

The Cyt P450 Enzyme CYP716A47 Catalyzes the Formation of Protopanaxadiol from Dammarenediol-II During Ginsenoside Biosynthesis in *Panax ginseng*

Jung-Yeon Han^{1,3}, Hyun-Jung Kim^{1,3}, Yong-Soo Kwon² and Yong-Eui Choi^{1,*}

¹Department of Forest Resources, College of Forest and Environmental Sciences, Kangwon National University, Chunchon 200-701, Republic of Korea

²College of Pharmacy, Kangwon National University, Chunchon 200-701, Republic of Korea

³These authors contributed equally to this work

*Corresponding author: E-mail, yechoi@kangwon.ac.kr; Fax, +82-33-252-8310

(Received August 29, 2011; Accepted October 24, 2011)

Ginseng (Panax ginseng C.A. Meyer) is one of the most popular medicinal herbs and contains pharmacologically active components, ginsenosides, in its roots. Ginsenosides, a class of tetracyclic triterpene saponins, are thought to be synthesized from dammarenediol-II after hydroxylation by the Cyt P450 (CYP) enzyme and then glycosylation by glycosyltransferase (GT). However, no genes encoding the hydroxylation and glycosylation in ginsenoside biosynthesis have been identified. Here, we identify protopanaxadiol synthase, which is a CYP enzyme (CYP716A47), to be involved in the hydroxylation of dammarenediol-II at the C-12 position to yield protopanaxadiol. Nine putative full CYP sequences were isolated from the expressed sequence tags (ESTs) of methyl jasmonate (MeJA)-treated adventitious ginseng roots. The CYP716A47 gene product was selected as the putative protopanaxadiol synthase because this gene was transcriptionally activated not only by MeJA treatment but also in transgenic ginseng that overexpresses squalene synthase and overproduces ginsenosides. In vitro enzymatic activity assays revealed that CYP716A47 catalyzed the oxidation of dammarenediol-II to produce protopanaxadiol. Ectopic expression of CYP716A47 in recombinant WAT21 yeasts that were fed dammarenediol-II vielded protopanaxadiol. Furthermore, co-expression of the dammarenediol synthase gene (PgDDS) and CYP716A47 in yeast yielded protopanaxadiol without adding dammarenediol-II. The chemical structures of the protopanaxadiol products from dammarenediol-II were confirmed using liquid chromatography-atmospheric pressure chemical ionization mass spectrometry (LC/APCIMS). Thus, CYP716A47 is a dammarenediol 12-hydroxylase that produces protopanaxadiol from dammarenediol-II.

Keywords: CYP716A47 • Cyt P450 • Dammarenediol • Ginsenoside • Protopanaxadiol.

Abbreviations: CYP, cytochrome P450; EST, expressed sequence tag; GC/MS, gas chromatography-mass

spectrometry; GT, glycosyltransferase; LC/APCIMS, liquid chromatography-atmospheric pressure chemical ionization mass spectrometry; MeJA, methyl jasmonate; ORF, open reading frame; RT-PCR, reverse transcription-PCR.

Introduction

Triterpenoid saponins are secondary metabolites of isoprenoidal compounds and are present in higher plants. They exhibit a wide range of structural diversity and biological activities among plant species. These molecules also have considerable commercial value and are exploited as medicinal drugs (Hostettmann and Marston 1995, Vogler et al. 1999, Shibata 2001). The natural role of saponins in plants is probably in defense against pathogen and pest attacks (Osbourn 1996). The primary components of triterpenoid saponins are oleanane (β -amyrin), ursane (α -amyrin), lupeol or dammarene-type triterpenoid skeletons.

Panax ginseng is well recognized by consumers for its notable pharmacological effects on cancer, diabetes mellitus and neurodegenerative disease, among others. Ginsenosides are considered the primary components of ginseng root responsible for its bioactivity. Panax ginseng roots contain at least 4% ginsenosides by dry weight (Shibata 2001). Seven dammarenetype tetracyclic triterpenes (ginsenoside Rb₁, Rb₂, Rc, Rd, Re, Rf and Rg_1) are reported to be the major ginsenoside constituents, and only ginsenoside Ro is an oleanane-type pentacyclic triterpene, which is found in minor amounts in P. ginseng. These dammarane-type ginsenosides are divided into two groups according to the aglycone structure: the panaxadiol (Rb₁, Rb₂, Rc and Rd) and panaxatriol groups (Rg₁, Re, Rf and Rg₂). It is well known that the dammarene-type triterpene is a major compound in the genera Panax (Kushiro et al. 1997) and Gynostemma (Cui et al. 1999).

The first step in biosynthesis of dammarane-type ginsenosides is the cyclization of 2, 3-oxidosqualene to

Plant Cell Physiol. 52(12): 2062–2073 (2011) doi:10.1093/pcp/pcr150, available online at www.pcp.oxfordjournals.org © The Author 2011. Published by Oxford University Press on behalf of Japanese Society of Plant Physiologists. All rights reserved. For permissions, please email: journals.permissions@oup.com Downloaded from https://academic.oup.com/pcp/article/52/12/2062/1821615 by guest on 16 August 2022





Fig. 1 Proposed biosynthetic pathway for ginsenosides in *P. ginseng*. Squalene epoxidase converts the squalene into 2,3-oxidosquane, which is converted into triterpene aglycones (dammarenediol-II or β -amyrin) through dammarenediol synthase or β -amyrin synthase. The triterpene aglycones subsequently undergo oxidation and glycosylation, and are finally converted into triterpene saponins (ginsenosides).

dammarenediol-II, a reaction that is catalyzed by a dammarenediol synthase (**Fig. 1**). Two homologous dammarenediol synthases from *P. ginseng* (DDS and PNA) have been functionally characterized (Han et al. 2006, Tansakul et al. 2006). Dammarenediol-II is thought to be converted to a ginsenoside after hydroxylation by Cyt P450 (CYP) enzymes (Shibuya et al. 2006) and subsequent glycosylation by glycosyltransferase (GT) (Kushiro et al. 1997, Choi et al. 2005, Shibuya et al. 2006). The ginsenoside Ro is thought to be synthesized from oleanolic acid, which is a product of β -amyrin (Shibata 1977).

Both CYP and GT are in plant genome supergene families. In plants, CYPs play critical roles in oxidative reactions during the biosynthesis of diverse plant secondary metabolites, lignins, terpenoids, sterols, fatty acids, hormones, pigments and defense-related phytoalexins (Schuler 1996). In *P. ginseng*, two CYP genes are thought to be involved in dammarene-type ginsenoside biosynthesis. One of these genes might be involved in dammarenediol hydroxylation at the C-12 position for protopanaxadiol synthesis. Another gene is involved in protopanaxadiol hydroxylation at the C-6 position for protopanaxatriol synthesis, and these two compounds are used as the backbones for dammarene-type ginsenosides. However, the CYP (protopanaxadiol and protopanaxatriol synthase) and GT genes remain uncharacterized. In the present study, we isolated nine putative full CYP gene sequences from the expressed sequence tag (EST) sequences of elicitor (MeJA)-treated adventitious roots. *CYP716A47* was selected as the putative CYP gene by profiling transcription in roots through reverse transcriptase–PCR (RT–PCR) analysis with and without MeJA treatment as well as transgenic cell lines that overexpressed squalene synthase (*PgSS1*), which overproduces ginsenosides. Ectopic expression of *CYP716A47* and co-expression of *PgDDS* and *CYP716A47* in yeast yielded protopanaxadiol, indicating that CYP716A47 is protopanaxadiol synthase, which is a critically important step in ginsenoside biosynthesis.

Results

Isolation and classification of CYP genes from the *P. ginseng* EST library

Elicitor (MeJA) treatment strongly activates triterpene saponin biosynthesis in many species, including *P. ginseng* (Gundlach et al. 1992, Suzuki et al. 2002, Han et al. 2006). In many studies, MeJA treatment produced the greatest stimulation of ginsenoside biosynthesis (Lee et al. 2004, Han et al. 2006). To isolate the CYP genes involved in ginsenoside biosynthesis, EST sequences



Table 1 The EST sequences that are homologous to the CYP genes from the MeJA-treated adventitious P. ginseng roots

CYP450 family	Typical EST ID	GenBank accession No. of gene	No. of ESTs	Nomenclature of CYPs	Homology to (species name)	Description	Accession No.	E-value
CYP71	KYERRC001_46-A08 KYERRC001_36-B03 KYERRC001_26-F07 KYERRC001_18-F12 KYERRC001_17-H02 KYERRC001_10-D12 KYERRC001_19-G01 KYERRC001_43-F04 KYERRC001_45-G06	JN604540 JN604541	7 5 1 2 1 1 1 1	CYP71D312 CYP71D313 CYP71D314 CYP71D315v2 CYP71D315v1 CYP71D316 CYP71D317 CYP71D318v1 CYP71D318v2	CYP71D55 (Hyoscyamus muticus)	Premnaspirodiene oxygenase	EF569601	2e-09
CYP72	KYERRC001_39-E08 KYERRC001_42-H03 KYERRC001_31-H09	JN604542	2 1 1	CYP72A219 CYP72A220 CYP72A221	CYP72A57 (Nicotiana tabacum)	Unknown	DQ350360	3e-64
CYP73	KYERRC001_19-F10	JN604543	5	CYP73A100	CYP73A5 (Arabidopsis thaliana)	Cinnamic acid 4-hydroxylase	D78596	3e-166
CYP82	KYERRC001_42-G11 KYERRC001_19-F02 KYERRC001_01-H12	JN604544 JN604545	3 1 1	CYP82H23 CYP82H24 CYP82D47	CYP82A1 (Pisum sativum)	Unknown wound- inducible P450 hydroxylase	AF175278	2e-09
CYP90	KYERRC001_12-D12		3	CYP90A29	CYP90 A2 (Pisum sativum)	Putative brassinosteroid C-23 hydroxylase	AB218762	9e-122
CYP94	KYERRC001_27-A10		1	CYP94A33	CYP94A1 (Vicia sativa)	Fatty acid omega-hydroxylase	O81117	0.0
CYP98	KYERRC001_26-B07 KYERRC001_08-H05		1 1	CYP98A58 CYP98A59	CYP98A13 (Ocimum basilicum)	<i>p-</i> Coumaroyl shikimate 3'-hydroxylase	AY082612	1×10^{-147}
CYP716	KYERRC001_23-C09	JN604536	7	CYP716A47	CYP716A2 (Medicago truncatula)	Oxidation of β-amyrin and erythrodiol at the C-28 position	FN 995112	6e-169
CYP734	KYERRC001_09-G04		1	CYP734A23	CYP734 A7 (Lycopersicon esculentum)	Brassinosteroid C-26 hydroxylase	AB223042	5e-49
CYP736	KYERRC001_22-F01	JN604539	3	CYP736A12	CYP736B (Vitis arizonica × Vitis rupestris)	Unknown disease responsive	FJ828518	5e-24
CYP749	KYERRC001_20-C10	JN604538	2	CYP749A22	CYP86A22 (Petunia × hybrida)	Fatty Acyl-CoA- hydroxylase	DQ099540	2e-17

were isolated from MeJA-treated adventitious roots. Of the 4,140 high-quality EST (2,966 unigenes) sequences, 53 EST sequences were homologous to CYP mRNAs (**Table 1**). They were grouped into 11 CYP families with single and multiple copies (**Table 1**). The most abundant CYP transcripts in ginseng adventitious roots were for the CYP71D subfamily, which included seven isoforms. The other CYP family genes with multiple transcripts were CYP716, CYP73, CYP72 and CYP82.

Because the EST sequences were isolated from MeJA-treated adventitious roots, we hypothesized that multiple transcripts might comprise the CYP genes involved in ginsenoside biosynthesis. Therefore, CYP genes with multiple copies were selected as the genes possibly responsible for ginsenoside biosynthesis. Nine full-length sequences (CYP71D312, CYP71D313, CYP72A219, CYP73A100, CYP82H23, CYP82D47, CYP716A47, CYP736A12 and CYP749A22) were obtained from the CYP mRNA transcript sequences with >2 EST copies, except for CYP82D47 because this gene is phylogenetically related to the Glycine max CYP93E1 gene (β -amyrin and sophoradiol 24-hydroxylase).

Phylogenic analysis of the full sequences of the nine genes with multiple EST transcripts

Analysis of the phylogenetic relationship between the nine putative CYP genes from *P. ginseng* and other plant CYP genes (**Fig. 2**) revealed that *CYP716A47* is similar to the recently identified *Medicago truncatula CYP716A12*, whose product is involved in β -amyrin and erythrodiol oxidation at the C-28 position (Carelli et al. 2011). The CYP716A47 deduced amino acid sequences were 49% homologous to that of CYP716A12. In addition, both CYP71D312 and CYP71D313 shared close amino acid sequence similarity with the CYP71 family genes involved in terpene biosynthesis (Takahashi et al. 2007). Both *CYP82H23*

Characterization of protopanaxadiol synthase (CYP716A47)





Fig. 2 Phylogenetic tree constructed based on the deduced amino acid sequences for the *P. ginseng* CYPs (bold letters) and other plant CYPs. The tree was constructed by the Neighbor–Joining method using the ClustalX program. Bar = 0.1 amino acid substitutions/site. The species abbreviations are *Gm*, *Glycine max*; *Sc*, *Solanum chacoense*; *At*, *Arabidopsis thaliana*; *Ge*, *Glycyrrhiza echinata*; *Sm*, *Solanum melongena*; *Gu*, *Glycyrrhiza uralensis*; *Sb*, *Sorghum bicolor*; *Vs*, *Vicia sativa*; *Cr*, *Catharanthus roseus*; *Hm*, *Hyoscyamus muticus*; *Ps*, *Pisum sativum*; *Ph*, *Petunia hybrida*; *Va*, *Vitis arizonica* × *Vitis rupestris*; *Mt*, *Medicago truncatula*; *Nt*, *Nicotiana tabacum*.



and CYP82D47 are phylogenetically related to the *G. max* CYP93E1 gene (β -amyrin and sophoradiol 24-hydroxylase). CYP73A100 showed significant amino acid sequence similarity to *Arabidopsis thaliana* CYP73A5 (cinnamate 4-hydroxylase), which is involved in phenylpropanoid biosynthesis (Urban et al. 1997). CYP736A12 is closely related to the uncharacterized disease response genes in *Vitis arizonica* × *Vitis rupestris* CYP736B and is phylogenetically related to *A. thaliana* ferulate 5-hydroxylase (CYP84A1). CYP72A219 is related to the secologanin synthase, which converts loganin into secologanin (Irmler et al. 2000), and CYP749A22 is related to *Arabidopsis* CYP72C1, which plays a role in inactivating brassinosteroids, although its biochemical activity is unknown (Thornton et al. 2010).

Transcriptional activities of the nine CYP mRNAs after elicitor (MeJA) treatment and in transgenic ginseng that overproduces ginsenosides

To narrow the range of candidate CYP genes involved in ginsenoside biosynthesis, the transcription activities of the nine isolated CYP mRNAs were investigated using RT–PCR analysis. Adventitious roots were exposed to $10 \,\mu$ M MeJA for 12 or 48 h, which has previously yielded the best activation results for genes involved in ginsenoside biosynthesis (Lee et al. 2004, Han et al. 2006, Han et al. 2010). CYP73A100, CYP749A22 and CYP716A47 were up-regulated by MeJA (**Fig. 3A**); CYP82D47 and CYP82H23 were down-regulated by MeJA; and CYP736A12, CYP71D312, CYP71D313 and CYP72A219 were unresponsive.

Previously, we reported that overexpression of the squalene synthase gene (PgSS1) in transgenic P. ginseng was followed by up-regulation of downstream genes, such as squalene epoxidase (SE), β -amyrin synthase (β -AS) and cycloartenol synthase (CS) (Lee et al. 2004). The enhanced activity of the PgSS1 enzyme resulted in a remarkable increase in the ginsenoside content (Lee et al. 2004). We hypothesized that the PgSS1-overexpressing transgenic ginseng roots up-regulate protopanaxadiol synthase transcription. Thus, the transcriptional activities of the nine CYP mRNAs in the transgenic ginseng roots that overexpress PgSS1 were monitored using RT-PCR (Fig. 3B). CYP716A47 and CYP749A22 were transcriptionally activated in the two transgenic lines compared with the wild type. CYP736A12 and CYP71D313 were transcriptionally activated only in transgenic line 5 and not in transgenic line 6. The other five CYPs showed no changed or decreased transcriptional activity in the transgenic lines.

Ectopic expression of CYP716A47 cDNA in WAT21 yeast

The full-length cDNA clone of CYP716A47 was 1,485 bp long with a 487 amino acid open reading frame (ORF), yielding a protein with a predicted molecular mass of 55.36 kDa. To examine functionally the hydroxylation activity of CYP716A47 in protopanaxadiol production, the ORF region of CYP716A47



Fig. 3 RT–PCR analysis of the nine CYP genes in the MeJA-treated roots and transgenic PgSS1-overexpressing ginseng roots. (A) RT–PCR analysis of the MeJA-treated roots. The arrows indicate activated transcripts after MeJA treatment. (B) RT–PCR analysis of the transgenic PgSS1-overexpressing ginseng roots. The arrows indicate activated transcripts in the transgenic roots.

cDNA was inserted into the pYES2.1 expression vector and expressed in WAT21 yeast under the control of the regulatable promoter (GAL1). The yeast extracts were analyzed using total ion chromatograms from liquid chromatography-atmospheric pressure chemical ionization mass spectrometry (LC/APCIMS). Because dammarenediol-II was exogenously applied to the yeast, both dammarenediol-II and protopanaxadiol were examined in the total ion chromatogram from LC/APCIMS (Fig. 4). The retention times of the standard dammarenediol-II and protopanaxadiol peaks were 24.70 and 32.59 min, respectively (Fig. 4). Both dammarenediol-II and protopanaxadiol were identified in CYP716A47-expressing yeast using LC/APCI-MS (Fig. 4B). A protopanaxadiol signal was not detected in the control yeast extract with the empty vector, although a peak from the exogenously added dammarenediol-II was detected (Fig. 4A).

The protopanaxadiol signal at the 32.59 min retention time in yeast expressing the CYP716A47 gene was analyzed using the MS fragmentation pattern (**Fig. 5**). The LC/APCIMS fragmentation pattern in yeast expressing the CYP716A47 gene included the *m*/*z* ratios of 407 [M-3H₂O+H]⁺, 425 [M-2H₂O+H]⁺ and 443 [M-H₂O+H]⁺, which are the same as for authentic protopanaxadiol (**Fig. 5**). This result indicates that CYP716A47 is a dammarenediol 12-hydroxylase gene and its product converts dammarenediol-II to protopanaxadiol.

In vitro CYP716A47 enzymatic activity assay

To examine the hydroxylation activity of CYP716A47 in protopanaxadiol production, microsomal fractions from



WAT21 yeast expressing CYP716A47 were incubated with dammarenediol-II for 2 h at 30°C. Gas chromatography-mass spectrometry (GC/MS) analysis revealed that CYP716A47 converted dammarenediol-II (37.95 min) to a product (42.46 min)



Fig. 4 The total ion chromatogram from LC/APCIMS analysis of the CYP716A47 product in yeast. (A) The LC chromatogram of the yeast cell extract with an empty vector as a control. (B) The LC chromatogram of the yeast cell extract with pYES2-CYP716A47. The arrow indicates a protopanaxadiol peak at a retention time of 32.59 min. (C) The LC chromatogram of a protopanaxadiol standard.

with the same GC retention time (42.46 min) and mass spectra as those of the authentic protopanaxadiol standard (**Fig. 6A**). No activity was observed in yeast transformed with the empty vector control (**Fig. 6C**) or in boiled enzyme preparations (**Fig. 6B**). This result suggests that CYP716A47 catalyzes conversion of dammarenediol-II to protopanaxadiol.

Co-expression of PgDDS and CYP716A47 in yeast

The *PgDDS* and *CYP716A47* genes were subcloned into the yeast expression vector pESC-URA with double gene expression cassettes, and the ethyl acetate extracts from the cells were analyzed by LC/APCIMS. The total ion chromatogram from LC/APCIMS revealed that co-expression of *PgDDS* and *CYP716A47* in yeast clearly produces a peak (**Fig. 7B**) at a retention time of 32.59 min, which is the same retention time as for the protopanaxadiol standard (**Fig. 7C**). In the control yeast with only an empty vector, no peak was detected at the retention time (32.59 min) of the standard protopanaxadiol peak (**Fig. 7A**).

The protopanaxadiol signal at a retention time of 32.59 min in yeast that expressed both the *PgDDS* and *CYP716A47* genes was analyzed using the MS fragmentation patterns (**Fig. 7D**). The LC/APCIMS spectra of a peak at the 32.59 min retention time from yeast that expressed both the *PgDDS* and *CYP716A47* genes included the *m*/z ratios of 407 [M-3H₂O+H]⁺, 425 [M-2H₂O+H]⁺ and 443 [M-H₂O+H]⁺, which are the same as for authentic protopanaxadiol (**Fig. 5B**). The protopanaxadiol yield on the galactose induction medium after 2 d of culture was approximately 17.32 μ g g⁻¹ FW.

HPLC analysis with UV fluorescent detection revealed that the chromatogram fraction pattern was similar between the recombinant (**Supplementary Fig. S1B**) and control yeast









Fig. 6 In vitro conversion of dammarenediol-II to protopanaxadiol by the reaction of microsomes from yeast expressing CYP716A47. (A) The GC/MS chromatograms of the reaction products of microsomes with dammarenediol-II. Line a reveals the protopanaxadiol (42.46 min) production from the reaction of microsomes with pYES2-CYP716A47 and dammarenediol-II. Lines b and c reveal the reaction products from boiled microsomes with dammarenediol-II and microsomes with empty vector with dammarenediol-II, respectively. Lines d and e are chromatograms for authentic standards of dammarenediol-II and protopanaxadiol, respectively. (B) The MS spectrum for a peak (42.46 min) detected in the reaction products of microsomes with CYP716A47 in line a of A. (C) The MS spectrum of protopanaxadiol standard.

(**Supplementary Fig. S1A**), except for a new peak at a retention time of 32.59 min, which corresponds to the standard protopanaxadiol peak (**Supplementary Fig. S1C**).

The authentic protopanaxatriol peak was found at a retention time of 19.20 min. Protopanaxatriol was not detected in yeast that expressed *PgDDS* and *CYP716A47* (data not shown). Moreover, a traceable LC/APCIMS signal for protopanaxatriol was not detected. This result indicates that CYP716A47 cannot further hydroxylate protopanaxadiol to biosynthesize protopanaxatriol.

To confirm the possibility that protopanaxadiol was catalyzed by yeast endogenous hydroxylation enzymes, yeast with *PgDDS* as a negative control was analyzed by GC/MS. In the yeast with *PgDDS* as a negative control, dammarenediol-II (**Supplementary Fig. S2A**) was detected at the same retention time (37.95 min) as for the dammarenediol-II standard, but no protopanaxadiol peak was detected (**Supplementary Fig. S2A**). Co-expression of *PgDDS* and *CYP716A47* in yeast clearly produces a peak (**Supplementary Fig. S2B**) at a retention time of 42.56 min, which is the same retention time (**Fig. 4D**) as for the protopanaxadiol standard. The major peaks (m/z = 109 and 424) in MS fragmentation (**Supplementary Fig. S2C**) were identical to those (**Fig. 4E**) of the authentic protopanaxadiol.

Discussion

Isolation of the CYP gene involved in ginsenoside biosynthesis

The CYP superfamily is a large and diverse group of enzymes. In *A. thaliana*, 246 CYP genes were reported (Nelson 2006). Therefore, it is difficult to analyze each CYP gene using a reverse





Fig. 7 LC/APCIMS analysis of protopanaxadiol production in WAT21 yeast that expressed both *PgDDS* and *CYP716A47*. (A) The total ion chromatogram of the yeast extracts with an empty vector as a control. (B) The total ion chromatogram of the yeast extracts with both *PgDDS* and *CYP716A47*; the arrow indicates the protopanaxadiol signal at a retention time of 32.59 min. (C) The total ion chromatogram of a protopanaxadiol standard. (D) The MS spectrum of a protopanaxadiol peak at a retention time of 32.59 min from yeasts with both *PgDDS* and *CYP716A47*.

genetics approach (Nelson 2006). There have been large-scale attempts to isolate the CYPs and GTs involved in ginsenoside biosynthesis in *P. ginseng* (Jung et al. 2003, Choi et al. 2005), *P. quinquefolius* (Sun et al. 2010) and *P. notoginseng* (Chen et al. 2011). In this work, we isolated *CYP716A47* from the EST sequences in MeJA-treated ginseng roots and successfully characterized CYP716A47 as a protopanaxadiol synthase.

Many studies have shown that MeJA treatment produces the greatest stimulation of ginsenoside biosynthesis (Lee et al. 2004, Han et al. 2006). We hypothesized that the CYP genes involved in ginsenoside biosynthesis might yield multiple transcripts in the EST sequences isolated from MeJA-treated roots. Among the 53 EST clones that were homologous to CYP genes, nine full sequences were obtained from ESTs with multiple transcripts. MeJA treatment activates not only the genes involved in triterpene biosynthesis but also various defensive mechanisms (Cheong and Choi 2003). Previously, we reported that upon PgSS1 overexpression in transgenic P. ginseng roots, all of the downstream genes tested were up-regulated, for example SE, β -AS and CAS. The enhanced activity of the PgSS1 enzyme yielded a remarkable increase in phytosterol and ginsenoside content. Two genes (CYP716A47 and CYP749A22) were finally selected as the best candidate CYP genes involved in ginsenoside biosynthesis that had elevated transcription both

in MeJA-treated roots and in the *PgSS1*-overexpressing transgenic ginseng roots.

Functional characterization of dammarenediol 12-hydroxylase

Using yeast expression analysis, CYP716A47 was characterized as a protopanaxadiol synthase with dammarenediol 12-hydroxylase activity, as demonstrated by the construction of a recombinant CYP716A47-expressing yeast, which yielded protopanaxadiol from dammarenediol after the yeast were fed dammarenediol-II. Co-expression of dammarenediol synthase and CYP716A47 in yeast yielded protopanaxadiol without feeding of dammarenediol-II.

CYP716A47 belongs to the CYP716 family in the CYP85 clan. Recently, it was shown that mutants in CYP716A12 from *M. truncatula* were unable to produce hemolytic saponins and only synthesized soyasaponins; thus, CYP716A12 lacked hemolytic activity (Carelli et al. 2011). Furthermore, in vitro enzymatic activity assays indicated that CYP716A12 catalyzes β -amyrin and erythrodiol oxidation at the C-28 position, yielding oleanolic acid, which indicates that it is a multifunctional oxidase (Carelli et al. 2011). Accordingly, CYP716 family genes are thought to be involved in triterpene saponin biosynthesis.



The first identification of a triterpene hydroxylase cDNA was CYP93E1, and the product of this gene converts β -amyrin into the sophoradiol in G. max (Shibuya et al. 2006). Because all of the CYP93 family members thus far identified have been flavonoid biosynthesis-related enzymes, it was unexpected that CYP93E1 would encode β -amyrin and sophoradiol 24-hydroxylase (Shibuya et al. 2006). Another CYP is β -amyrin 11-oxidase (CYP88D subfamily in the CYP85 clan), which converts β -amyrin to 11-oxo- β -amyrin during biosynthesis of the triterpene sweetener glycyrrhizin (Seki et al. 2008). These two characterized enzymes, CYP93E1 and CYP88D6, are involved in oleanane-type saponin biosynthesis. The major type of ginsenoside is the tetracyclic dammarene-type triterpene, which is structurally similar to sterols. Jung et al. (2003) suggested that the CYP involved in ginsenoside biosynthesis might have evolved from sterol biosynthesis. In this work, CYP716A47 in P. ginseng is categorized into the CYP716 family, which is involved in dammarene-type saponin biosynthesis.

In addition to the triterpene glycosides, triterpene aglycones also have interesting biological activities. For example, soyasapogenol B has hepatoprotective activity (Sasaki et al. 1997), and oleanolic acid and ursolic acid have anti-inflammatory and anti-tumor activities (Banno et al. 2004), among others. In P. ginseng, > 30 ginsenosides have been identified and divided into two groups according to the aglycone structure: the panaxadiol (Rb1, Rb2, Rc and Rd) and panaxatriol (Rg1, Re, Rf and Rg₂) groups. Each ginsenoside has been shown to have different pharmacological effects, including immune system modulation as well as anti-stress, anti-hyperglycemic, anti-inflammatory, anti-oxidant and anti-cancer effects (Briskin 2000, Shibata 2001). However, biotransformation may be required before these ginsenosides can be active in mammalian systems. Recent studies have demonstrated that ginsenoside metabolites have greater biological effects than ginsenosides (Jia et al. 2004, Popovich and Kitts 2004). The anti-tumor activities of the ginsenoside Rh₂ and protopanaxadiol are more potent than those of the ginsenoside Rg₃ (Bae et al. 2004). Recently, protopanaxadiol, a triterpene aglycone hydrolyzed from various ginsenosides, has been shown to exhibit apoptotic effects on cancer cells through various signaling pathways and has been reported as cytotoxic to multidrug-resistant tumors (Jia et al. 2004, Popovich and Kitts 2004, Li et al. 2006). Therefore, PanaGin Pharmaceuticals (British Columbia, Canada) is developing protopanaxadiol (Pandimex) for the treatment of lung cancer and other solid tumors and has completed pre-investigational new drug communications with the US Food and Drug Administration. In 2011, the company is preparing to file an investigational new drug for HemoMex S (PBD-1226) as an adjuvant treatment for lung cancer and other solid tumors (PanaGin Pharmaceuticals, Inc. Further information available at http:// www.panagin.com).

Because the production of triterpene is not practical through organic synthesis, these compounds must be

isolated from natural sources or by ginsenoside hydrolysis. Protopanaxdiol production through co-expression of *PgDDS* and *CYP716A47* in yeast in this study might represent a promising way to produce useful dammarene-type triterpenes using genetic engineering.

Materials and Methods

Plant materials and EST library construction

Total RNA was isolated from in vitro cultured adventitious roots. $Poly(A)^+$ RNA was isolated using a Poly(A) quick mRNA isolation kit (Stratagene). cDNA libraries were constructed using the Creator SMART cDNA library construction kit (Clontech). DH10B was used as a host strain, and pDNR-LIB was used as a cloning vector.

Single-pass partial sequences were determined using an automated DNA sequencer (model ABI Prism 3700, Applied Biosystems). A total of 4,226 cDNA strands (NCBI GenBank dbEST accession Nos. HS076062–HS080287) were randomly isolated and sequenced from the 5' end using the ABI 3730 XL automatic DNA sequencer (Applied Biosystems). The resulting ESTs were compared with GenBank and dbEST sequences using the BLASTX algorithm on the NCBI website.

Isolation and comparison of CYP protein sequences

Nine complete genes were isolated by sequencing the nine selected putative CYP ESTs. The GenBank accession numbers of the nine putative CYP sequences from P. ginseng were the following: CYP72A219 (JN604542), CYP73A100 (JN604543), CYP736A12 (JN604539), CYP82H23 (JN604544), CYP82D47 (JN604545), CYP71D312 (JN604540), CYP71D313 (JN604541), CYP749A22 (JN604538) and CYP716A47 (JN604536). Their nucleotide and predicted amino acid sequences were analyzed using the DNASIS program (Hitachi Software Engineering Co.). To analyze the phylogenic relationships among these gene sequences, amino acid sequences were obtained from EMBL, GenBank and DDBJ sequence data. Multiple sequence alignments were generated using the CLUSTAL W program (Thompson et al. 1994). Phylogenetic analysis of the deduced amino acid alignments was performed using the Neighbor-Joining method with TreeView software (Page, 1996). Bootstrap analysis with 1,000 replicates was used to assess the strength of the nodes in the tree (Felsenstein 1985).

RT-PCR analysis

Total RNA was isolated from the MeJA-treated adventitious roots or different organs of ginseng plants and reverse transcribed using the ImProm-II Reverse Transcription System (Promega). First-strand cDNA was used as the template for RT–PCR analysis, which was performed as follows: 96° C for 5 min, 30 cycles of 96° C for 30 s, 60° C for 30 s, 72° C for 1 min, and a final 10 min extension at 72° C. The cDNA for β -actin was



used as a control for RNA integrity and loading accuracy. Electrophoresis of the products was performed using 1% agar/ $0.5 \times$ TBA buffer. The primer sequences are summarized in **Supplementary Table S1**. The RT–PCR analyses were repeated twice, and representative data are shown in the figures.

Expression of CYP716A47 in yeast

To construct an expression plasmid vector for yeast, ORFs from CYP716A47 were amplified from cDNA using PCR (25 cycles of 40 s at 94°C, 40 s at 55°C and 2 min at 72°C) with Pfu DNA polymerase (Stratagene) and cloned into pYES2.1 using the TOPO TA expression kit (Invitrogen). The primer pairs used to isolate the cDNAs were 5'-ATG GTG TTG TTT TTC TCC CTA TCT-3' and 5'-TTA ATT GTG GGG ATG TAG ATG AAT-3' for CYP716A47. The PCR products were cloned into the pYES2.1/V5-His-TOPO vector and transformed into Escherichia coli. The ORFs were then ligated to the GAL1 promoter in the sense orientation. The nucleotide sequence of the inserted DNA was confirmed by sequencing. CYP716A47 and an empty vector were expressed in the Saccharomyces cerevisiae strain WAT21, which carries A. thaliana NADPH-CYP reductase (Urban et al. 1997).

WAT21 yeast cells were transformed using a modified lithium acetate procedure, as described previously (Gietz et al. 1992). Transformed cells were selected by SC-U (SC minimal medium lacking uracil) and after 3 d of growth were subcultured on YPG medium (Kribii et al. 1997). The culture conditions and methods for induction by galactose and preparation of the triterpene monoalcohol fraction have been described previously (Kushiro et al. 1998), with the exception that dammarenediol-II (50 mg I^{-1}) dissolved with ethanol and Tween-80 (1:1, v/v) was added to the medium as a substrate. After galactose induction for 1 d, the cells were collected by centrifugation at $500 \times g$ for 5 min and refluxed with 2 ml of 20% KOH/50% EtOH for 5 min. After extraction with the same volume of hexane, the extracts were analyzed by LC/APCIMS.

Co-expression of PgDDS and CYP716A47 in yeast

The coding region fragment for *PgDDS* (GenBank accession No. GU183405) was amplified by PCR using a primer set with restriction enzyme sites at the 5' end (*Notl-PgDDS-Fw* 5'-<u>GCGG</u> <u>CCGC</u> ATG TGG AAG CTG AAG GTT GCT CAA GGA-3' and *Sacl-PgDDS-Rv* 5'-<u>GAGCTC</u> TTA AAT TTT GAG CTG CTG GTG CTT AGG), cloned into the pGEM-Teasy vector and sequenced. The ORF fragments were excised through *Notl* and *Sacl* and inserted into the *Notl* and *Sacl* sites of the pESC-URA vector (Stratagene). Similarly, the coding region fragment of the CYP716A47 gene was amplified and subcloned into the pGEM-T Easy vector as an *Xhol–KpnI* fragment using the primer pairs *Xhol-CYP716A47-Fw* 5'-<u>CTCGAG</u> ATG GTG TTT TTC TCC CTA TCT and *KpnI-CYP716A47-Rv* 5'-<u>GGTACC</u> TTA ATT GTG GGG ATG TAG ATG ATG AAT-3'. The plasmid DNA was

digested by *Xhol* and *Kpnl* and then ligated into the *Xhol* and *Kpnl* sites of the pESC-URA vector including the *PgDDS* gene. The resulting plasmids were designated pESC-*PgDDS-CYP716A47-URA* and pESC-*PgDDS-URA*. WAT21 cells were transformed with these plasmids using the same method as described above.

In vitro enzymatic activity assay

To examine in vitro conversion of dammarenediol to protopanaxadiol from yeast expressing *CYP716A47*, microsomal extraction of proteins was done as described by Olsen et al. (2010) with some modifications. The microsomal fractions from WAT21 yeast cells expressing *CYP716A47* were collected by ultracentrifugation at 100,000 \times g for 60 min. The enzymatic activity of CYP716A47 was tested in a total volume of 500 µl of 100 mM potassium phosphate buffer (pH 7.4) containing 1 mM NADPH, 20 µg of substrate (dammarenediol-II or protopanaxadiol) and 1 mg of microsomal fraction protein. The reaction mixture was incubated for 2 h at 30°C, and the reaction was extracted twice with same volume of hexane.

GC/MS analysis

The hexane extract was evaporated and dissolved in methanol (1 ml). A 10 µl aliquot of the solution was analyzed by GC (Agilent 7890A) linked to an inert MSD system (Agilent 5975C) with its Triple-Axis detector, and equipped with a HP-5MS capillary column $(30 \text{ m} \times 0.25 \text{ mm}, \text{ film thickness})$ 0.25 mm). The injection temperature was 250°C, and the column temperature program was as follows: 150°C for 5 min, followed by a rise to 300° C at a rate of 5° C min⁻¹, and a hold at 300°C for 20 min. The carrier gas was He, and the flow rate was 1.2 ml min^{-1} . The interface temperature was 300° C, with a split injection (10:1). The temperature of the ionization chamber was 250°C, and ionization was by electron impact at 70 eV. The dammarenediol-II used as the standard for GC/MS analysis was provided by Professor Yong Soo Kwon of the Kangwon National University in Korea. The protopanaxadiol was obtained from Sigma-Aldrich Co.

LC/APCIMS analysis of the recombinant yeast

Yeasts (1 g) were extracted by sonication with 15 ml of 20% KOH containing 50% ethanol for 10 min. LC/APCIMS analysis was performed on a surveyor LC system (Thermo Finnigan Co.), which consisted of four solvent pumps, a Rheodyne injector (5 ml loop) and a HTP Pal autosampler (CTC Analytics). The analytical column was a YMC pack-pro C18 RS (5 mm, 2.0×150 mm, YMC Co. Ltd.) maintained at 40°C. The time of water and acetonitrile gradient application and the ratio of the constituents were as follows: 0 min, 20% acetonitrile and 80% water; 30 min, 90% acetonitrile and 10% water; 32 min, 95% acetonitrile and 5% water; 36 min, 20% acetonitrile and 80% water; and 45 min, 20% acetonitrile and 80% water at a flow rate of 0.2 ml min⁻¹. A triple quadrupole mass spectrometer Finnigan TSQ



Quantum Ultra (Thermo Electron Co.) was fitted with an atmospheric pressure chemical ionization (APCI) system, which was used for detection. The analysis was performed in the positive mode with a 5.0 mA discharge current, 320°C vaporizer temperature and 320°C ion-transfer capillary temperature. Nitrogen was used as the sheath (15 p.s.i.) and auxiliary gas (10 p.s.i.). For HPLC-UV detection, ginsenosides were monitored at a wavelength of 202 nm. Authentic dammarenediol and protopanaxadiol were subjected to the same conditions.

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the Ministry of Education, Science & Technology (MEST) [WCU project (R33-10157)]; Rural Development Administration, Republic of Korea [the Next-Generation BioGreen 21 Program (PJ008088)]; the National Research Foundation of Korea [a fellowship from the Fostering next-generation researchers program to J.-Y.H.].

Acknowledgments

The authors are grateful to Mrs. Kang JY (Central Laboratory, Kangwon National University) for the LC/APCIMS analysis.

References

- Bae, E.A., Han, M.J., Kim, E.J. and 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.
- Banno, N., Akihisa, T., Tokuda, H., Yasukawa, K., Higashihara, H., Ukiya, M. et al. (2004) Triterpene acids from the leaves of *Perilla frutescens* and their anti-inflammatory and antitumor-promoting effects. *Biosci. Biotechnol. Biochem.* 68: 85–90.
- Briskin, D.P. (2000) Medicinal plants and phytomedicines. Linking plant biochemistry and physiology to human health. *Plant Physiol.* 124: 507–514.
- Carelli, M., Biazzi, E., Panara, F., Tava, A., Scaramelli, L., Porceddu, A. et al. (2011) *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. *Plant Cell* 23: 3070–3081.
- Chen, S., Luo, H., Li, Y., Sun, Y., Wu, Q., Niu, Y. et al. (2011) 454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in *Panax ginseng. Plant Cell Rep.* 30: 1593–1601.
- Cheong, J.J. and Choi, Y.D. (2003) Methyl jasmonate as a vital substance in plants. *Trends Genet.* 19: 409-413.
- Choi, D.W., Jung, J.D., Ha, Y.I., Park, H.W., In, D.S., Chung, H.J. et al. (2005) Analysis of transcripts in methyl jasmonate-treated ginseng hairy roots to identify genes involved in the biosynthesis of

ginsenosides and other secondary metabolites. *Plant Cell Rep.* 23: 557–566.

- Cui, J.F., Eneroth, P. and Bruhn, J.G. (1999) *Gynostemma pentaphyllum*: identification of majors sapogenins and differentiation from *Panax* species. *Eur. J. Pharm. Sci.* 8: 187–191.
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.
- Gietz, D., St Jean, A., Woods, R.A. and Schiestl, R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* 20: 1425.
- Gundlach, H., Müller, M.J., Kutchan, T.M. and Zenk, M.H. (1992) Jasmonic acid is a signal transducer in elicitor-induced plant cell cultures. *Proc. Natl Acad. Sci. USA* 89: 2389–2393.
- Han, J.Y., Kwon, Y.S., Yang, D.C., Jung, Y.R. and Choi, Y.E. (2006)
 Expression and RNA interference-induced silencing of the dammarenediol synthase gene in *Panax ginseng*. *Plant Cell Physiol*. 47: 1653–1662.
- Hostettmann, K.A. and Marston, A. (1995) *In* Saponins. Chemistry and Pharmacology of Natural Products. Cambridge University Press, Cambridge.
- Irmler, S., Schröder, G., St-Pierre, B., Crouch, N.P., Hotze, M., Schmidt, J. et al. (2000) Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 *CYP72A1* as secologanin synthase. *Plant J.* 24: 797–804.
- Jia, W., Yan, H., Bu, X., Liu, G. and Zhao, Y. (2004) Aglycone protopanaxadiol, a ginseng saponin, inhibits P-glycoprotein and sensitizes chemotherapy drugs on multidrug resistant cancer cells. J. Clin. Oncol. ASCO Annual Meeting Proceedings (Post-Meeting Edition). 22(Suppl.): 9663.
- Jung, J.D., Park, H.W., Hahn, Y., Hur, C.G., In, D.S., Chung, H.J. et al. (2003) Discovery of genes for ginsenoside biosynthesis by analysis of ginseng expressed sequence tags. *Plant Cell Rep.* 22: 224–230.
- Kribii, R., Arró, M., Del Arco, A., González, V., Balcells, L., Delourme, D. et al. (1997) Cloning and characterization of the Arabidopsis thaliana SQS1 gene encoding squalene synthase–involvement of the C-terminal region of the enzyme in the channeling of squalene through the sterol pathway. *Eur. J. Biochem.* 249: 61–69.
- Kushiro, T., Ohno, Y., Shibuya, Y. and Ebizuka, Y. (1997) In vitro conversion of 2,3-oxidosqualene into dammarenediol by *Panax ginseng* microsomes. *Biol. Pharm. Bull.* 20: 292–294.
- Kushiro, T., Shibuya, M. and Ebizuka, Y. (1998) β-Amyrin synthase: cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. *Eur. J. Biochem.* 256: 238–244.
- Lee, M.H., Jeong, J.H., Seo, J.W., Shin, C.G., Kim, Y.S., In, J.G. et al. (2004) Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* overexpressing squalene synthase gene. *Plant Cell Physiol.* 45: 976–984.
- Li, G., Wang, Z., Sun, Y., Liu, K. and Wang, Z. (2006) Ginsenoside 20(S)-protopanaxadiol inhibits the proliferation and invasion of human fibrosarcoma HT1080 cells. *Basic Clin. Pharm. Toxicol.* 98: 588–592.
- Nelson, D. (2006) Plant cytochrome P450s from moss to poplar. *Phytochem. Rev.* 5: 193–204.
- Olsen, K.M., Hehn, A., Jugdé, H., Slimestad, R., Larbat, R., Bourgaud, F. et al. (2010) Identification and characterisation of CYP75A31, a new flavonoid 3'5'-hydroxylase, isolated from *Solanum lycopersicum*. *BMC Plant Biol.* 10: 21.
- Osbourn, A.E. (1996) Preformed antimicrobial compounds and plant defense against fungal attack. *Plant Cell* 8: 1821–1831.



- Page, R.D. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357-358.
- Popovich, D.G. and Kitts, D.D. (2004) Ginsenosides 20(S)protopanaxadiol and Rh2 reduce cell proliferation and increase sub-G1 cells in two cultured intestinal cell lines (Int-407 and Caco-2). *Can. J. Physiol. Pharmacol.* 82: 183–190.
- Sasaki, K., Minowa, N., Kuzuhara, H., Nishiyama, S. and Omoto, S. (1997) Synthesis and hepatoprotective effects of soyasapogenol B derivatives. *Bioorg. Med. Chem. Lett.* 7: 85–88.
- Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H. et al. (2008) Licorice beta-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc. Natl. Acad. Sci. USA* 105: 14204–14209.
- Schuler, M. (1996) Plant cytochrome P450 monooxygenases. Crit. Rev. Plant Sci. 15: 235–284.
- Shibata, S. (1977) Saponins with biological and pharmacological activity. *In* New Natural Products and Plant Drugs with Pharmacological or Therapeutical Activity. Edited by Wagner, H. and Wolff, P. pp. 177–196. Springer-Verlag, New York.
- Shibata, S. (2001) Chemistry and cancer preventing activities of ginseng saponins and some related triterpenoid compounds. J. Korean Med. Sci. 16: S28–S37.
- Shibuya, M., Hoshino, M., Katsube, Y., Hayashi, H., Kushiro, T. and Ebizuka, Y. (2006) Identification of beta-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. *FEBS J.* 273: 948–959.
- Sun, C., Li, Y., Wu, Q., Luo, H., Sun, Y., Song, J. et al. (2010) De novo sequencing and analysis of the American ginseng root transcriptome using a GS FLX Titanium platform to discover putative genes involved in ginsenoside biosynthesis. *BMC Genomics* 11: 262.

- Suzuki, H., Achnine, L., Xu, R., Matsuda, S.P.T. and Dixon, R.A. (2002) A genomics approach to the early stages of triterpene saponin biosynthesis in *Medicago truncatula*. *Plant J.* 32: 1033–1048.
- Takahashi, S., Yeo, Y.S., Zhao, Y., O'Maille, P.E., Greenhagen, B.T., Noel, J.P. et al. (2007) Functional characterization of premnaspirodiene oxygenase, a cytochrome P450 catalyzing regio- and stereo-specific hydroxylations of diverse sesquiterpene substrates. *J. Biol. Chem.* 282: 31744–31754.
- 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.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.
- Thornton, L.E., Rupasinghe, S.G., Peng, H., Schuler, M.A. and Neff, M.M. (2010) Arabidopsis CYP72C1 is an atypical cytochrome P450 that inactivates brassinosteroids. *Plant Mol. Biol.* 74: 167–181.
- Urban, P., Mignotte, C., Kazmaier, M., Delorme, F. and 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.
- Vogler, B.K., Pittler, M.H. and Ernst, E. (1999) The efficacy of ginseng. A systematic review of randomised clinical trials. *Eur. J. Clin. Pharmacol.* 55: 567–575.