Bangladesh J. Bot. **44**(5): 897-908, 2015 (December), Supplementary

CLONING, CHARACTERIZATION AND BACTERIAL EXPRESSION OF THE β-AMYRIN SYNTHASE GENE FROM *PANAX JAPONICUS* C. A. MEY

ZHIWEI HUANG[1], LIXING WANG[1], MINGSHU GUO, SHIQIANG LIN, ZUXIN CHENG AND JINGUI ZHENG*

Agricultural Product Quality Institute, Fujian Agriculture and Forestry University, Fuzhou 350002, China

Key words: *Panax japonicus* C. A. Mey, Ginsenosides, *β*-Amyrin synthase, Gene cloning, Sequence analysis, Prokaryotic expression

Abstract

Panax japonicus C. A. Mey. is one of the rare traditional Chinese herbal medicines; its main active ingredient is ginsenosides. *β*-amyrin synthase (*βAS*) is the rate-limiting enzyme in the biosynthesis pathway of oleanane-type ginsenosides. We cloned for the first time the full-length cDNA of *βAS* from *P. japonicus* (GenBank: KP658156) using reverse transcription-polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE), and we named it *PjβAS*, which contained an open reading frame (ORF) of 2,286 bp in length and encoded a protein of 761-amino acids in length and about 87.8 kDa in molecular weight. Bioinformatics analysis showed that *PjβAS* had more than 80% homology with *βAS* from *P. ginseng*, *Aralia elata, Betula platyphylla,* and *Malusx domestica*, etc. *PjβAS* consists of six functional domains of the terpene synthase family and the functional domain of the prenyltransferase/squalene oxidase family. The coding sequence of *PjβAS* was cloned into a prokaryotic expression vector pET-41a(+); and the recombinant vector pET-AS was transformed into the *Escherichia coli* BL21(DE3) strain. Real-time fluorescence quantitative PCR (qRT-PCR) and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis showed that *PjβAS* could be over-expressed by IPTG-induction in BL21(DE3). High performance liquid chromatography (HPLC) analysis showed that pET-AS could synthesize *β*-amyrin in BL21 (DE3) cells. These results indicated that *PjβAS* functions as a *β*AS gene.

Introduction

Panax japonicus C.A. Mey., belonging to Araliaceae *Panax*, is one of the rare traditional Chinese herbal medicines. All plants of the *Panax* genus have high medicinal value and their main active ingredient is ginseng saponins (Yun 2001). The total saponin content of the roots of *P. ginseng* is 2 - 7%, whereas the total saponin content in the roots of *P. japonicus* can be as high as 15%, which is 2 to seven-fold higher than that observed in *P. ginseng* and three-fold higher than reported in *P. quinquefolius*. Ginsenosides are further classified into the oleanane type and the dammarane type; and the oleanane-type saponins are the main saponins in *P. japonicus* (Haralampidis *et al.* 2002, Lichtenthaler *et al.* 1997).

Ginsenosides are triterpenoid saponins of plant secondary metabolites, i.e. the product of the triterpenoid saponin biosynthesis branch in the isoprenoid pathway (Fig. 1). Triterpenoid saponins are formed by different cyclizations of squalene (Haralampidis *et al.* 2002; Lichtenthaler *et al.* 1997). There are two pathways for the synthesis of isopentenyl pyrophosphate (IPP) in the isoprenoid pathway, one is mevalonate (MVA) pathway (Chappell 1995, Goldstein and Brown 1990), and the other is the 1-deoxy-D-xylulose-5-phosphate (DOXP) pathway, also known as the non-mevalonate pathway (Lichtenthaler 1999, Wanke *et al.* 2001). IPP in the triterpenoid biosynthesis pathway comes from the MVA pathway, and IPP can further synthesize squalene.

*Author for correspondence: <zjg231@126.com>. [1]College of Food Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China.

Squalene epoxidase (SQE) catalyzes linear squalene to form cyclic 2, 3-oxidosqualene. 2,3-oxidosqualene cyclases (OSCs) are rate-limiting enzymes in the triterpenoid biosynthesis pathway, which catalyzes the further cyclization of 2,3-oxidosqualene to form the triterpenoid skeleton, which then can be further modified to form a variety of triterpenoid saponins (Haralampidis *et al.* 2002, Lichtenthaler *et al.* 1997). OSCs of ginseng plants mainly consist of *β*-amyrin synthase (*βAS*) and dammarenediol-II synthase (DS), which catalyze 2, 3-oxidosqualene to produce the respective oleanane-type and dammarane-type saponins, *β*-amyrin and dammarenediol-II (Haralampidis *et al.* 2002, Liang and Zhao 2008, Lichtenthaler *et al.* 1997).

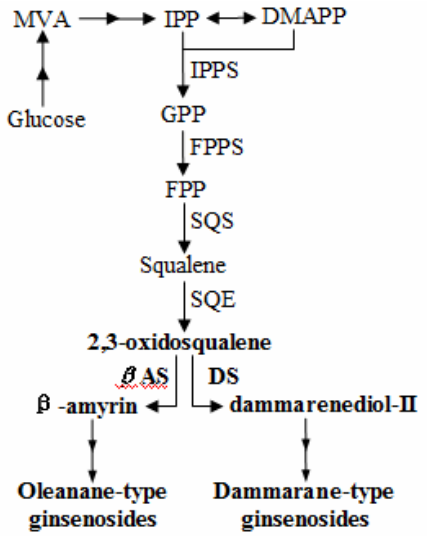

Fig. 1. Schematic of the biosynthetic pathway of ginsenosides (Haralampidis *et al*. 2002, Liang and Zhao 2008, Lichtenthaler *et al.* 1997). Intermediates: MVA, mevalonic acid; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate. Enzymes: IPPS, isoprenyl diphosphate synthase; FPPS, famesyl diphosphate synthase; SQS, squalene synthase; SQE, squalene epoxidase; DS, dammarenediol-II synthase; *βAS*, *β*-amyrin synthase.

As a perennial herb, Ginseng requires long period of cultivation, and its cultivation and growing conditions are special. It is sensitive to soil and the climate, and its quality can be affected by continuous cropping. All these difficulties indirectly hinder the large-scale application of ginseng. Thus, the isolation and cloning of *βAS* from *P. japonicus* is of important theoretical significance and application value in conducting studies on its metabolic regulation of ginsenosides synthesis and development of ginsenosides medicine source.

Zhao *et al.* (2011) studied the expression of ginseng *βAS* and the regulation of ginsenosides biosynthesis using the antisense RNA technology. When antisense *βAS* was introduced into the hairy roots of the ginseng plant, the transcription levels of the *βAS* in the transgenic hairy roots significantly decreased. In addition, *β*AS activity also decreased, and the content of oleanane-type ginsenosides Ro was reduced by up to 40%. On the other hand, the DS activity of these *βAS* antisense lines increased, and the content of dammarane-type ginsenosides increased by up to 30%. These findings indicated that regulation of the synthesis of ginsenosides can be achieved by altering the expression of *βAS* using genetic engineering techniques.

βAS has previously been cloned from *P. ginseng* (GenBank: AB014057.1, AB009030.1) (Haralampidis *et al.* 2002), *Pisum sativum* (GenBank: AB034802.1) (Haralampidis *et al.* 2002), *Betula platyphylla* (GenBank: AB055512.1), *Centella asiatica* (GenBank: AY520818.1), *Malus domestica* (GenBank: FJ032007.1), and *Aralia elata* (GenBank: HM219225.1). In addition, some researchers have conducted bioinformatics analysis on the cloned *βAS* and performed heterologous expression assays to verify their function (Chen *et al.* 2013, Haralampidis *et al.* 2001, Suzuki *et al.* 2002).

We designed primers based on the conserved regions of reported nucleotide sequences of *βAS* from other plant species for use in a combined reverse transcription-polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) assay (Frohman *et al.* 1988, Chenchik *et al.* 1996, Etienne *et al.* 2000). We cloned the full-length cDNA of *βAS* from *P. japonicus*, conducted sequence analysis and functional prediction and expressed the *βAS* in *E. coli* to verify its function.

Materials and Methods

P. japonicus was collected from the Jinggangshan National Nature Reserve in Jiangxi province, China. The roots of three-year-old plants were collected, immediately frozen in liquid nitrogen, and stored in a –80°C freezer. TRIzol (Invitrogen, 15596026, USA) was used to extract total RNA from the roots of *P. japonicus*, and its integrity was determined by agarose gel electrophoresis. A ND-2000C Ultramicro UV spectrophotometer (Thermo Fisher Scientific, USA) was used to determine the RNA concentration and $OD_{260/280}$ values of the samples. The extracted RNA is stored at –80°C.

Based on the results of multiple sequence alignment of *βAS* cDNA from *P. ginseng* and other plants, we designed specific primers, namely ASRT1-1, ASRT2-1 and ASRT2-2 (Table 1) using the software Vector NTI 7.0. RevertAid™ First Strand cDNA Synthesis Kit (Fermentas, K1621, Canada) was used to reversely transcribe total RNA of *P. japonicus* into first strand cDNA, and PCR was conducted to amplify the conserved fragments of *βAS* using *TaKaRa Taq*™ (Takara, DR001A, China). The PCR reaction system consisted of a 50 µl volume containing 5.0 µl 10 × PCR Buffer (Mg^{2+} Plus), 4.0 µl dNTP mixture (2.5 mM for each nucleotide), 1.0 µl 20 µM primer ASRT1-1, 1.0 µl 20 µM primer ASRT2-1, 2.0 µl first-strand cDNA (diluted 10-fold), 0.5 µl *TaKaRa Taq* (5U/µl), and 36.5 µl ddH_2O. The PCR reaction conditions were as follows: 94°C for 5 min; 94°C for 30 s, 58°C for 30 s, 72°C for 60 s/kb, for 35 cycles; and 72°C for 10 min. After the PCR products were recovered by using the EZNA® Gel Extraction Kit (Omega Bio-Tek, D2500-02, USA), these were ligated to a pMD® 18-T Simple Vector (Takara, D103A), transformed into *E. coli* DH5α cells, and the positive clones were picked up for sequencing (Sambrook *et al.* 2001). Homology analysis of the sequencing results was performed using BLAST (http://www.ncbi.nlm.nih.gov). Based on the obtained conserved sequences, the specific forward primer, ASF4, was designed (Table 1), and PCR was conducted to amplify the downstream conserved region of *βAS* using the first strand cDNA as template and ASF4/ASRT2-2 as primers. The recovery, ligation, transformation, and sequencing of PCR products were based on the methods above.

Based on the conserved sequence of *βAS* obtained by RT-PCR analysis, the specific primers ASF5, ASF6 and ASR4, ASR5 were designed for 3'-RACE and 5'-RACE (Table 1). SMART ™ RACE cDNA Amplification Kit (Clontech, 634923, USA) was used to reversely transcribe total RNA into the first strand cDNA for 3'-RACE and 5'-RACE. Following the PCR reaction system and reaction conditions described by the manufacturer, we performed 3'-terminus amplification and 5'-terminus amplification of *βAS*, and the primers used in the first round of PCR in 3'-RACE were ASF5/UPM, and the nested PCR primers were ASF6/NUP. The primers used in the first round of PCR in 5'-RACE were UPM/ASR4, and the nested PCR primers were NUP/ASR5. The methods

for the recovery, ligation, transformation, and sequencing of PCR products were similar to those employed in RT-PCR analysis.

Table 1. List of primers used in the study.

Primers	Sequence (5′ − 3′)	Direction
RT-PCR		
ASRT1-1	GGAAGRCAGACATGGGAGTTTG	Forward
ASRT2-1	TCHACCCAACAAGCAAGCATAC	Reverse
ASF4	TATTTACGGAGCCTTTCTTGACT	Forward
ASRT2-2	TCTCTTTCTTCCTATGCCCTGG	Reverse
RACE		
ASF5	ATGCTTGCTTGTTGGGTTGAGG	Forward
ASF6	AGATCATGGATGGCAAGTTTCG	Forward
UPM*	CTAATACGACTCACTATAGGGC	
NUP*	AAGCAGTGGTATCAACGCAGAGT	
ASR4	CCCCACAAACCTCTTCCCATAC	Reverse
ASR5	GAAGGAAAGGAGGAAGGAT	Reverse
Coding region		
βASF	TTGAAGATGTGGAGGCTAATGA	Forward
βASR2	CATTTGAGTATTGGCTGACCGT	Reverse
Expression		
ASF*Sac* I	GAGCTCATGTGGAGGCTAATGACAGCCA	Forward
ASR*Not* I	GCGGCCGCTGTTCAGACGCTTTTAGGTGGT	Reverse
qRT-PCR		
qASF	TGCCAGAGCAAGAAAA TGGA	Forward
qASR	CATAGGAAGGAAAGGAGGAAGGA	Reverse
qACTF	CATCTTGGCATCTCTCAGCAC	Forward
qACTR	AACTTTGTCCACGCTAATGAA	Reverse

*Primers marked with asterisk were derived from kit. Restriction sites that were introduced into the pET-41a(+) vector to facilitate cloning are underlined.

We obtained the full-length cDNA sequence of *βAS* by splicing the sequences obtained by RT-PCR, 3'-RACE, and 5'-RACE. Specific forward and reverse primers βASF/βASR2 were designed (Table 1), and PCR was performed to amplify the coding region of *βAS* using the first strand cDNA as the template and PrimeSTAR® HS DNA Polymerase (Takara, DR010A). The PCR reaction system and reaction conditions were set up according to the instruction provided in the kit. The methods for the recovery, ligation, transformation, and sequencing of PCR products were similar to those employed in RT-PCR analysis.

Homology analysis of the full-length cDNA sequence and the deduced amino acid sequence of *βAS* was conducted using BLAST. Using the multiple sequence alignment software CLUSTALX 2.1, we aligned the amino acid sequence of *P. japonicus βAS* with those of the *Panax* genus plants, etc. reported in the GenBank and other databases and constructed a phylogenetic tree to determine the genetic relationships of *βAS* among different plants. By searching Conserved Domain Database (CDD), Conserved Domain Architecture Retrieval Tool (CDART), Blocks (http://blocks.fhcrc.org), and other databases, we analyzed the structural features and functional domains of *P. japonicus βAS* in order to predict its function.

Forward and reverse primers ASFSac I and ASRNot I (Table 1), which were specific to the termini of the open reading frame (ORF) of *βAS*, were designed, and PCR was carried out using the first strand cDNA as template and PrimeSTAR® HS DNA polymerase to amplify the ORF of *βAS*. The PCR products were ligated to a pMD 18-T Simple Vector, resulting in plasmid pMD-AS, which was sequenced to confirm the ORF of *βAS*. The ORF of *βAS* was then cloned into the prokaryotic expression vector pET-41a (+) (Merck KGaA, 70556-3, Germany), and a *βAS* prokaryotic expression vector was obtained. The recombinant plasmid pET-AS was transformed into *E. coli* BL21 (DE3) cells, and after enzyme digestion and PCR validation, the recombinant bacteria BL21/pET-AS carrying plasmid pET-AS was obtained (Sambrook *et al.* 2001). The recombinant strain BL21/pET-AS was inoculated into LB liquid medium and cultured by shaking at 37°C until its OD_{600} reached 0.6. After a non-induced sample was removed as control, IPTG was added to the remaining culture at a final concentration of 0.5 mmol/l. After 2 hrs or 4 hrs of induction, a 1 ml aliquot of the bacteria was isolated and the cells were collected from the suspension by centrifugation at 1,500 g for 10 min. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyze the expression products (Sambrook *et al.* 2001).

TRIzol was used to extract total RNA from samples of the uninduced, 2 hrs or 4 hrs induced recombinant bacteria BL21/pET-AS (i.e. CK, A2 or A4), and the DNase enzyme (Takara, D2210) was used to remove residual DNA from the RNA samples. PrimeScript® RT reagent kit (Takara, DRR037A) was used to reversely transcribe total RNA into the first-strand cDNA. The conditions for reverse transcription reaction were as follows: 37°C for 15 min and 85°C for 5 s. SYBR® Premix Ex Taq™ II (Takara, RR041A) and ABI 7500 Real-Time PCR System (Applied Biosystems, USA) were used to conduct real-time fluorescence quantitative PCR (qRT-PCR) analysis of *βAS* in the recombinant strain, and *ACTIN* was used as the internal reference gene. The primers used were qASF/qASR and qACTF/qACTR (Table 1). The PCR reaction system was a 20-μl solution that included 10.0 μl of SYBR® Premix Ex Taq™ II (2×), 0.4 μl of each forward and reverse primer (20 μM), 0.4 μl of ROX Reference Dye II (50×), 4.0 μl of the first strand cDNA (diluted 5 times), and 4.8 μL of ddH_2O. PCR reaction conditions were as follows: 95°C for 30 s; 95°C for 5 s, 60°C for 34 s, for a total of 40 cycles. The relative expression of *βAS* was calculated using the equation: $RQ = 2^{-\Delta\Delta Ct}$. Five replicates were analyzed for each sample (Huang *et al.* 2015; Huang *et al.* 2015).

Each 10 ml of bacteria from samples of the uninduced, 2 hrs or 4 hrs induced recombinant bacteria BL21/pET-AS (i.e. CK, A2 or A4) was harvested and collected by centrifugation at 1,2000 g for 1 min. Methanol was used as solvent to reflux extract *β*-amyrin for 10 hrs, and a 1525-type High Performance Liquid Chromatography instrument (Waters, USA) was used to determine *β*-amyrin content. The column employed was a Hypersil ODS2 C18 column (250 nm × 4.6 nm, 5 μm, Dalian Elite Company, China), and the chromatographic conditions were as follows: the mobile phase was methanol-0.05 mol/l NaH_2PO_4 (85 :15), the flow rate was 1.0 ml/min, the column temperature was 25°C, and the detection wavelength was 210 nm. The standard, *β*-amyrin (CAS: 559-70-6, Lot #: 0001441054), was purchased from Sigma (USA).

Results and Discussion

The total RNA of *P. japonicus* showed three clear bands on electrophoresis, the ratio of brightness of the 23S and 16S bands was about 2 : 1. These results indicated that the total RNA we obtained was intact and could therefore be used for RT-PCR and RACE experiments. The concentration of total RNA was 2261.5 ng/μL and $OD_{260/280}$ was 2.13.

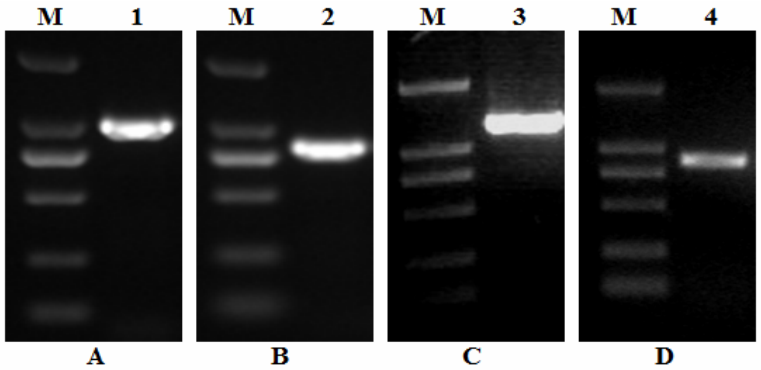

Fig. S1. RT-PCR and RACE PCR products of *βAS* from *P. japonicus*. (M), DNA Marker DL2000.
(1, 2), RT-PCR product of *βAS*. (3), 3'-RACE PCR product of *βAS*. (4), 5'-RACE PCR product of *βAS*.

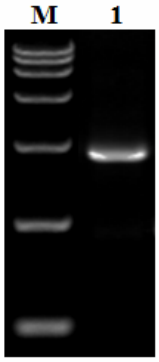

Fig. S2. Coding region PCR product of *βAS* from *P. japonicus*. (M), DNA Marker DL15000.
(1), coding region of PCR product of *βAS*.

After two runs of RT-PCR, we obtained two conserved fragments of 1,079 bp and 771 bp in size of the *βAS* (Fig. S1-A and B). By 3'-RACE and 5'-RACE, a 3' end fragment 1,167 bp in size (Fig. S1-C) and a 5' end fragment 846 bp in size (Fig. S1-D) were obtained, respectively. BLAST analysis showed that the sequences of these four fragments had >80% homology with the reported *βAS* sequences from *P. ginseng*, *B. platyphylla*,, and *G. glabra*. We spliced the two conserved sequences, the 3'-terminal sequence and the 5'- terminal sequence, and obtained the full-length cDNA sequences of the *βAS*. Using the first strand cDNA as template and βASF/βASR2 as primers, we amplified the coding region of *βAS*, which generated a fragment with a size of 2,354 bp (Fig. S2). This size of the fragment was consistent with that obtained by splicing. BLAST analysis showed that the homology between this coding sequence and OSCPNY2 (*βAS2*) of *P. ginseng*, *βAS* of *A. elata*, OSCPNY1 (*βAS1*) of *P. ginseng*, and OSCBPY of *B. platyphylla* were 99, 83, 82, and

81%, respectively. The full-length cDNA of *P. japonicus βAS* was thereby named *PjβAS* (2,715 bp in size, GenBank: KP658156), which consisted of an ORF of 2,286 bp in size (nucleotide positions 99 - 2,384); a fragment from positions 165 - 1,854 (1,690 bp) was obtained by RT-PCR, a fragment from positions 1,526 - 2,715 (1,190 bp) was obtained by 3'-RACE, and a fragment from positions 1 - 846 (846 bp) was obtained by 5'-RACE. The deduced amino acid sequence of the *PjβAS* coding sequence (Fig. 2) has a total of 761 amino acid residues, its molecular weight was 8,7782 Da, and its isoelectric point was 5.85.

```
1   MWRLMTAKGG NDLYLYSTNN FIGRQTWEFD PDYGTPAERA EVEEARLHFW NNRYQVKPSG
61  DVLWRMQFLK EKNFKQIIPQ VKVEDGEEIS YEAATTTLRR AVHYFSALQA DDGHWPAENA
121 GPLFFLPPLV MCLYITGHLN TVFPAEYRIE ILRYIYCHQN EDGGWGLHIE GHSTMFCTAL
181 SYICMRILGE GRDGGENNAC ARARKWILDH GSVTAIPSWG KTWLSILGLF DWSGSNPMPP
241 EFWILPPFLP MHPAKMWCYC RMVYMPMSYL YGKRFVGPIT PLILQLREEL YAQAYDEINW
301 RKVRHNCAKE DLYYPHPLIQ DIMWDSLYIF TEPFLTRWPF NKLREKALQT TMKHIHYEDE
361 NSRYITIGCV EKVLCMLACW VEDPNGDYFK KHLARIPDYI WVAEDGMKMQ SFGSQEWDTG
421 FAIQALLASD LTDEIRPTLM KGHDFIKKSQ VKENPSGDFK SMHRHISKGS WTFSDQDHGW
481 QVSDCTAEAL KCCLLFSRMP TEIVGDKMED SQLFDAVNIL LSLQSKNGGL AAWEPAGSSE
541 WLELLNPTEF FEDIVIEHEY VECTSSAIQA MVMFKKLYPG HRKKEIEVSI TNAVQYLEDI
601 QMPDGSWYGN WGVCFTYGTW FAMGGLTAAG KTYNNSQTLH KAVDFLIKWQ RSDGGWGESY
661 LSCPNKEYTP LEGNRSNLVH TSWAMMGLIH SGQAERDPTP LHRAAKLLIN SQMESGDFPQ
721 QEITGVFMKN CMLHYAAYRN IYPLWALAEY RKNVRLPPKS V
```

Fig. 2. The deduced amino acid sequence of *PjβAS* from *P. japonicus*. The colored background area is the six functional domains of the terpene synthase family, the boxed area is the functional domain of the prenyltransferase/squalene oxidase family.

BLAST analysis showed that *PjβAS* had >80% homology with *βAS* from *P. ginseng, A. elata, B. platyphylla, M. domestica*, and other plants at the nucleotide and amino acid sequences levels.

Although different OSCs catalyze the same substrate, 2, 3-oxidosqualene, different types of products were produced. For example, OSCs in *P. japonicus* are mainly *β*AS and DS; they catalyze 2,3-oxidosqualene to synthesize the respective oleanane-type and dammarane-type ginsenosides substrates, *β*-amyrin and dammarenediol-II, whereas rice OSC is a cycloartenol synthase (CS), which catalyzes 2,3-oxidosqualene to produce cycloartenol (Augustin *et al.* 2011, Haralampidis et al. 2001, Haralampidis *et al.* 2002, Suzuki *et al.* 2002). However, the amino acid sequences or the coding sequences of different types of OSCs are highly homologous (Haralampidis *et al.* 2002, Hayashi *et al.* 2000, Kushiro *et al.* 1998). BLAST analysis indicated that the homology between *PjβAS* and the lupeol synthase (*LUS*) gene of *Euphorbia tirucalli*, *CS* of *Ricinus communis*, *LUS* of *Bruguiera gymnorrhiza*, and *LUS2* of *Arabidopsis thaliana* were 81, 75, 75, and 71%, respectively. The results were consistent to those of previous studies (Haralampidis *et al.* 2001, Tansakul *et al.* 2006).

Multiple sequence alignment of the amino acid sequences of different *β*AS (Fig. S3) showed that the amino acid sequence of *PjβAS* had high homology with those of reported *βAS* genes of the *Panax* genus plants, etc. The phylogenetic tree of *βAS* from different plant species is shown Fig. 3, which shows that Pj*βAS* was closely related to BAMS2_PANGI (Beta-Amyrin Synthase 2, Genbank: O82146.1) of *P. ginseng*.

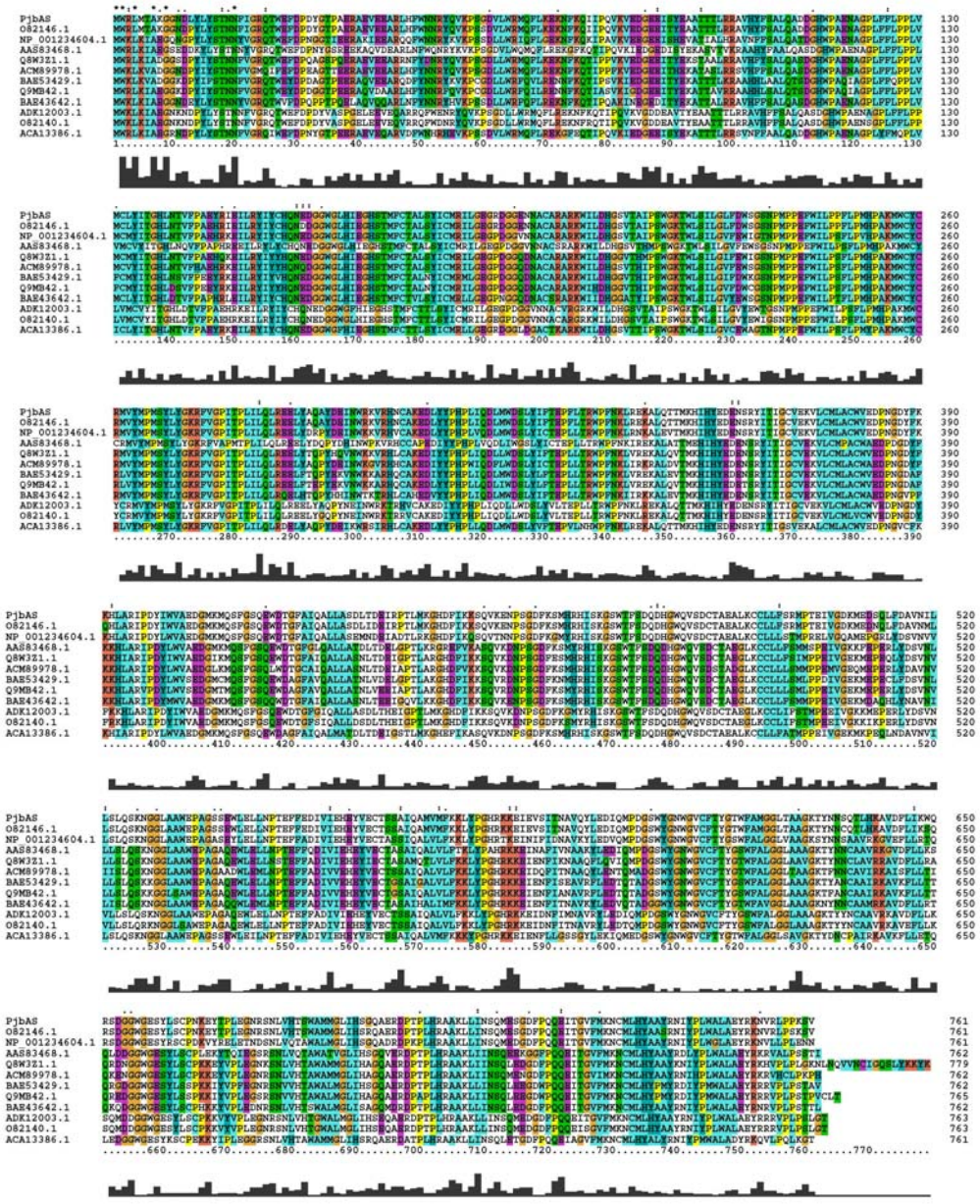

Fig. S3. Multiple sequence alignment of the deduced amino acid sequences of *PjβAS* with its homologs using CLUSTALX 2.1. Intermediates: PjbAS, *P. japonicus* beta-amyrin synthase; O82146.1, *P. ginseng* beta-amyrin synthase 2; NP_001234604.1, *Solanum lycopersicum* beta-amyrin synthase; AAS83468.1, *Bupleurum kaoi* beta-armyrin synthase; Q8W3Z1.1, *Betula platyphylla* beta-amyrin synthase; ACM89978.1, *Malusx domestica* putative beta-amyrin synthase; BAE53429.1, *Lotus japonicus* beta-amyrin synthase; Q9MB42.1, *Glycyrrhiza glabra* beta-amyrin synthase; BAE43642.1, *Euphorbia tirucalli* beta-amyrin synthase; ADK12003.1, *Aralia elata* beta-amyrin synthase; O82140.1, *P. ginseng* beta-amyrin synthase 1; ACA13386.1, *Artemisia annua* beta-amyrin synthase.

BLAST analysis showed that the homology between *PjβAS* and *OSCPNY2* of *P. ginseng* (*βAS*2, Genbank: AB014057.1) at the nucleotide level and amino acid level were up to 96% and 98%, respectively, whereas the homology between *PjβAS* and *OSCPNY1* of *P. ginseng* (*βAS*1, Genbank: AB009030.1) was only 82 and 85%. Multiple sequence alignments and phylogenetic trees indicated that the kinship between Pj*β*AS and OSCPNY2 of *P. ginseng* (beta-amyrin synthase 2, UniProtKB/Swiss-Prot: O82146.1) was closer than that between Pj*β*AS and OSCPNY1 of *P. ginseng* (beta-amyrin synthase 1, UniProtKB/Swiss-Prot: O82140.1). Therefore, the *PjβAS* obtained in the present study belongs to beta-amyrin synthase 1 gene.

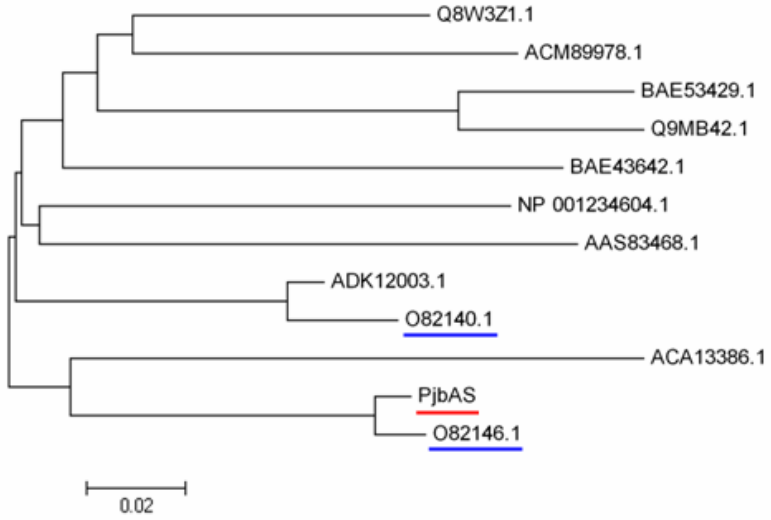

Fig. 3. Phylogenetic analysis of the deduced amino acid sequences of *PjβAS* using the CLUSTALX 2.1.

Sequence analysis also showed that the homology between the cDNA of the gene (Genbank: AB122080.1, beta-amyrin synthase gene) with those of the *DS* of *Panax* genus plants was as high as 98 - 99%, whereas its homology with *βAS* was < 80%. Therefore, we inferred that AB122080.1 was not *βAS*, but *DS*, and its function requires further experimental verification.

CDD and CDART analysis showed that *PjβAS* harbored functional domains of class II terpene cyclases and β subunits of protein prenyltransferases. 1. Subfamily domain of squalene cyclases in class II terpene cyclases (with α6 - α6 spiral fold): eukaryotes OSCs contain the complete transmembrane protein structure and can catalyze the cyclization cascade of cations, converting linear triterpenoids to polycyclic compounds. Plant OSCs (e.g., CS and *βAS*) can catalyze 2, 3-oxidosqualene to form cycloartenol and *β*-amyrin. These enzymes have a secondary domain, which can be inserted into the α-α helix of the main structure. 2. β subunits of protein prenyltransferases, which include farnesyl transferase, and types I and II geranylgeranyltransferase; these enzymes catalyze lipidation of the carboxyl termini of Ras, Rab, and other cellular signal transduction proteins, and promotes the combination of membranes and specific interactions between proteins.

Blocks analysis showed that *PjβAS* has six functional domains of the terpene synthase family and the functional domain of prenyltransferase/squalene oxidase family (Fig. 2). The six functional domains of the terpene synthase family were as follows: the first domain covered amino acids 150 - 198, with a total length of 49 amino acids; the second domain included amino acids 202 - 248, and showed a total length of 47 amino acids; the third domain included amino acids 467 - 492, with a

total length of 26 amino acids; the fourth domain of 604 - 620, a total of 17 amino acids; the fifth domain consisted of amino acids 676 - 689, with a total length of 14 amino acids; the sixth domain included amino acids 725 - 750, and presents a total length of 26 amino acids. The functional domain of the prenyltransferase/squalene oxidase family consisted of amino acids 605 - 618 and a total length of14 amino acids; this domain overlapped with the fourth domain of the terpene synthase family.

In summary, *PjβAS* has functional domains of the terpene synthase family and prenyltransferase/squalene oxidase family. *PjβAS* also has high homology with those of reported *βAS* protein sequences of the *Panax* genus plants, etc. thus we speculated that *PjβAS* functions as a *β*-amyrin synthase.

The ORF of *PjβAS*, whose both ends were added appropriate restriction sites, was cloned into the expression vector pET-41a (+) after a double digestion (Fig. S4) and verification by sequencing to generate a *PjβAS* prokaryotic expression vector, pET-AS. The recombinant vector pET-AS was transformed into *E. coli* BL21 (DE3) cells after restriction enzyme digestion and PCR validation to produce the recombinant bacteria BL21/pET-AS (Sambrook *et al*. 2001). SDS-PAGE analysis showed that when induced by IPTG, the recombinant strain BL21/pET-AS can successfully express the fusion protein with the expected size of approximately 88 kDa (Fig. 4).

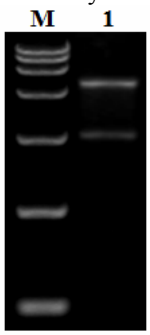

Fig. S4. Identification of recombinant expression vector pET-AS by enzyme digestion.
(M), DNA Marker DL15000. (1) Digested by *Sac* I and *Not* I.

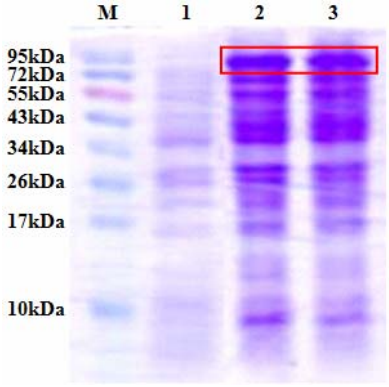

Fig. 4. SDS-PAGE analysis of the expressed products of recombinant vector pET-AS.
(M), protein Marker; (1), uninduced; (2, 3), induced for 4 h and 2h, respectively.

qRT-PCR analysis showed that when induced by IPTG, *PjβAS* could be transcribed in large amounts in the strain; its relative expression levels were 4,613.9 and 5,914.9 after 2 hrs or 4 hrs of

IPTG induction (Fig. 5). Using *β*-amyrin as a standard, we assayed the *β*-amyrin content of recombinant strain BL21/pET-AS by HPLC (Fig. S5). The results showed that after 2 hrs or 4 hrs of IPTG induction, the content of *β*-amyrin in the recombinant bacteria was 3.8 mg/ml and 5.1 mg/ml, respectively (Fig. 6).

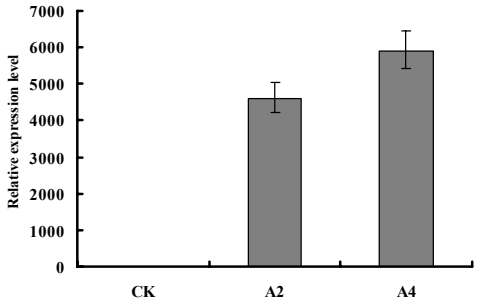

Fig. 5. qRT-PCR analysis of *PjβAS* expression in recombinant *E. coli* BL21/pET-AS. CK: uninduced; A2, A4, induced for 4hRS, 2hRS, respectively.

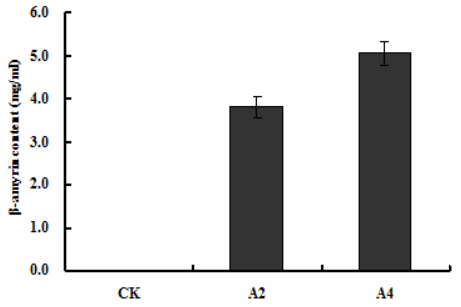

Fig. 6. Production of *β*-amyrin as analyzed by HPLC. CK: uninduced; A2, A4, induced for 4, 2hrs, respectively

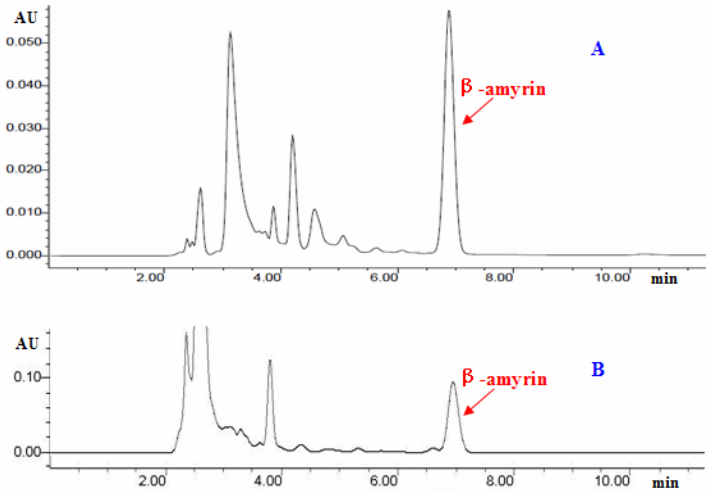

Fig. S5. HPLC chromatograms of *β*-amyrin. (A) *β*-amyrin standard. (B) HPLC chromatograms of *β*-amyrin in recombinant *E. coli*.

Acknowledgements

The National Natural Science Fund (30671276), the National Transgenic Organism New Variety Culture Key Project (2009ZX08001-032B), the National Key Technology R&D Program (2013BAD01B05) of China financially supported this research study, and Fujian Agriculture and Forestry University Project of Constructing High-level University (612014042).

References

Augustin JM, Kuzina V, Andersen SB and Bak S 2011. Molecular activities, biosynthesis and evolution of triterpenoid saponins. Phytochemi. **72**(6): 435-457.

Chappell J 1995. Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. **46**: 521-547.

Chen HH, Liu Y, Zhang XQ, Zhan XJ and Liu CS 2013. Cloning and characterization of the gene encoding β-amyrin synthase in the glycyrrhizic acid biosynthetic pathway in *Glycyrrhiza uralensis*. Acta Pharmaceutica Sin. B **3**(6): 416-424.

Chenchik A, Diachenko L, Moqadam F, Tarabykin V, Lukyanov S and Siebert PD 1996. Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA. Biotechniques **21**(3): 526-534.

Etienne P, Petitot AS, Houot V, Blein JP and Suty L 2000. Induction of tcI 7, a gene encoding a beta-subunit of proteasome, in tobacco plants treated with elicitins, salicylic acid or hydrogen peroxide. Febs. Lett. **466**(2-3): 213-218.

Frohman MA, Dush MK and Martin GR 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. P. Natl. Acad. Sci. USA **85**(23): 8998-9002.

Goldstein JL and Brown MS 1990. Regulation of the mevalonate pathway. Nature **343**(6257): 425-430.

Haralampidis K, Bryan G, Qi X, Papadopoulou K, Bakht S, Melton R and Osbourn A 2001. A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots. P. Natl. Acad. Sci. USA **98**(23): 13431-13436.

Haralampidis K, Trojanowska M and Osbourn AE 2002. Biosynthesis of triterpenoid saponins in plants. Adv. Bio. Eng./Biotech. **75**: 31.

Hayashi H, Hiraoka N, Ikeshiro Y, Kushiro T, Morita M, Shibuya M and Ebizuka Y 2000. Molecular cloning and characterization of a cDNA for *Glycyrrhiza glabra* cycloartenol synthase. Biol. Pharm. Bull. **23**(2): 231-234.

Huang ZW, Lin JC, Cheng ZX, Xu M, Guo MS, Huang XY, Yang ZJ and Zheng JG 2015. Production of oleanane-type sapogenin in transgenic rice via expression of *β*-amyrin synthase gene from *Panax japonicus* C. A. Mey. BMC Biotechnol. **15**: 45.

Huang ZW, Lin JC, Cheng ZX, Xu M, Huang XY, Yang ZJ, Zheng JG 2015. Production of dammarane-type sapogenins in rice by expressing the dammarenediol-II synthase gene from *Panax ginseng* C. A. Mey. Plant Sci. **239**: 106–114.

Kushiro T, Shibuya M and Ebizuka Y 1998. Beta-amyrin synthase - cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. Eur. J. Biochem. **256**(1): 238-244.

Liang YL and Zhao SJ 2008. Progress in understanding of ginsenoside biosynthesis. J. Plant Biol. **10**(4): 415-421.

Lichtenthaler HK, Rohmer M and Schwender J 1997. Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. Physiol. Plantarum **101**(3): 643-652.

Lichtenthaler HK 1999. The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. Annu. Rev. of Plant Physiol. and Plant Mol. Biol. **50**: 47-65.

Sambrook J, Russell DW 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press.

Suzuki H, Achnine L, Xu R, Matsuda SPT and Dixon RA 2002. A genomics approach to the early stages of triterpene saponin biosynthesis in Medicago truncatula. Plant J. **32**(6): 1033-1048.

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**(22): 5143-5149.

Wanke M, Skorupinska-Tudek K and Swiezewska E 2001. Isoprenoid biosynthesis via 1-deoxy-D-xylulose 5-phosphate/2-C-methyl-D-erythritol 4-phosphate (DOXP/MEP) pathway. Acta Biochim. Pol. **48**(3): 663-672.

Yun TK 2001. Panax ginseng - a non-organ-specific cancer preventive? Lancet Oncol. **2**(1): 49-55.

Zhao SJ, Hou CX, Xu LX, Liang YL, Qian YC and Yao S 2011. Effects of suppressing oleanane-type ginsenoside biosynthesis on dammarane-type ginsenoside production. J. Jilin University Engineering and Technology Edition) **41**(3): 865-868. (in Chinese)

(Manuscript received on 11 September, 2015; revised on 29 November, 2015)