2.1-UGTPg1 as the template with primers gaGTF and TD2T-GTR, primers GT-TDH2tF and FB1p-TDH2tR using BY4742 chromosome DNA as the template, primers TD2t-FBA1pF and CYP-FBA1pR using chromosome DNA as the template, and then with primers gaGTF and CYP-FBA1pR using the PCR fragment mentioned before as template, the following PCR fragment was transformed with pHCD (*Nco*I digested) into BA21, and the plasmid pHCD-UGTPg1 was constructed via homologous recombination and a *Asi*SI digestion site inserted between *UGTPg1* ORF and *TDH2* terminator, the result strain was AK1.
pHCD-UGTPgs: The other full-length ORFs of UDP-glucosyltransferases (15

UGTPgs) were cloned into pYES2.1 vector in the downstream of *GAL1* promoter (6×His tag were added to the C-terminal of the UGTPgs). The UGTPgs-6×His were amplified by PCR and then co-transformed with the vector pHCD-UGTPg1(after *Asi*SI digestion) into the strain BA21 to construct the expression vector pHCD-UGTPgs via homologous recombination. The resulted BA21strains harboring pHCD-UGTPgs were cultured and screened for the production of novel protopanaxadiol (PPD)-type ginsenosides.

pHCDA: Fragment TEF2p-ATR21-TPI1t was PCR amplification as follows. PCR amplification of *TEF2* promoter was performed with primers p-TEF2PF and A21-TEF2PR using BY4742 chromosome DNA as the template. PCR amplification of *ATR2-1* ORF was performed with primers TEF2p-A21F and TPI1t-A21R using plasmid pKS-ATR21 as template. PCR amplification of *TPI1* terminator was performed with primers A21-TPI1tF and p-TPItR using BY4742 chromosome DNA as the template. For the secondary PCR, 25 ng each of the purified PCR products mentioned above were used as the template and PCR amplified with primers p-TEF2pF and p-TPItR to give TEF2p-ATR21-TPI1t. The fragment was co-transformed with the plasmid pHCD (after *Mfe*I digestion) into BY4742 to construct the plasmid pHCDA via homologous recombination.

pHSCDA: Fragment GPM1p-rc-TPI1p-SQE1-ENO2t-TEF1p was PCR amplification as follows. PCR amplification of *GPM1* promoter was performed with primers TPI1p-GPM1pF and CYP-GPM1pR using BY4742 chromosome DNA as the template. PCR amplification of *TPI1* promoter was performed with primers GPM1p-TPI1pF and SQE-TPI1pR using BY4742 chromosome DNA as the template. PCR amplification of *SQE1* ORF was performed with primers TPI1p-SQE1F and ENO2t-SQE1R using plasmid pKS-SQE1 as the template. PCR amplification of *ENO2* terminator was performed with primers SQE-ENO2tF and TEF1p-ENO2tR using BY4742 chromosome DNA as the template. PCR amplification of *TEF1*promoter was performed with primers ENO2t-TEF1PF and DDS-TEF1PR using plasmid pLLeu-tHMGR-UPC2.1 as the template. For the secondary PCR, 25 ng each of the purified PCR products mentioned above were used as the template and PCR amplified with primers CYP-GPM1pR and DDS-TEF1pR to give GPM1p-rc-TPI1p-SQE1-ENO2t-TEF1p. The fragment was co-transformed with the plasmid pHCDA (after *Nco*I and *Not*I digestion) into BY4742 to construct the plasmid pHSCDA via homologous recombination.

pHCDR1 and pHCDR1: being constructed similarly as pHCDA, modified by using two NADPH-cytochrome P450 reductase encoding genes from *P. ginseng*, PgCPR1 and PgCPR2, to replace the *ATR2-1* ORF from *A. thaliana* in pHCDA.

pHSCDAG: Fragment TDH2t-FBA1p-UGTPg1 was PCR amplification as follows. PCR amplification of *TDH2* terminator was performed with primers CYP-TDH2tF and FB1p-TDH2tR using BY4742 chromosome DNA as the template. PCR amplification of *FBA1* promoter was performed with primers TD2t-FBA1pF and UGTPg1-FBA1pR using BY4742 chromosome DNA as the template. PCR amplification of *UGTPg1* ORF was performed with primers FB1p-UGTPg1F and CYC1t-UGTPg1R using pYES2.1-UGTPg1 as the template. For the secondary PCR,

25 ng each of the purified PCR products mentioned above were used as the template and PCR amplified with primers CYP-TDH2tF and CYC1t-UGTPg1R to give TDH2t-FBA1p-UGTPg1. The fragment was co-transformed with the plasmid pHSCDA (after *Xho*I digestion) into BY4742 to construct the plasmid pHSCDA via homologous recombination and the result strain BK1.

Yeast cultivation

To determine the performance of the genetically engineered yeast strains, individual clones were inoculated into the SC medium lacking the corresponding nutrition (2% glucose as the carbon source) and cultivated at 30°C, 250 rpm for 24 h. Aliquots were diluted to an initial OD_{600} of 0.05 in 50 mL of the SC medium in 250 mL flasks and grown at 30°C, 250 rpm for 4 days. The cell pellets were harvested and extracted by *n*-butanol and subjected to further analysis. The strains AK1 and AKE were cultured in SC medium with 2% galactose, while the strains BK1 and BKE were cultured in SC medium with 2% glucose.

Preparation of yeast microsomes and enzyme assays for CYP716A47

The preparation of yeast microsomes and the enzymatic assays of CYP716A47 were performed as described by Han et al.(2011)[6] with some modifications. The cells of the strain AC1 (BA21 harboring pESC-HIS-CYP) was broken by mortar grinding in liquid nitrogen. The *in vitro* activities of the CYP716A47 was assayed in 500 μL reactions containing 100mM PBS buffer (pH7.4), 1 mM NADPH, 0.5mM substrate (dammarenediol-II (DM) or 20(*S*)-O-β-(D-glucosyl)-dammarenediol II (DMG)) and 1 mg of microsome preparation. The reactions were incubated at 30°C for 36h and products were extracted by adding 500μL *n*-butanol. The upper layer was collected and evaporated under vacuum. The residues were dissolved in methanol for HPLC analysis. For comparison of activities of the CYP716A47 on DM and DMG, 0.5 mM DM and DMG were both added to reactions and were incubated at 30°C for 0, 0.5, 1, 2 and 36 hours.

TLC analysis

The TLC analysis was performed using silica gel 60 F254 plates (Merck KGaA, Darmstadt, Germany) with chloroform/methanol/water (30/10/1, v/v) as the developing solvent. The spots on the TLC plates were visualized by spraying with 1% (w/v) vanillin in H_2SO_4 /ethanol (6/100, v/v) followed by heating at 110°C for 5 min. For identification, a mixture of ginsenoside authentic samples containing compound K, Rh2, F2, Rg3, Rb1, Rb2, Rb3, Rc and Rd and PPD of concentration 0.2 mg/mL (dissolved in methanol) was also spotted on each plate. Ginsenoside authentic samples were purchased from Nanjing Zelang Medical Technology Co., Ltd. (China). DM was purchased from BioBioPha Co.,Ltd. (China). DMG was obtained from the glucosylation of dammarenediol II catalyzed by UGTPg1.

HPLC analysis

HPLC analysis was performed on a Shimadzu LC-20A prominence system (Shimadzu,

Kyoto, Japan) equipped with a binary pump, an online degasser, an autoplate-sampler and a thermostatically controlled column compartment. Chromatographic separation was carried out at 35°C on an Shodex C18-120-5 4E column (5μm , 4.6 mm × 250mm). The gradient elution system consisted of water (A) and acetonitrile (B). Separation for extract of enzyme reaction and yeast fermentation was achieved using the following gradient: 0 min (35% B), 55 min(90% B), 50-55min(90% B) and 55-65min(35% B). The flow rate was kept at 0.8 mL/min. Authentic ginsenosides were subjected to the same conditions.

Semi-preparative HPLC was performed on a Waters 1525 binary pump system with a Waters 2489 detector (210 nm) using a YMC-Pack ODS-A (250 × 10 mm, S-5 μm) or a Daicel CHIRALPAK® AD-H (250 × 10 mm, S-5 μm) columns, and the mobile phases were 70% and 80% acetonitrile in water for CK and DMG, respectively. The products were monitored at 203 nm.

HPLC/ESIMS analysis

Chromatographic analysis was performed on an Agilent 1200 Series LC system (Agilent Technologies, Waldbronn, Germany) equipped with a Shodex C18-120-5 4E column. The ESIMS data were acquired using a LCT Premier XEmass spectrometer (Waters Corporation, Massachusetts, USA) in the positive ionization mode.

NMR analysis

NMR experiments were performed in pyriding-d_5 for CK or CD_3OD for DM and DMG on a Bruker Avance III 400 (for ^{1}H NMR) or a Bruker Avance III 500 (for ^{13}C NMR) (Bruker, Billerica, MA, USA) referenced to the solvent peaks. The 1D and 2D NMR spectra of the related compounds were provided in Figure S4 and Figure S7.

Supplementary Information, Table S6. Primers used in this study.

Primers	Sequence 5'-3'
P-TRPUF	CTTCTGTTCCATGTCGACGCCCGGGCCCTATAGTGAGT CGAGTTAGGGACAGTTAGAGGC
TRPUR	AATAAAAAAAAAATGATGAATTGAATTGAAAAGCTAGC ATAATACCCAGCAAGTCAGCAT
URAF-TPR1	TGCTCTAGATTCCGATGCTGACTTGCTGGGTATTAT GCTAGCTTTTCAATTCAATTCATC
OURAR	GAGGGCTGTCACCCGCTCGGCGGCTTCTAATCCGT TCCTTTTTCAATGGGTAATAACTGA
OGAL1F	TCAATTTAATTATATCAGTTATTACCCATTGAAAAAGGA ACGGATTAGAAGCCGCCGAGC
OGAL1R	GATCGATCATGGAGGTTGACGAAGAAGAAGAAGAGGAC AT AGTAGTCCGATCCGGGGTTT
OA21F	TTTAACGTCAAGGAGAAA AAACCCCGGATCGGACTACT ATGTCCTCTTCTTCTTCTTCGT
OA21R-TRP1	CCAATCCAAAAGTTCACCTGTCCCACCTGCT TTACCATACATCTCTAAGATATCTTCCAC
TRPDF	TGCAAACGAGTGGAAGATATCTTAGAGATGTATGGTAA AGCAGGTGGGACAGGTGAACTT
P-TRPDR	AATCCATCGATACTAGTGCGGCCGCCCTTTAGTGAGGG TTGACGACTTGAGGCTGATGGT
SQE1-F	ATTGCAGGAGAACAAACGTG
SQE1-R	CGCCCTCGACTTATAGTCATA
CPRF	GCTACAACATCTACAACGCC
CPRR	TTTGGTGAAAGTGCTACATA
CYPC-F	TTA GGATCC ATGGTGTTGTTTTTCTCCCTAT
CYPC-R	ATT CTCGAG TTAATTGTGGGGATGTAGATGAA
ESC-PgDDSF	AAA GCGGCCGC ATGTGGAAGCTGAAGGTTGCTCAAGGA
ESC-PgDDSR	AAA GAGCTC TTAAATTTTGAGCTGCTGGTGCTTAGG
gaGTF	TCCTCCGTGCGTCCTCGT
TD2T-GTR	CATTAAAGTAACTTAAGGAGTTAAATTTAGCGATCGC TTACATAATTTCC
GT-TDH2tF	GCGAAGCTATTTGAGGAAATTATGTAAGCGATCGC TAAATTTAACTCCTTAAGTTACTT
FB1p-TDH2tR	AGGCTGGTATTGTTGTTCAAGCCAGCGGTGCCAGTTGGA GCGAAAAGCCAATTAGTGTG
TD2t-FBA1pF	TTTCGATAAAGCACTTAGTATCACACTAATTGGCTTTTCG C TCCAACTGGCACCGCTGG
CYP-FBA1pR	AAGAAGAAGAAGAGATAGGGAGAAAAACAACACCAT TTTTGAATATGTATTACTTGGTT
p-TEF2pF	GGTAGTCTAGTACCTCCTGTGATATTATCCCATTCCATGC GGGGCCGTATACTTACATAT

A21-TEF2pR	TCGATCATGGAGGTTGACGAAGAAGAAGAAGAGGACAT GTTTAGTTAATTATAGTTCGTT
TEF2p-A21F	TAGAATATACGGTCAACGAACTATAATTAACTAAACATG TCCTCTTCTTCTTCTTCGTCA
TPI1t-A21R	AAGAAGATAATATTTTTATATAATTATATTAATCTTACCA TACATCTCTAAGATATCTTC
A21-TPI1tF	GAGTGGAAGATATCTTAGAGATGTATGGTAAGATTAATA TAATTATATAAAAATATTATC
p-TPItR	CACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGC CTATATAACAGTTGAAATTTGG
TPI1p-GPM1pF	ACGGGTAATCTTCCACCAACCTGATGGGTTCCTAGATAT A TAGTCGTGCAATGTATGACT
CYP-GPM1pR	GGAAGAAGAAGAAGAGATAGGGAGAAAAACAACACCAT TTATTGTAATATGTGTGTTTGT
GPM1p-TPI1pF	TTCCTGCTCACAAATCTTAAAGTCATACATTGCACGACTA TATATCTAGGAACCCATCAG
SQE-TPI1pR	TGCAACGTATCAGTAGTACTAGAAGAAGATGAATTCAT TTTTTAGTTTATGTATGTGTTT
TPI1p-SQE1F	TAAATCTATAACTACAAAAAACACATACATAAACTAAAA A ATGAATTCATCTTCTTCTAG
ENO2t-SQE1R	AAGCAGAAAAGACTAATAATTCTTAGTTAAAAGCACTTT A TTAGTGAATGGGGGGAGCTC
SQE-ENO2tF	CTGCTTACTACAGAGCTCCCCCCATTCACTAA TAAAGTGCTTTTAACTAAGAATTATTAG
TEF1p-ENO2tR	CTGGAAGAGTAAAAAAGGAGTAGAAACATTTTGAAGCT AT AGGTATCATCTCCATCTCCC
ENO2t-TEF1pF	GACCACAGTGATATGCATATGGGAGATGGAGATGATACC T ATAGCTTCAAAATGTTTCTA
DDS-TEF1pR	TATGGATCATTTCCTTGAGCAACCTTCAGCTTCCACAT TTTGTAATTAAAACTTAGATTA
CYP-TDH2tF	GCGCATGGACTTCCAATTCATCTACATCCCCACAAT TAAATTTAACTCCTTAAGTTACTT
UGTPg1-FBA1pR	GCGGGCAAGAATATCAATTCTGACTTCATGCGATCGC TTTTGAATATGTATTACTT
FB1p-UGTPg1F	AACCATAACCAAGTAATACATATTCAAAAGCGATCGC ATGAAGTCAGAA
CYC1t-UGTPg1R	TCGGTTAGAGCGGATCTTAGCTAGCCGCGGTACCAAGC TTACATAATTTCCTC

Supplementary Information, Table S7. Plasmids used in this study.

Plasmids	Genotype or characteristic	Source
pBlu2KSp		Invitrogen
pYES2.1	2MICRON, URA3	Invitrogen
pESC-HIS	2MICRON, HIS3	Invitrogen
pLLeu-tHMGR-UPC2.1	*ADH1p-tHMGR-ADH1t*, *TEF1p-UPC2.1-ADH1t*, CEN6/ARSH4, LEU2	Dai *et al*. (2012) [7]
pKS-ATR21	ATR2-1	This work
pKS-SQE1	SQE1	This work
pMD18T-UGTPgx	Cloned UGTPgx (x represents 1-16)	This work
pYES2.1-UGTPgx	C-terminally 6xHis-tagged UGTPgx (x represents 1-16)	This work
pESC-HIS-CYP	*GAL1p-CYP-CYC1t*, 2MICRON, HIS3	This work
pHCD	*GAL1p-CYP-CYC1t*, *GAL10p-DDS-ADH1t*, 2MICRON, HIS3	This work
pHCD-UGTPg1	*FBA1p-CYP-CYC1t*, *GAL10p-DDS-ADH1t*, *GAL1p-UGTPg1-TDH2t*, 2MICRON, HIS3	This work
pHCDA	*GAL1p-CYP-CYC1t*, *GAL10p-DDS-ADH1t*, *TEF2p-ATR2-1-TPI1t*, 2MICRON, HIS3	This work
pHSCDA	*TPI1p-SQE1-ENO2t*, *GPM1p-CYP-CYC1t*, *TEF1p-DDS-ADH1t*, *TEF2p-ATR2-1-TPI1t*, 2MICRON, HIS3	This work
pHCDR1	*GAL1p-CYP-CYC1t*, *GAL10p-DDS-ADH1t*, *TEF2p-PgCPR1-TPI1t*, 2MICRON, HIS3	This work
pHCDR2	*GAL1p-CYP-CYC1t*, *GAL10p-DDS-ADH1t*, *TEF2p-PgCPR2-TPI1t*, 2MICRON, HIS3	This work
pHSCDAG	*TPI1p-SQE1-ENO2t*, *GPM1p-CYP-TDH2t*, *TEF1p-DDS-ADH1t*, *TEF2p-ATR2-1-TPI1tFBA1p-UGTPg1-CYC1t*, 2MICRON, HIS3	This work
pET28a-UGTPg1	harboring C-terminally 6xHis-tagged UGTPg1	This work

Supplementary Information, Table S8. Strains used in this study.

Strains	Genotype or characteristic	Source
BY4742	*MATα*, *his3*Δ1, *leu2*Δ0, *lys2*Δ0, *ura3*Δ0	Brachmann CB *et al.* (1998) [8]
BA21	*MATα, his3Δ1, leu2Δ0, lys2Δ0, ura3Δ0, trp1Δ*::*URA3_GAL1p-ATR2-1*	This work
AC1	BA21 haroring pESC-HIS-CYP	This work
AP1	BA21 harboring pHCD	This work
AK1	BA21 harboring pHCD-UGTPg1	This work
AKE	BA21 harboring pHCD-UGTPg1 and pLLeu-tHMGR-UPC2.1	This work
BK1	BY4742 harboring pHSCDAG	This work
BKE	BY4742 harboring pHSCDAG and pLLeu-tHMGR-UPC2.1	This work
BD1	BY4742 harboring pHCDA	This work
BD2	BY4742 harboring pHCDR1	This work
BD3	BY4742 harboring pHCDR2	This work

References

1 Huang X, Madan A. CAP3: A DNA sequence assembly program. *Genome Res* 1999; **9**:868-877.

2 Grabherr MG, Haas BJ, Yassour M *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011; **29**:644-652.

3 Vogt T, Jones P. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci* 2000; **5**:380-386.

4 Lin W, Peng Y, Li G *et al.* Isolation and functional characterization of PgTIP1, a hormone-autotrophic cells-specific tonoplast aquaporin in ginseng. *J Exp Bot* 2007; **58**:947-956.

5 Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D. Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *J Biol Chem* 1997; **272**:19176-19186.

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

7 Dai Z, Liu Y, Huang L, Zhang X. Production of miltiradiene by metabolically engineered *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 2012; **109**:2845-2853.

8 Brachmann CB, Davies A, Cost GJ et al. Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 1998; 14:115-132.