Metabolic Engineering 29 (2015) 97–105

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Production of bioactive ginsenosides Rh2 and Rg3 by metabolically engineered yeasts

Pingping Wang [a,1], Yongjun Wei [a,1], Yun Fan [a,c,1], Qunfang Liu [b], Wei Wei [a], Chengshuai Yang [a], Lei Zhang [a], Guoping Zhao [a,c], Jianmin Yue [b,*], Xing Yan [a,**], Zhihua Zhou [a,**]

[a] CAS-Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, China
[b] State Key Laboratory of Drug Research, Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, 201203, China
[c] State Key Laboratory of Genetic Engineering, Department of Microbiology, School of Life Sciences and Institute of Biomedical Sciences, Fudan University, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 24 September 2014
Accepted 2 March 2015
Available online 11 March 2015

Keywords:
Ginsenoside Rh2
Ginsenoside Rg3
UDP-glycosyltransferase
Protopanaxadiol producing chassis
Panax plants

ABSTRACT

Ginsenosides Rh2 and Rg3 represent promising candidates for cancer prevention and therapy and have low toxicity. However, the concentrations of Rh2 and Rg3 are extremely low in the bioactive constituents (triterpene saponins) of ginseng. Despite the available heterologous biosynthesis of their aglycone (protopanaxadiol, PPD) in yeast, production of Rh2 and Rg3 by a synthetic biology approach was hindered by the absence of bioparts to glucosylate the C3 hydroxyl of PPD. In this study, two UDP-glycosyltransferases (UGTs) were cloned and identified from *Panax ginseng*. UGTPg45 selectively transfers a glucose moiety to the C3 hydroxyl of PPD and its ginsenosides. UGTPg29 selectively transfers a glucose moiety to the C3 glucose of Rh2 to form a 1–2-glycosidic bond. Based on the two UGTs and a yeast chassis to produce PPD, yeast cell factories were built to produce Rh2 and/or Rg3 from glucose. The turnover number (k_{cat}) of UGTPg29 was more than 2500-fold that of UGTPg45, which might explain the higher Rg3 yield than that of Rh2 in the yeast cell factories. Building yeast cell factories to produce Rh2 or Rg3 from simple sugars by microbial fermentation provides an alternative approach to replace the traditional method of extracting ginsenosides from *Panax* plants.

1. Introduction

It has been reported widely that ginsenosides Rh2 and Rg3 induce tumor cell apoptosis (Min et al., 2006; Park et al., 1997), inhibit tumor cell proliferation (Kim et al., 2004) and restrain tumor invasion and metastasis (Mochizuki et al., 1995; Shinkai et al., 1996). Moreover, the use of Rh2 or Rg3 in combination with anti-tumor drugs, increases the efficacy of tumor suppression and potently reduces the side effects of therapy in tumor-bearing mice (Liu et al., 2009; Nakata et al., 1998). Rh2 and Rg3 are also potentially effective candidates for preventing metabolic disorders, such as diabetes and obesity, *via* activating the AMPK signaling pathway (Hwang et al., 2007; Park et al., 2008). Rg3 is now used as the major active component in an anti-tumor and anti-cancer drug in China (Z20030043). BST204, a purified ginseng dry extract containing high concentrations of a Rh2 and Rg3 mixture, is being developed as a supportive-care agent for cancer patients in Korea (Bae et al., 2014).

Despite the versatile bioactivities of Rh2 and Rg3, their contents in ginseng are extremely low. Rh2 is almost undetectable in the total ginsenosides of *Panax* plants (*Panax ginseng*, *Panax quinguefolium*, and *Panax notoginseng*), and Rg3 is only detected in the total ginsenosides of *P. ginseng*; its content in the dry ginseng is about 0.0003% (Shibata, 2001). The contents of Rh2 and Rg3 can be increased markedly to 0.001% and 0.015%, respectively, in red ginseng, which is prepared by treatment with steaming and air-drying (sometimes, the process is repeated) (Shibata, 2001). The hot water reflux extraction of total ginsenosides from *P. quinguefolium* also generates relatively higher contents of Rh2 and Rg3 (Popovich and Kitts, 2004). Rh2 and Rg3 are both protopanaxadiol (PPD)-type ginsenosides, bearing one and two glucose moieties at the C3 hydroxy of PPD, respectively. The red ginseng treatment and hot water treatment are believed to result in the partial release of the glycosyl moieties that are originally attached

* Corresponding author. Fax: +86 21 50806718.
** Corresponding authors. Fax: +86 21 54924049.
E-mail addresses: jmyue@mail.shcnc.ac.cn (J. Yue), yanxing@sibs.ac.cn (X. Yan), zhouzhihua@sippe.ac.cn (Z. Zhou).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ymben.2015.03.003
1096-7176/

at the hydroxyls at C20 and C3 of the major PPD-type ginsenosides, such as Rb1, Rb2, Rc and Rd, and thus yield ginsenosides Rh2 and Rg3. However, the contents of Rh2 and Rg3 in red ginseng are still considerably lower than those of the major ginsenosides (Shibata, 2001).

Thus, the preparation of Rh2 and Rg3 for commercial use relies on chemical or biological deglycosylation of the major protopanaxadiol-type ginsenosides. The major PPD-type ginsenosides can be converted to Rh2 and Rg3 under mild acidic conditions (Han et al., 1982). In biological approaches, microorganisms or glycoside hydrolases were used to hydrolyze the glycosides of the major ginsenosides to produce Rh2 and Rg3 (Bae et al., 2004; Cheng et al., 2008; Su et al., 2006; Yan et al., 2008). For example, BST204, a ginsenoside extract containing a high concentration of Rh2 and Rg3 mixtures, was prepared from crude ginseng with a ginsenoside-β-glucosidase combined with acid hydrolysis to enrich Rh2 and Rg3 (Bae et al., 2014). Both of the chemical and biological deglycosylation approaches require a supply of PPD-type ginsenosides, which depends on plantations of *Panax* plants. However, the repeated cultivation of *Panax* plants makes soil unsustainable, and can result in decreasing productivity and diseases. Indeed, the fields for *Panax* plants cultivation generally require a period of more than 10 years for rotational tillage before the next replantation (Huang and Madan, 1999). Thus, problems associated with the cultivation of *Panax* plants are becoming limiting in the preparation of Rh2 and Rg3 at a large scale. Although Rh2 and Rg3 capsules have been available commercially on the Chinese market, their price is too high for most customers.

Previously, we assembled 158 UGT contigs from *Panax* EST datasets, which were clustered into 42 operational taxonomic units (OTUs) (Yan et al., 2014). Some of them were cloned from *P. ginseng* and expressed in *Escherichia coli* and one of them was characterized as a novel UDP-glycosyltransferase (UGTPg1) that could selectively glucosylate the C20 hydroxyl of dammarane-type triterpenoids. Through synthetic biology, a yeast microbial cell factory to produce compound K (CK) was constructed by integrating the known pathway to produce PPD in combination with this novel glycosylation step. In this study, more *Panax* EST datasets were used to assemble UGTs with full-length coding DNA sequences (CDS). Also more UGTs were cloned from *P. ginseng* and expressed. We sought to identify UGTs that could selectively glucosylate the C3 hydroxyl of PPD and further elongate the sugar chain. In this study, we not only identified two UGTs able to sequentially transfer two glucose moieties to the C3 hydroxyl of dammarane-type triterpenoids but also provided a technique for *de novo* synthesis of Rh2 and Rg3 from glucose using a one-pot approach.

2. Materials and methods

2.1. Prediction of UDP-glycosyltransferases

A *Panax* cDNA database was constructed in our previous study (Yan et al., 2014). To expand the *Panax* cDNA database here a further four *Panax* EST datasets were collected from the NCBI database (Table S1) and used for cDNA assembly by the method as described previously (Yan et al., 2014). The BESTORF software (http://linux1.softberry.com/berry.phtml?topic=bestorf&group=programs&subgroup=gfind) was used for prediction of ORFs in cDNA sequences in the database. ORFs encoding UDP-glycosyltransferases were identified by the conserved PSPG (Plant Secondary Product Glycosyltransferase) box (Vogt and Jones, 2000). Protein sequences of the selected ORFs were aligned with MAFFT (http://mafft.cbrc.jp/alignment/software/), and were then classified into OTUs using the DOTUR software (95% cutoff).

2.2. Cloning of UDP-glycosyltransferase encoding cDNAs

The *P. ginseng* cDNA was prepared from *P. ginseng* callus and *P. ginseng* leaves using the PrimeScript RT reagent Kit with gDNA eraser (Takara, Dalian, China). The cDNAs potentially encoding UDP-glycosyltransferase with long ORFs (likely to be full-length) were PCR amplified using the *P. ginseng* cDNA as the template and cloned into the pMD18T vector (Takara, Dalian, China). These cloned genes were sequenced; the corresponding GenBank accession numbers range from KM401909 to KM401937.

2.3. Phylogenetic analysis of P. ginseng UGT genes

Multiple alignments of amino acid sequences of the *Panax* UGTs were carried out using MAFFT. The alignment result output was used to create a phylogenetic tree by the MEGA 5.2.2 software using the Neighbor-Joining method, with Poisson correction. The bootstrap confidence values were obtained based on 1000 replicates and the pairwise-deletion option was used to address gaps in the alignment of the amino acid sequences.

2.4. Heterologous expression of the UGT genes in E. coli BL21 (DE3)

Heterologous expression of the UGT genes in *E. coli* BL21 (DE3) was carried out as described previously (Yan et al., 2014). Briefly, UGT genes with a C-terminally 6 × His-tag was ligated into the pET28a vector with the GBclonart kit (Genebank Biosciences Inc., Suzhou, China) and expressed in *E. coli* BL21 (DE3) under IPTG induction. The cells were collected by centrifugation and suspended in 100 mM phosphate buffer (pH 8.0) supplemented with 1 mM PMSF and then disrupted with a French Press (25 kpsi). The cell debris was removed by centrifugation (12,000*g*, 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.

Quantification of the UGTPg29 and UGTPg45 in the crude cell extract was carried out as described previously (Yan et al., 2014). Serially diluted samples (0, 2, 4, 6, 8, 10 ng/μL) of a C-terminal 6 × His-tagged xylanase (53 kDa) were used to generate a standard curve for quantitative reference.

2.5. Enzymatic assays for UGTPgs

Enzymatic assays for glycosyltransferases were performed as described previously (Yan et al., 2014). Generally, the reaction was carried out in a 100 μL volume containing 100 mM phosphate buffer (pH 8.0), 1% Tween-20, 5 mM UDP-glucose, 0.5 mM acceptor substrate and recombinant *E. coli* extract for 2 h in a 35 °C water bath and was terminated by adding 100 μL *n*-butanol. The product was extracted in the organic phase which was then evaporated. The residue was dissolved in methanol for TLC and HPLC analysis.

For the kinetic study of UGTPg45, the reaction mixture contained 100 mM phosphate buffer (pH 7.5), 1% Tween-20, 5 mM UDP-glucose, acceptor substrate (40–500 μM PPD) and 400 ng UGTPg45 in a final volume of 100 μL. The reaction was incubated at 40 °C for 20 min. For UGTPg29, the reaction mixture contained 100 mM phosphate buffer (pH 8.0), 1% Tween-20, 5 mM UDP-glucose, acceptor substrate (40–500 μM Rh2) and 100 ng UGTPg29 in a final volume of 100 μL. The reaction was incubated at 40 °C for 15 min. HPLC analysis was used to quantify the target product in each reaction. The Michaelis–Menten parameters were calculated by least-squares fitting of the kinetic model using Prism 5 (GraphPad, San Diego, CA, USA). All data are presented as means ± SD of three independent repeats.

Table 1
Strains and plasmids used in this study.

	Genotype or characteristic	Source
Strains		
BY4742	*MATα*, *his3*Δ1, *leu2*Δ0, *lys2*Δ0, *ura3*Δ0	Brachmann et al. (1998)
ZW-DM-A	*TEF1p-PgDDS-CYC1t*, *PGK1p-tHMG1-ADH1t*, and *LEU2* marker gene integrated into rDNA site of BY4742	This work
ZW-PPD-A	*TEF1p-synPgPPDS-FBA1t*, *TDH3p-ATR2.1-ENO2t*, *PGK1p-tHMG1-ADH1*t, and *URA3* marker gene integrated into *δ DNA* site of ZW-DM-A	This work
ZW-PPD-B	*GPM1p-ERG20-CYC1t*, *TPI1p-PgSQE1-ENO2t*, *PGK1p-ERG9-FBA1*t, and *HIS3* marker gene integrated into *YPRCτ3* site of ZW-PPD-A	This work
D20RH18	*UGTPg45* gene integrated into *δ DNA* site of ZW-PPD-B	This work
D20RG1	*UGTPg45* and *UGTPg29* genes integrated into *δ DNA* site of ZW-PPD-B	This work
Plasmids		
pYES2	2MICRON, *URA3*	Invitrogen
pAG32	*HPH*	EUROSCARF
pKS-ATR21	*ATR2.1*	Yan et al. (2014)
pHSCDAG	*TPI1p-SQE1-ENO2t*, *GPM1p-CYP-TDH2t*, *TEF1p-DDS-ADH1t*, *TEF2p-ATR2-1-TPI1t*, *FBA1p-UGTPg1-CYC1t*, 2MICRON, *HIS3*	Yan et al. (2014)
p2M	*TEF1*p-*CYC1*t, *HPH*	This work
p2M-GT45	*TEF1p-UGTPg45-CYC1t*, *HPH*	This work
p2M-GG	*TEF1p-UGTPg45-TPI1t-HXT7p-UGTPg29-CYC1t*, *HPH*	This work

2.6. Construction of vectors and yeast strains

All the *Saccharomyces cerevisiae* strains and plasmids used or constructed in this study are listed in Table 1, and the primers used for the construction of the yeast strains are listed in Table S2. Genome integration of gene cassettes was performed using DNA assembler methods *via* yeast homologous recombination (Shao et al., 2009). The detailed procedures to construct the strains are provided in the Supplementary Data.

2.7. Yeast cultivation and metabolite extraction

To determine the performance of the genetically engineered yeast strains, individual clones were inoculated into the YPD medium and cultivated at 30 °C, 250 rpm for 24 h. Aliquots were diluted to an initial OD_{600} of 0.05 in 10 mL of YPD medium in 50 mL flasks and grown at 30 °C, 250 rpm for 6 days. Cell biomass was determined by spectrophotometer under 600 nm and convert to dry cell weight (1 OD_{600} = 0.513 g DCW L^{-1}) (Dai et al., 2013). The cell pellets were harvested, disrupted and extracted with *n*-butanol and used for the analysis of intracellular dammarane-type triterpenoids (dammarenediol II, PPD, Rh2 and Rg3). The supernatant was extracted directly with *n*-butanol and used for the analysis of extracellular dammarane-type triterpenoids. Unless stated otherwise, the yield of dammarane-type triterpenoids of yeasts referred to the sum of intracellular and extracellular contents. The cell pellets were also lysed with 20% (w/v) KOH and 50% (v/v) ethanol by boiling for 30 min, and extracted with hexane and used for the analysis of squalene, 2,3-oxidosqualene, lanosterol and ergosterol by HPLC.

2.8. TLC analysis

TLC analysis was performed using silica gel 60 F254 plates (Merck KGaA, Darmstadt, Germany) as described previously (Yan et al., 2014). For identification, a mixture of PPD-type ginsenoside authentic samples (CK, Rh2, F2, Rg3, Rd, Rb1, Rb2 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. (Nanjing, China).

2.9. HPLC analysis

A Shimadzu LC-20 A prominence system (Shimadzu, Kyoto, Japan) was used for the HPLC analysis. Chromatographic separations were carried out at 35 °C on a Shodex C18-120-5 4E column (5 μm, 4.6 mm × 250 mm) as described previously (Yan et al., 2014). Authentic dammarenediol II (DM) was purchased from BioBioPha Co., Ltd. (Kunming, China). Authentic squalene, 2,3-oxidosqualene, lanosterol and ergosterol were purchased from Sigma-Aldrich.

Semi-preparative HPLC was carried out on a Waters 1525 binary pump system with a Waters 2489 detector using a YMC-Pack ODS-A column (5 μm, 10 mm × 250 mm), and the mobile phases were 70% and 80% acetonitrile in water for Rg3 and Rh2, respectively. The products were monitored at 203 nm.

2.10. HPLC/ESIMS analysis

The chromatographic analysis was performed using an Agilent 1200 Series LC system (Agilent Technologies, Waldbronn, Germany) equipped with a Shodex C18-120-5 4E column. ESIMS data were acquired using a Bruker Daltonics Esquire 3000plus (Bruker Daltonics, MA, USA) in the positive ionization mode.

2.11. NMR analysis

NMR experiments were performed in pyridine-d_5 for Rg3 and CD_3OD for Rh2 on a Bruker Avance III 400 (for ^1H NMR) or a Bruker Avance III 500 (for ^{13}C NMR) (Bruker, Billerica, MA, USA) with reference to the solvent peaks.

3. Results and discussion

3.1. Phylogenetic analysis of cloned *P. ginseng* UGT genes

A *Panax* cDNA database was constructed in our previous study (Yan et al., 2014). In this study, the database was expanded by adding more *P. ginseng* datasets for cDNA assembly (Table S1). Contigs encoding UGTs were predicted by the consensus sequence of PSPG box. ORFs > 1320 bp in length from these contigs with the PSPG motif were used for further analyses. All of these UGT contigs were clustered into 59 OTUs, based on protein sequences at a 95% similarity cutoff including UGTPg1-16 which were cloned, expressed and characterized partially in our previous study (Yan et al., 2014). The UGT genes of the remaining 43 OTUs were PCR amplified and cloned from cDNA prepared from a *P. ginseng* callus or leaves. However, only 29 UGTs (designated as UGTPg17–45) were successfully cloned. The cDNA of the other 14 OTUs (designated as UGT_Panax46–59) could not be cloned from *P. ginseng*, which may have been due to the transcript of these genes having very low abundances in the *P. ginseng* callus and

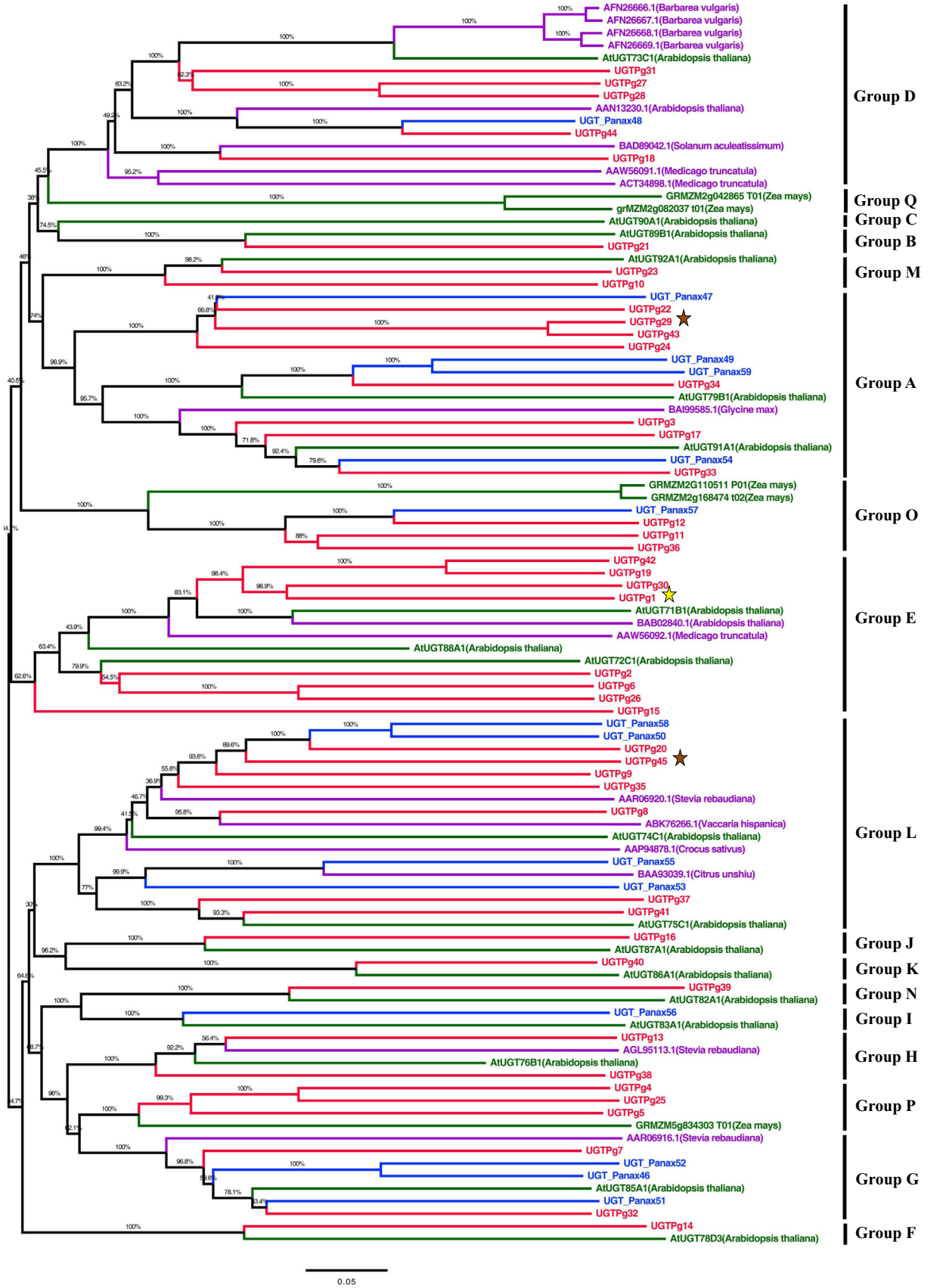

Fig. 1. Phylogenetic analysis of *Panax* UGTs with an unrooted tree. A neighbor-joining tree was constructed using the MEGA 5.2.2 software (bootstrap values: 1000) with peptide sequences of 59 representative UGTs from *Panax* plants (45 UGTs cloned from *P. ginseng*, UGTPg1-45 and 14 UGTs assembled from a *Panax* cDNA database, UGT_Panax46-59). In total, 18 UGTs of *Arabidopsis thaliana*, 5 UGTs of *Zea mays*, and 17 characterized terpene UGTs were used as reference sequences. The UGTs cloned from *P. ginseng* are labeled in red. The UGTs assembled from the *Panax* cDNA database are labeled in blue. The UGTs of *A. thaliana* and *Z. mays* used to classify UGTs in groups are labeled in green. The characterized terpene UGT sequences are labeled in purple. UGTPg1, previously identified as a triterpene UGT, is marked by a yellow asterisk. UGTPg29 and UGTPg45, which were also identified as triterpene UGTs in this study, are marked by brown asterisks. The bootstrap values are marked above the nodes and the scale bar is indicated under the tree. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

leaves used as the template or that these genes were specifically present in other *Panax* plants, such as *P. notoginseng* and *P. quinquefolium*, since their transcriptome datasets were also used to assemble the UGT contigs (Table S1). These 14 DNA sequences are accessible at http://www.genoportal.org/bbdb/ under the accession numbers of SBB_00486 to SBB_00499.

By phylogenetic analysis, 59 OTUs assembled from *Panax* plants, which might represent 59 UGTs, were clustered into 15 groups defined by the UGTs from *Arabidopsis thaliana* and *Zea mays* (Li et al., 2014). The results demonstrated a high diversity of UGTs in *Panax* plants as in other plants. Some UGTs of *Panax* plant, such as UGTPg20, UGTPg45, UGTPg9, UGTPg35, UGTPg8, UGT_Panax55 and UGT_Panax53, were phylogenetically close to characterized UGTs with known activities to terpenoids in the phylogenetic tree, suggesting they may be involved in glycosylation of ginsenosides (Fig. 1).

3.2. Characterizations of UGTPg45 and UGTPg29

The 29 *P. ginseng* UGTs newly cloned in this study were expressed in *E. coli* BL21 (DE3) for characterization of their activities. Only two UGTs, UGTPg45 and UGTPg29, were found to be active in the glycosylation of PPD and PPD-type ginsenosides (Figs. S1–3). UGTPg45 belongs to group L and show 41% similarity to a UGT of *Stevia rebaudiana* (SrUGT74G1, GenBank accession no. AAR06920.1). UGTPg29 belongs to group A. And the similarities of both were low to UGTPg1 (GenBank accession no. AIE12479.1), the first characterized ginsenosides glycosyltransferase (Yan et al., 2014), which belongs to group E.

Crude cell extract of the recombinant *E. coli* expressing UGTPg45 ORF can convert 20(*S*)–PPD into 20(*S*)–Rh2 as monitored by TLC (Fig. S2), HPLC (Fig. 2 A) and HPLC/ESIMS analysis (Fig. 2B). After purification by semi-preparative HPLC, the structure of the product was further confirmed to be that of 20(*S*)–Rh2 by ^{1}H and ^{13}C NMR (Table S4). Thus, UGTPg45 was demonstrated to have a novel function to glucosylate the C3 hydroxyl of 20(*S*)–PPD. Similarly, incubation of UGTPg45 with UDP-glucose and CK, a PPD-type ginsenoside with a free C3 hydroxyl, yielded ginsenoside F2 *in vitro* (Fig. S2). However, UGTPg45 could not catalyze the glucosylation of other PPD-type ginsenosides, such as Rh2 and Rg3, which already harbor one or more glucose moieties at the C3 hydroxyl. These results indicate that UGTPg45 selectively transfers a glucose moiety to the free C3 hydroxyl of PPD and PPD-type ginsenosides *in vitro*.

The substrate specificity of UGTPg29 was analyzed *in vitro* using the crude cell extract of the recombinant *E. coli* expressing the

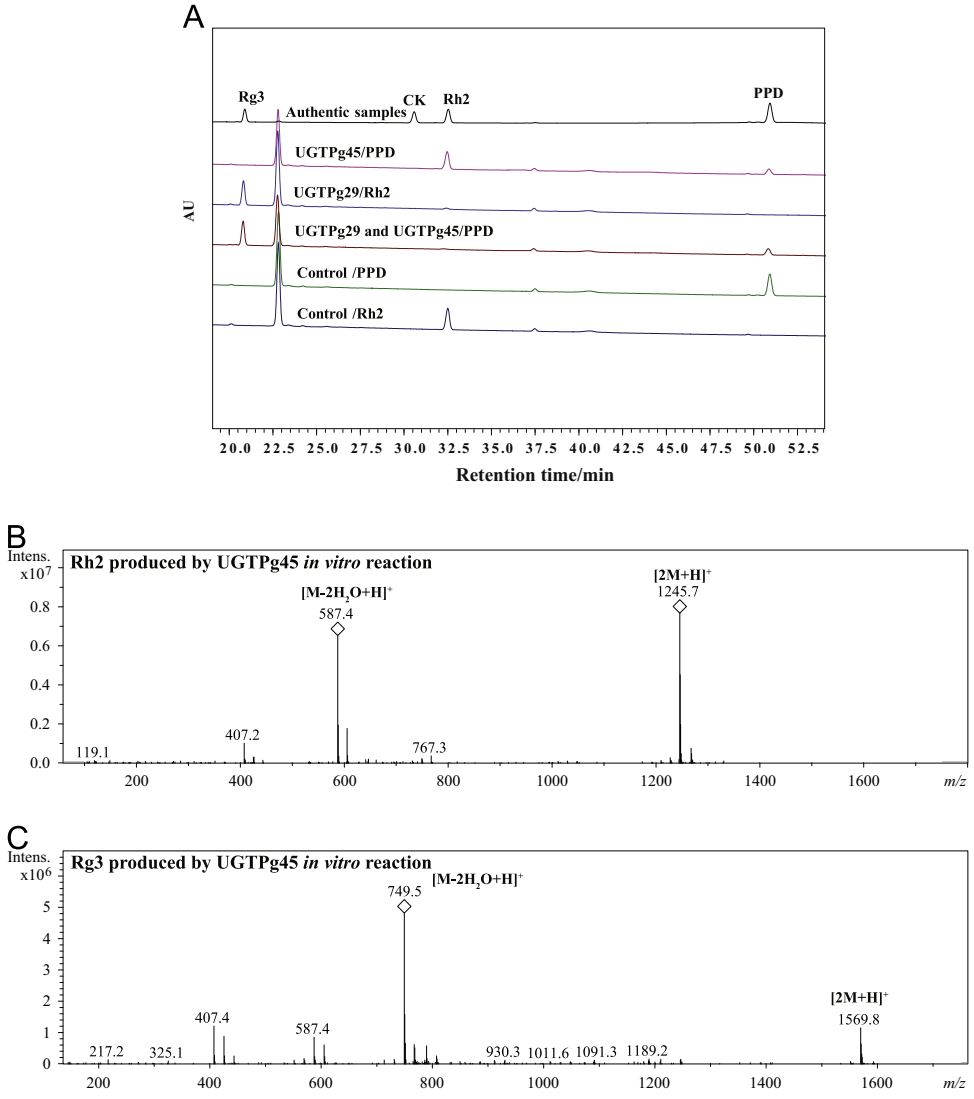

Fig. 2. Functional characterization of UGTPg45 and UGTPg29 *in vitro*. (A) HPLC analysis of *in vitro* products from the incubation of UGTPg45 and (or) UGTPg29 with PPD or Rh2 as substrates. (B) MS spectrum of Rh2 produced from the *in vitro* reaction of PPD catalyzed by UGTPg45. (C) MS spectrum of Rg3 produced from the *in vitro* reaction of Rh2 catalyzed by UGTPg29.

UGTPg29 ORF. The expressed UGTPg29 could readily convert 20(*S*)–Rh2 into 20(*S*)–Rg3, as shown by TLC (Fig. S3A), HPLC (Fig. 2A) and HPLC/ESIMS analysis (Fig. 2C), and the structure of the product was further confirmed to be that of 20(*S*)–Rg3 by ^{1}H and ^{13}C NMR following purification by semi-preparative HPLC (Table S4). Similarly, incubation of UGTPg29 with ginsenoside F2 and UDP–glucose yielded ginsenoside Rd (Fig. S3A). Moreover, incubating equivalent amounts UGTPg45 and UGTPg29 proteins with PPD and UDP–glucose directly yield Rg3 (Fig. S3A). By contrast, when UGTPg29 and UDP–glucose were incubated with PPD, CK, Rg3 and Rd, no new compound was detected (Fig. S3B). These results demonstrated that UGTPg29 selectively transfer a glucose moiety to the C3 glucose of Rh2 to form a 1–2-glycosidic bond. To our knowledge, this is the first reported UGT identified to elongate the sugar chain of tetracyclic triterpenoids.

3.3. Kinetic study of UGTPg45 and UGTPg29

The V_{max}, K_m, k_{cat} and k_{cat}/K_m for UGTPg45 to catalyze the glucosylation of the C3 hydroxyl of PPD to produce Rh2 were 80 nmol/min/mg, 209 μM, 0.4 s^{-1} and 0.002 μM^{-1} s^{-1}, respectively (Table 2). These values for UGTPg29 towards Rh2 to produce Rg3 were 24,624 nmol/min/mg, 149 μM, 1111.5 s^{-1} and 7.460 μM^{-1} s^{-1}, respectively (Table 2). The substrate affinity of UGTPg29 was slightly higher than that of UGTPg45. However, the turnover number (k_{cat}) of UGTPg29 was more than 2500-fold that of UGTPg45, indicating that the catalytic efficiency of UGTPg29 was much higher than that of UGTPg45. This is consistent with the results of *in vitro* reactions, in which few Rh2 accumulated when an equivalent UGTPg45 and UGTPg29 were used to catalyze PPD (Fig. S3). Based on these results, we suggested that the single glucose moiety at C3 hydroxyl of PPD-type ginsenosides, such as Rh2 and F2, could be rapidly elongated to form a Glc(2–1)Glc glycan by UGTPg29 and produce Rg3 and Rd, respectively. The rare ginsenosides Rh2, F2 and Rg3 could be intermediates for the major ginsenosides Rb1, Rb2, Rc and Rd in *P. ginseng*, all of which carry a Glc(2–1)Glc glycan at their C3 hydroxyl group. Moreover, the k_{cat} of UGTPg45 was also only one twenty-fifth of that of UGTPg1 which catalyzes the glucosylation of the C20 hydroxyl of PPD-type ginsenosides with high specificity (Yan et al., 2014), suggesting that the glucosylation of C20 hydroxyl was possibly prior to that of C3 hydroxyl. Although Rh2 and Rg3 could also possibility be generated through sequential glucosylation of the C3 hydroxyl of PPD by UGTPg45 and UGTPg29, respectively; they would be rapidly glycosylated by UGTPg1 to produce ginsenoside F2 and Rd as illustrated in our previous study (Yan et al., 2014). Thus, the contents of Rh2 and Rg3, which bear a free C20 hydroxyl, are markedly lower than those of the major PPD-type ginsenosides in *P. ginseng*.

Table 2
The kinetic parameters of UGTPg45 and UGTPg29 towards PPD and Rh2 as substrates, respectively.

Enzyme[a]	Substrate/ product	V_{max} (nmol/ min/mg)	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m (μM^{-1} s^{-1})
UGTPg45	PPD/Rh2	80 ± 4	209 ± 14	0.384	0.0018
UGTPg29	Rh2/Rg3	24,624 ± 1429	149 ± 16	1111.5	7.46

[a] All given data in this table representing mean values from three repeats with corresponding standard deviations.

3.4. Construction of protopanaxadiol producing chassis yeasts

PPD is a precursor of dammarane-type ginsenosides and its biosynthetic pathway has been determined (Han et al., 2006, 2011; Tansakul et al., 2006). The biosynthetic pathway of PPD was integrated into the chromosome of *S. cerevisiae* BY4742 by Dai and coworkers (Dai et al., 2013). The strengthening of the precursor supplying pathway of PPD, including MVA pathway, brought relative high PPD yield (Dai et al., 2013). A yeast chassis to produce PPD was built according to Dai et al. with some modifications: we used the cytochrome P450 reductase gene of *A. thaliana ATR2.1*, instead of *ATR1* (Urban et al., 1997) and some steps of assembling the genes were modified (Fig. 3).

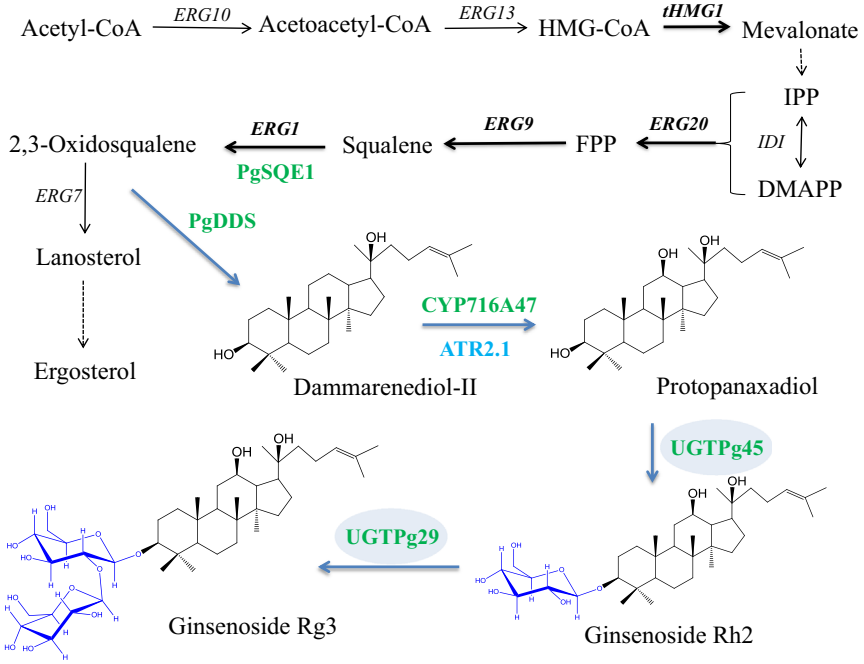

Fig. 3. Biosynthetic pathway of Rh2 and Rg3 in the genetically engineered yeasts. Black arrows indicate the intrinsic *S. cerevisiae* pathway; multiple-step reactions are shown by dashed arrows and enhanced reactions are shown by bold arrows. Blue arrows represent the *P. ginseng* pathway. Black-colored genes are from *S. cerevisiae*, green-colored genes are from *P. ginseng*, with newly characterized UGTs in this study shown in opal; the blue-colored gene is from *A. thaliana*. HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A, IPP, isopentenyl pyrophosphate, DMAPP, dimethylallyl pyrophosphate, FPP, farnesyl diphosphate. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

First, the *PgDDS* gene together with the *tHMG1* gene under the control of the constitutive promoters *TEF1* and *PGK1*, were integrated into the rDNA sites of BY4742 chromosome. The DM yield of the resulting strain ZW-DM-A was 2.60 μmol/g DCW. Second, a codon-optimized *CYP716A47* gene (*SynPgPPDS*) (Dai et al., 2013), *ATR2.1* gene and *tHMG1* gene under the control of the constitutive promoters *TEF1*, *TDH3* and *PGK1*, respectively, were integrated into the *δ DNA* sites of ZW-DM-A chromosomes, and 20 transformants were selected and verified by PCR. The PPD yields of these transformants varied greatly, suggesting different copies of genes had been integrated into the recombinant yeast chromosomes (Fig. S4 A). The highest PPD producing yeast strain, 2.49 μmol/g DCW, was designated as ZW-PPD-A (Fig. 4, Table 3). To further improve the productivity of PPD, strain ZW-PPD-B was constructed by integrating *PgSQE1*, *ERG9* and *ERG20* genes (under the control of the constitutive promoters *TPI1*, *PGK1* and *GPM1*, respectively) into the *YPRCτ3* site of strain ZW-PPD-A. Genes integrated into this site often have higher expression levels than those in other commonly used sites, such as the *URA3* locus (Flagfeldt et al., 2009). As expected, production of PPD in the resulting strain ZW-PPD-B, 8.29 μmol/g DCW, was over 3-fold of that of ZW-PPD-A (Fig. 4 and Table 3). Major part of PPD (> 80%) was retained within the cells of the strains constructed (Table S3), which was consistent with the previous report (Dai et al., 2013). All the yeast strains constructed in our study produced very little DM, much lower than that of the strains constructed in a previous study (Dai et al., 2013). For example, the ratio of PPD/DM in strain ZW-PPD-B was more than 100-fold *versus* those of the strains constructed by Dai et al. (2013). This result may be due to the utilization of the cytochrome P450 reductase gene (*ATR2.1*) in this study, which may be more efficient than *ATR1*. Accumulating less intermediate metabolite would be a benefit of further engineering of this strain.

3.5. Cell factories for the production of Rh2 and Rg3

To establish a yeast cell factory to synthesize ginsenoside Rh2, UGTPg45 (under the control of the constitutive promoter *TEF1*) was

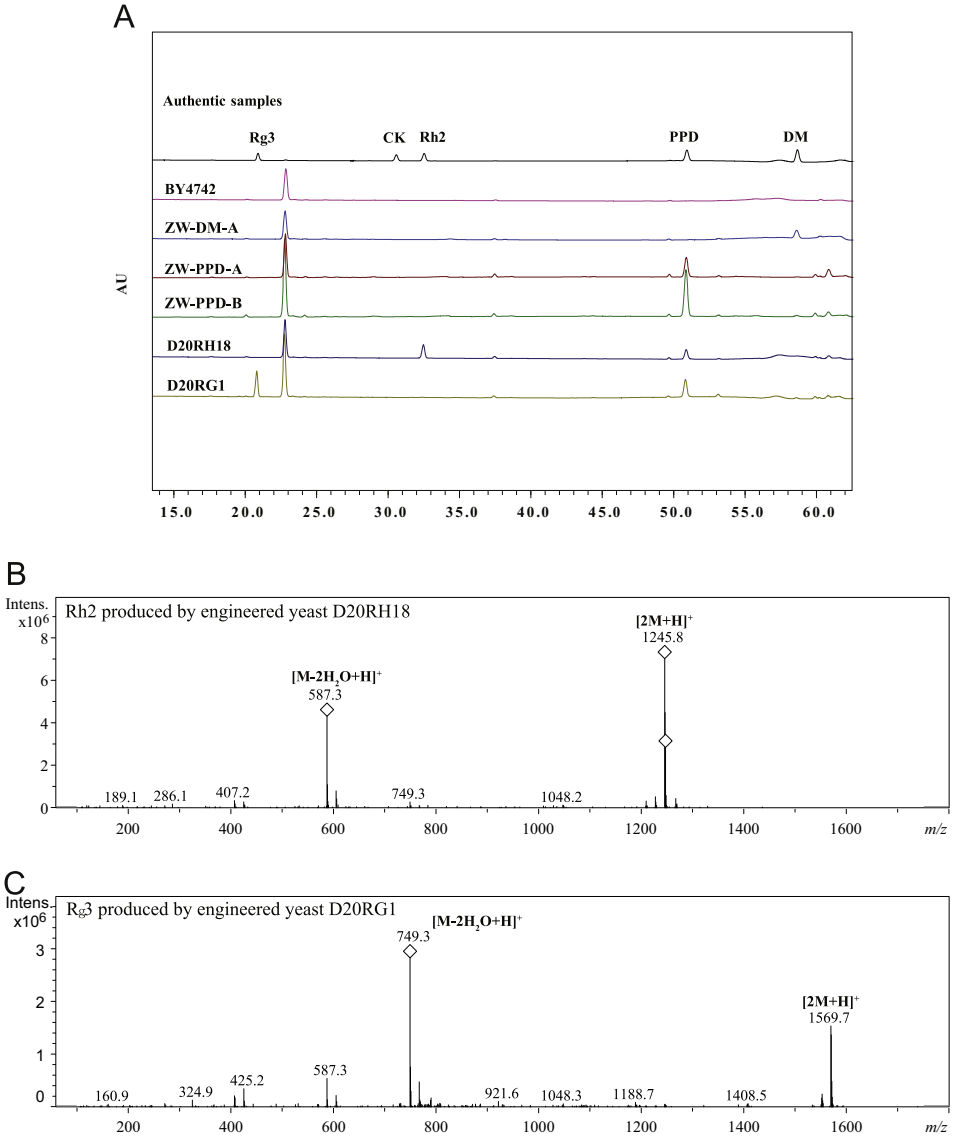

Fig. 4. Chemical analyses of DM, PPD, Rh2, and Rg3 produced from the genetically engineered yeasts: (A) HPLC analysis of DM, PPD, Rh2, and Rg3 produced from ZW-DM-A, ZW-PPD-A, ZW-PPD-B, D20RH18 and D20RG1. Black: authentic samples. Red: product of BY4742. Blue: product of ZW-DM-A. Brown: product of ZW-PPD-A. Green: product of ZW-PPD-B. Purple: product of D20RH18. Light green: product of D20RG1. (B) MS spectrum of Rh2 produced by the genetically engineered yeast D20RH18. (C) MS spectrum of Rg3 produced by the genetically engineered yeast D20RG1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Production of ginsenosides and precursor products in four genetically engineered strains.

Biomass or yield[a]	ZW-PPD-A	ZW-PPD-B	D20RH18	D20RG1
DCW (g/L)	19.85 ± 0.24	19.77 ± 0.53	18.79 ± 0.31	18.19 ± 0.08
Rg3 (μmol/g DCW)	0	0	0	3.49 ± 0.14
Rh2 (μmol/g DCW)	0	0	1.45 ± 0.27	0.07 ± 0.02
PPD (μmol/g DCW)	2.49 ± 0.62	8.29 ± 1.24	1.02 ± 0.11	3.68 ± 0.12
DM (μmol/g DCW)	0.06 ± 0.02	0.14 ± 0.01	0	0.18 ± 0.004
SQ (μmol/g DCW)	4.09 ± 0.42	3.18 ± 0.44	2.92 ± 0.20	2.52 ± 0.38
LST (μmol/g DCW)	1.34 ± 0.05	1.25 ± 0.23	1.38 ± 0.03	3.05 ± 0.38
ER (μmol/g DCW)	9.01 ± 0.39	9.66 ± 1.29	11.47 ± 0.37	12.21 ± 1.00
Total dammarenediol triterpenoids[b] (μmol/g DCW)	2.55 ± 0.63	8.43 ± 1.25	2.47 ± 0.33	7.42 ± 0.22

[a] All given data in this table representing mean values from three repeats with corresponding standard deviations. Yield referred to the sum of intracellular and extracellular contents. PPD, protopanaxadiol; DM, dammarenediol-II; SQ, squalene; LST, lanosterol; ER, ergosterol.
[b] Total dammarenediol triterpenoids including DM, PPD, Rh2 and Rg3.

integrated into the *δ DNA* sites of strain ZW-PPD-B, the PPD producing chassis, and 20 transformants harboring UGTPg45 were selected for fermentation. Based on HPLC analysis of product of fermentation, all transformants produced Rh2, among which transformant D20RH18 produced the highest Rh2, 1.45 μmol/g DCW (Fig. S4B, Table 3). The structure of Rh2 produced by D20RH18 was further confirmed by HPLC/ESIMS (Fig. 4B).

Similarly, both UGTPg45 and UGTPg29 (under the control of the constitutive promoters *TEF1* and *HXT7*, respectively) were integrated into the *δ DNA* sites of strain ZW-PPD-B to construct a yeast cell factory to synthesize ginsenoside Rg3. Based on HPLC analysis, the highest Rg3 production, 3.49 μmol/g DCW, was detected in transformant D20RG1 (Fig. 4A, Fig. S4C). The product of D20RG1 was further confirmed to be Rg3 by HPLC/ESIMS (Fig. 4B).

The cell biomass and the yields of total dammarane-type triterpenoids and ergosterol biosynthesis branch products (including lanosterol and ergosterol), and the variation of the components in the PPD producing chassis, ZW-PPD-A, ZW-PPD-B and the two cell factories, transformants D20RG1 and D20RH18,were compared (Table 3). The cell biomass of D20RH18 and D20RG1 after fermentation was only slightly lower than that of their chassis strain ZW-PPD-B, suggesting that the production of ginsenosides Rh2 and Rg3 did not obviously disturb the growth of these strains. The total dammarane-type triterpenoids (including DM, PPD, Rh2 and Rg3) produced by strain D20RH18 was significantly lower than that of chassis strain ZW-PPD-B and the Rg3 producing strain D20RG1 (Table 3). The total dammarane-type triterpenoids of the other 19 Rh2 producing transformants were all considerably lower than that of ZW-PPD-B and D20RG1 (data not shown). Squalene and 2,3-oxidosqualene are precursors of both dammarane-type triterpenoids and sterols (including lanosterol and ergosterol). No significant difference in accumulated squalene (2,3-oxidosqualene was not detected in all strains) or the total yield of lanosterol and ergosterol was detected between D20RH18 and ZW-PPD-B (Table 3). Thus, it seems that the introduction of UGTPg45 into D20RH18 resulted in the inhibition of the synthesis of dammarane-type triterpenoids. However, compared with the chassis strain ZW-PPD-B, the total dammarane-type triterpenoids of D20RG1, in which both of UGTPg45 and UGTPg29 were introduced, did not decrease markedly. It seems likely that the production of Rh2 in D20RH18 severely inhibits its ability to synthesize dammarane-type triterpenoids. In fact, little Rh2 was detected in D20RG1, due to the strong catalytic efficiency of UGTPg29. Indeed, the k_{cat} of UGTPg29 was more than 2500-fold that of UGTPg45 (Table 2). Thus, the Rh2 produced in D20RG1 cells could be rapidly transformed to Rg3. It thus is important to examine the effect of Rh2 on the synthesis of dammarane-type triterpenoids.

Considerable PPD accumulated within both cell factories, D20RG1 and D20RH18 (Table 3), which may be due to the low catalytic efficiency of UGTPg45 (Table 2). Overexpression of UGTPg45 in the two cell factories by introduction of increased numbers of copies of UGTPg45 or use of stronger promoters is expected to improve their Rg3 and Rh2 yields further. Moreover, ergosterol yields in D20RG1 and D20RH18 were higher than those of the chassis strains, ZW-PPD-A and ZW-PPD-B. Production of ergosterol is believed to be required for yeast growth. However, it seems the lower ergosterol yield in ZW-PPD-A and ZW-PPD-B was sufficient for normal growth (Table 3). Limitation of the metabolic flow from 2,3-oxidosqualene to lanosterol and then to ergosterol in D20RG1 and D20RH18 is also expected to increase the metabolic flow to Rg3 and Rh2, respectively. Moreover, the majority of the Rh2 and Rg3 produced by strains D20RH18 and D20RG1 was stored in the cells, extracellular transport of Rh2 and Rg3 would likely decrease feedback inhibition (or other unknown inhibition mechanisms) of the biosynthesis of dammarane-type triterpenoids or ginsenosides.

4. Conclusion

Two novel UDP-glycosyltransferases, UGTPg45 and UGTPg29, which selectively transfer two glucose moieties to the C3 hydroxyl of PPD or PPD-type ginsenosides step-by-step were identified from *P. ginseng*. The introduction of the two UGTs and UGTPg45 alone into the PPD producing yeast chassis brought the construction of two yeast cell factories to produce bioactive ginsenoside Rg3 and Rh2 from glucose, respectively. Analysis of the specificity and kinetics of the two UGTs also gives hint to uncover the metabolic processes of ginsenoside synthesis by sequential glucosylation of PPD in *P. ginseng*.

Acknowledgment

This work was financially supported by the National Basic Research Program of China (973 program: 2012CB721103 and 2012CB721105), the Knowledge Innovation Program from the Chinese Academy of Sciences (Nos. KSCX2-EW-G-13-1 and KSCX2-EW-J-12). The hormone-autotrophic (H; habituated) *Panax ginseng* callus line, the C-terminal 6 × His-tagged xylanase and the *A. thaliana*, cDNA were kindly provided by Prof. Weiming Cai, Dr. Ning Liu and Dr. Wenjuan Cai, respectively. All of them are from the Institute of Plant Physiology and Ecology, SIBS, CAS.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2015.03.003.

References

Bae, E.-A., Han, M.J., Kim, E.-J., Kim, D.-H., 2004. Transformation of ginseng saponins to ginsenoside Rh2 by acids and human intestinal bacteria and biological activities of their transformants. Arch. Pharm. Res. 27, 61–67.

Bae, S.H., Park, J.B., Zheng, Y.F., Jang, M.J., Kim, S.O., Kim, J.Y., Yoo, Y.H., Yoon, K.D., Oh, E., Bae, S.K., 2014. Pharmacokinetics and tissue distribution of ginsenoside Rh2 and Rg3 epimers after oral administration of BST204, a purified ginseng dry extract, in rats. Xenobiotica 44, 1099–1107.

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., Boeke, J.D., 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14, 115–132.

Cheng, L.-Q., Na, J.R., Bang, M.H., Kim, M.K., Yang, D.-C., 2008. Conversion of major ginsenoside Rb1 to 20(*S*)-ginsenoside Rg3 by *Microbacterium* sp. GS514. Phytochemistry 69, 218–224.

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

Flagfeldt, D.B., Siewers, V., Huang, L., Nielsen, J., 2009. Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. Yeast 26, 545–551.

Han, B., Park, M., Han, Y., Woo, L., Sankawa, U., Yahara, S., Tanaka, O., 1982. Degradation of ginseng saponins under mild acidic conditions. Planta Med. 44, 146–149.

Han, J.Y., Kim, H.J., Kwon, Y.S., 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.

Han, J.Y., Kwon, Y.S., Yang, D.C., Jung, Y.R., Choi, Y.E., 2006. Expression and RNA interference-induced silencing of the dammarenediol synthase gene in *Panax ginseng*. Plant Cell Physiol. 47, 1653–1662.

Huang, X., Madan, A., 1999. CAP3: A DNA sequence assembly program. Genome Res. 9, 868–877.

Hwang, J.-T., Kim, S.-H., Lee, M.-S., Kim, S.H., Yang, H.-J., Kim, M.-J., Kim, H.-S., Ha, J., Kim, M.S., Kwon, D.Y., 2007. Anti-obesity effects of ginsenoside Rh2 are associated with the activation of AMPK signaling pathway in 3T3-L1 adipocyte. Biochem. Biophys. Res. Commun. 364, 1002–1008.

Kim, H.-S., Lee, E.-H., Ko, S.-R., Choi, K.-J., Park, J.-H., Im, D.-S., 2004. Effects of ginsenosides Rg3 and Rh2 on the proliferation of prostate cancer cells. Arch. Pharm. Res. 27, 429–435.

Li, Y., Li, P., Wang, Y., Dong, R., Yu, H., Hou, B., 2014. Genome-wide identification and phylogenetic analysis of Family-1 UDP glycosyltransferases in maize (*Zea mays*). Planta 239, 1265–1279.

Liu, T.-G., Huang, Y., Cui, D.-D., Huang, X.-B., Mao, S.-H., Ji, L.-L., Song, H.-B., Yi, C., 2009. Inhibitory effect of ginsenoside Rg3 combined with gemcitabine on angiogenesis and growth of lung cancer in mice. BMC Cancer 9, 250.

Min, J.-K., Kim, J.-H., Cho, Y.-L., Maeng, Y.-S., Lee, S.-J., Pyun, B.-J., Kim, Y.-M., Park, J.H., Kwon, Y.-G., 2006. 20 (*S*)-Ginsenoside Rg3 prevents endothelial cell apoptosis via inhibition of a mitochondrial caspase pathway. Biochem. Biophys. Res. Commun. 349, 987–994.

Mochizuki, M., Yoo, Y.C., Matsuzawa, K., Sato, K., Saiki, I., TONo-oKA, S., Samukawa, K.-i., Azuma, I., 1995. Inhibitory effect of tumor metastasis in mice by saponins, ginsenoside-Rb2, 20 (*R*)-and 20 (*S*)-ginsenoside-Rg3, of red ginseng. Biol. Pharm. Bull. 18, 1197–1202.

Nakata, H., Kikuchi, Y., Tode, T., Hirata, J., Kita, T., Ishii, K., Kudoh, K., Nagata, I., Shinomiya, N., 1998. Inhibitory effects of ginsenoside Rh2 on tumor growth in nude mice bearing human ovarian cancer cells. Cancer Sci. 89, 733–740.

Park, J., Lee, K.Y., Oh, Y.J., Kim, K.-W., Lee, S.K., 1997. Activation of caspase-3 protease via a Bcl-2-insensitive pathway during the process of ginsenoside Rh2-induced apoptosis. Cancer Lett. 121, 73–81.

Park, M.W., Ha, J., Chung, S.H., 2008. 20 (*S*)-ginsenoside Rg3 enhances glucose-stimulated insulin secretion and activates AMPK. Biol. Pharm. Bull. 31, 748–751.

Popovich, D.G., Kitts, D.D., 2004. Generation of ginsenosides Rg3 and Rh2 from North American ginseng. Phytochemistry 65, 337–344.

Shao, Z., Zhao, H., Zhao, H., 2009. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. Nucl. Acids Res. 37, e16.

Shibata, S., 2001. Chemistry and cancer preventing activities of ginseng saponins and some related triterpenoid compounds. J. Korean Med. Sci.16 (Suppl.), S28S37.

Shinkai, K., Akedo, H., Mukai, M., Imamura, F., Isoai, A., Kobayashi, M., Kitagawa, I., 1996. Inhibition of *in vitro* tumor cell invasion by ginsenoside Rg3. Cancer Sci. 87, 357–362.

Su, J.-H., Xu, J.-H., Lu, W.-Y., Lin, G.-Q., 2006. Enzymatic transformation of ginsenoside Rg_3 to Rh_2 using newly isolated *Fusarium proliferatum* ECU2042. J. Mol. Catal. B: Enzym. 38, 113–118.

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

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

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

Yan, Q., Zhou, W., Li, X., FENG, M., Zhou, P., 2008. Purification method improvement and characterization of a novel ginsenoside-hydrolyzing β-glucosidase from *Paecilomyces bainier* sp. 229. Biosci. Biotechnol. Biochem. 72, 352–359.

Yan, X., Fan, Y., Wei, W., Wang, P., Liu, Q., Wei, Y., Zhang, L., Zhao, G., Yue, J., Zhou, Z., 2014. Production of bioactive ginsenoside compound K in metabolically engineered yeast. Cell Res. 24, 770–773.