



Cloning and characterization of a β -amyrin synthase gene from the medicinal tree *Aralia elata* (Araliaceae)

Y. Wu, H.D. Zou, H. Cheng, C.Y. Zhao, L.F. Sun, S.Z. Su, S.P. Li and Y.P. Yuan

College of Plant Science, Jilin University, Changchun, China

Corresponding author: Y.P. Yuan
E-mail: yapingyuan@yahoo.com.cn

Genet. Mol. Res. 11 (3): 2301-2314 (2012)
Received January 11, 2012
Accepted March 9, 2012
Published August 13, 2012
DOI <http://dx.doi.org/10.4238/2012.August.13.4>

ABSTRACT. *Aralia elata* is an important medicinal plant in China; it produces large amounts of oleanane type triterpene saponins. A full-length cDNA encoding β -amyrin synthase (designated as *AeAS*) was isolated from young leaves of *A. elata* by reverse transcription-PCR. The full-length cDNA of *AeAS* was found to have a 2292-bp open reading frame, encoding a protein with 763 amino acid residues. The deduced amino acid sequence of *AeAS* showed the highest identity (97%) to *Panax ginseng* β -amyrin synthase. When *AeAS* cDNA was expressed in *Escherichia coli*, an 87.8-kDa recombinant protein was detected by SDS-PAGE and Western blotting. The sequence was also heterologously expressed in the yeast *Pichia pastoris*, and production of β -amyrin was detected by HPLC. Tissue expression pattern analysis by real-time reverse transcription-PCR revealed that *AeAS* is strongly expressed in leaves and stems, and weakly expressed in roots and flowers.

Key words: *Aralia elata*; β -amyrin synthase; *Pichia pastoris*; PNY1; Triterpenoid saponins; Real-time RT-PCR

INTRODUCTION

Aralia elata (Miq.) Seem. is an upright deciduous small tree or shrub belonging to the Araliaceae family. It is widely distributed in China, Far East Russian, Japan, and Korea. The young shoot of *A. elata* is a popular edible plant in China, Japan and Korea. The stem bark and root bark are used as traditional medicine for neurasthenia, rheumatism, hepatitis, diabetes, stomach spasm, and as a tonic. It has been reported that oleanane type triterpenoid saponins were the main active components (New Medical College of Jiangsu, 1977). Numerous new oleanane type triterpene saponins were isolated from the leaves (Kim et al., 2005), stem barks (Nhiem et al., 2011), root barks (Song et al., 2001), and buds (Li et al., 2006, 2007) of this plant. These saponins were reported to show anti-diabetic activity (Kim et al., 1993), cytoprotective effect on carbon tetrachloride-induced hepatic injury (Saito et al., 1993), ethanol absorption inhibitory effect (Yoshikawa et al., 1996a), hypoglycemic activity (Yoshikawa et al., 1995, 1996b), antioxidant activity and hypocholesterolemic and hypolipidemic effect (Chuang and Jung, 2003). Recently, many studies were focused on antiulcer, anti-inflammatory and anticancer activity of these oleanane type triterpene saponins from *A. elata* (Zhang et al., 2006; Lee et al., 2009; Nhiem et al., 2011).

Oleanane type triterpenoid saponins are synthesized starting from the isoprenoid pathway. Two molecules of isopentenyl diphosphate (IPP) are condensed with one molecule of dimethylallyl diphosphate (DMAPP) by farnesyl diphosphate synthase (FPS), creating the molecule farnesyl diphosphate (FPP). Two FPP molecules are joined by squalene synthase (SS), originating squalene. Squalene is oxidized by squalene epoxidase (SE) to 2,3-oxidosqualene. 2,3-oxidosqualene is cyclized to produce β -amyrin by β -amyrin synthase (β -AS) belonging to the oxidosqualene cyclase (OSC) family. β -amyrin is thought to be a precursor for oleanolic acid, which is the aglycone of oleanane type triterpenoid saponins. β -amyrin is produced from oxidosqualene by product specific β -AS (monofunctional) or as one of the multifunctional OSCs (Yendo et al., 2010). Up to now, a number of product specific β -AS have been cloned and their function have been identified by heterologous expression in yeast: *PNY1* (Kushiro et al., 1998a) and *PNY2* (Kushiro et al., 1998b) from *Panax ginseng*, *GgbAS1* from *Glycyrrhiza glabra* (Hayashi et al., 2001), *PSY* from *Pisum sativum* (Morita et al., 2000), *AsbAS1* (*Sad1*) from *Avena strigosa* (Haralampidis et al., 2001), *MtAMY1* from *Medicago truncatula* (Iturbe-Ormaetxe et al., 2003), *BPY* from *Betula platyphylla* (Zhang et al., 2003), *EtAS* from *Euphorbia tirucalli* (Kajikawa et al., 2005), *LjOSC1* from *Lotus japonicas* (Sawai et al., 2006), *SvBS* from *Saponaria vaccaria* (Meesapyodsuk et al., 2007), *BgbAS* from *Bruguera gymnorhiza* (Basyuni et al., 2007), *AsOXAL* from *Aster sedifolius* (Cammarelli et al., 2008), *AtBAS* from *Arabidopsis thaliana* (Shibuya et al., 2009), *GsAS1* from *Gentiana straminea* (Liu et al., 2009), *NsbAS1* from *Nigella sativa* (Scholz et al., 2009).

Although extraction, structure analysis and pharmacological tests of oleanane type triterpenoid saponins from *A. elata* have been extensively studied, the understanding of their biosynthesis is still limited. With the aim to enhance the production of oleanane type triterpenoid saponins and to understand their biosynthesis pathway in *A. elata*, an *AeAS* gene was first cloned and characterized. Furthermore, expression patterns of *AeFPS*, *AeSS*, *AeSE*, and *AeAS* genes involved in the oleanane type triterpenoid saponin synthesis pathway were also determined by real-time RT-PCR.

β -AS, is to our knowledge the key synthase connected with oleanane type triterpenoid

saponin production. Therefore, we decided to isolate the gene encoding β -amyrin towards improving saponin yield. Here, we report on the first cloning and characterization of a β -AS gene from *A. elata*.

MATERIAL AND METHODS

Plant materials

A. elata was cultivated in the Division of Agriculture of Jilin University in Changchun, China. Young leaves, young stems, young later roots, and flowers were collected from 3-year-old plants in August. Harvested plant materials were frozen with liquid nitrogen and stored at -80°C .

Strains, plasmids and media

Escherichia coli strain DH5 α and the pGM-T vector (Tiangen, Beijing, China) were used for cloning and vector construction. *E. coli* strain BL21 (DE3) and the pET28a vector (Novagen, Madison, WI, USA) were used for prokaryotic expression. Bacterial cells were cultured in Luria Bertani medium [LB, 1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract, pH 7.5]. *Pichia pastoris* strain GS115 (His⁻ Mut⁺) and the pPIC9K vector (Invitrogen, Carlsbad, CA, USA) were used for eukaryotic expression. For *P. pastoris* cultivations, the following media were used at appropriate steps: yeast extract peptone dextrose medium (YPD, 1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose), yeast nitrogen base (YNB), minimal dextrose medium (MD, 1.34% YNB, 4 x 10⁻⁵% biotin, 2% dextrose), minimal medium (MM, 1.34% YNB, 4 x 10⁻⁵% biotin, 0.5% methanol), buffered complex glycerol medium (BMGY, 1% yeast extract, 2% peptone, 100 mM phosphate buffer, pH 6.0, 1.34% YNB, 4 x 10⁻⁵% biotin, 1% glycerol), and buffered complex methanol medium (BMMY, 1% yeast extract, 2% peptone, 100 mM phosphate buffer, pH 6.0, 1.34% YNB, 4 x 10⁻⁵% biotin, 1% glycerol, 0.5% methanol). All chemicals were from Sangon (Shanghai, China).

A. elata cDNA synthesis

Total RNA was extracted from *A. elata* young leaves using RNAiso Plus (Takara, Tokyo, Japan), according to manufacturer instructions. The quality and concentration of RNA were assessed by ethidium bromide-stained agarose 1% (w/v) gel electrophoresis and spectrophotometry. The RNA was converted to cDNA by providing 1 μg total RNA as the template in a 20 μL reaction using M-MLV reverse transcriptase (Takara) in the presence of 1 μL 0.5 $\mu\text{g}/\mu\text{L}$ oligo(dT)₁₈ (Takara) and 2 μL 10 mM deoxyribonucleoside triphosphate (dNTP) (Takara) at 42 $^{\circ}\text{C}$ for 1 h.

Cloning of *AeAS* cDNA

Based on the mRNA sequence of *PNY1* (GenBank accession No. AB009030) from *P. ginseng*, *AeAS*-specific primers (forward, 5'-ATGTGGAAGCTTAAGATAGCGG-3' and reverse, 5'-TTAGGTGCCTAGGGACGGTAAT-3') were used in the amplification reaction.

Using cDNA from young leaves of *A. elata* as the template, a 35-cycle PCR using LA *Taq* (Takara) with proof reading activity was conducted. The cycling parameters were 94°C for 1 min, 55°C for 1 min, 72°C for 2.5 min and a final 10-min extension at 72°C using a PTC-200 Peltier Thermal Cycler (Bio-Rad, Hercules, CA, USA). The PCR product was recovered by electrophoresis on 1% (w/v) agarose gel, cloned into pGM-T easy vector and transformed into competent *E. coli* DH5 α cells for sequencing (Sangon).

Sequence and phylogenetic analysis

Multiple sequence alignments were generated using the CLUSTAL W program (Thompson et al., 1994). Phylogenetic analysis of deduced amino acid alignments was performed using the neighbor-joining method with the program and TreeView software (Page, 1996). Bootstrap analysis with 1000 replicates was used to assess the strength of nodes in the tree. The GenBank accession Nos. of β -AS of other plants were as follows: BAA33461 (*P. ginseng* PNY1), BAA33722 (*P. ginseng* PNY2), BAA89815 (*G. glabra* GgbAS1), BAA97558 (*P. sativum* PSY), CAC84558 (*A. strigosa* AsbAS1), CAD23247 (*M. truncatula* MtAMY1), BAB83088 (*B. platyphylla* BPY), BAE43642 (*E. tirucalli* EtAS), BAE53429 (*L. japonicas* LjOSC1), ABK76265 (*S. vaccaria* SvBS or *Vaccaria hispanica* BS), BF80443 (*B. gymnorhiza* BgbAS), AAX14716 (*A. sedifolius* AsOXA1), BAG82628 (*A. thaliana* AtBAS), ACO24697 (*G. straminea* GsAS1), and ACH88048 (*N. sativa* NsbAS1).

Construction of *AeAS* prokaryotic expression vector

The entire ORF of *AeAS* cDNA was PCR amplified using forward primer (5'-GG GAATTCCATATGATGTGGAAGCTTAAGATAGCGG-3') (*Nde*I restriction site underlined and translation start codon in bold), and reverse primer (5'-ACGCGTCGACTTAGGTGCC TAGGGACGGTAAT-3') harboring an *Sac*I site (underlined) and a stop codon (bold). The PCR product was gel purified and cloned into pGM-T easy vector for sequencing. After the sequence was confirmed to be correct, the target fragment was digested with *Nde*I and *Sac*I, gel purified, and ligated into the *Nde*I and *Sac*I sites in pET28a. The constructed plasmid pET28a-*AeAS* was transformed into *E. coli* strain BL21 (DE3) according to manufacturer instructions (Novagen). The empty pET28a vector was also transformed for a negative control.

Heterologous expression of *AeAS* in *E. coli*

E. coli BL21 (DE3) cells harboring recombinant plasmid pET28a-*AeAS* were grown in 10 mL LB medium containing 100 μ g/mL kanamycin. Cultures were shaken at 200 rpm and held at 37°C to an absorbance of 0.7 to 1.0 at 600 nm before which isopropyl thio- β -D-galactoside (IPTG) was added to a final concentration of 0.1 mM. After 4 h of additional incubation, bacterium cells were collected by centrifugation and resuspended in 2X sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer (80 mM Tris-HCl, pH 6.8, 100 mM dithiothreitol, 2% SDS, 0.006% bromophenol blue, and 15% glycerol). Then, samples were boiled for 10 min and separated on 12% SDS-PAGE. Proteins were visualized by staining in Coomassie brilliant blue R-250 (Sangon). For the Western blot procedure, proteins were transferred onto nitrocellulose membrane (0.45 μ m; Invitrogen) using a MiniTrans blot system (Bio-

Rad). Proteins transferred to the membrane were blocked with 5% bovine serum albumin in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl (TBS buffer) for at least 2 h. The membrane was twice washed with TBS buffer containing 0.05% Tween-20 (TBST buffer) for 10 min each, and then incubated for 2 h in TBS buffer and anti-His antibody (mouse monoclonal antibody; Tiangen) diluted 1:2000. The membrane was then washed three times with TBST buffer and exposed to horseradish peroxidase conjugated goat anti-mouse IgG antibody (Sigma, St. Louis, MO, USA) diluted 1:3000 for 2 h at 37°C. Finally, the membrane was washed three times with TBST buffer for 10 min each, and color development was performed with the TMB kit (Tiangen).

Construction of *AeAS* eukaryotic expression vector

Based on the recombination plasmid pET28a-*AeAS*, the fusion sequence of entire ORF of *AeAS* cDNA and His-Tag sequence at 5' region was PCR amplified using forward primer (5'-GCTACGTACATCATCATCATCACAGCAG-3') (*Sna*BI restriction site underlined and His-Tag in bold), and reverse primer (5'-ATTTGCGGCCGCTTAGGTGCCTAGGGACGGTAAT-3') harboring an *Not*I site (underlined) and a stop codon (bold). The PCR product was gel purified and cloned into pGM-T easy vector for sequencing. After the sequence was confirmed to be correct, the target product was digested with *Sna*BI and *Not*I, gel purified, and ligated into the *Sna*BI and *Not*I sites in pPIC9K (Invitrogen). The constructed plasmid pPIC9K-*His₆-AeAS* was then linearized with *Bgl*II and purified for transformation. The empty pPIC9K was also linearized for a negative control.

Transformation of the *P. pastoris* strain

Transformation of the *P. pastoris* strain GS115 (His⁻ Mut⁺) was performed using the electroporation method according to pPIC9K vector manual (Invitrogen). Approximately 20 μ L linearized plasmid DNA (about 10 μ g) was used for electroporation in 0.2 cm cuvettes, using Gene Pulse Xcell™ Electroporation System (Bio-Rad) at 1.5 kV, 25 μ F and 200 Ω for 5 s. Immediately after pulsing, 1 mL 1 M ice-cold sorbitol was added to the cuvette cells, and then 200 to 600 μ L of the mixture was spread on MD plates at 30°C for screen of His⁺ transformants. To screen for methanol utilization, each colony on MD plates was first spotted onto an MM plate and then onto a new MD plate. After 4 h, Mut⁻ and Mut⁺ colonies were identified. The clone screening was performed using PCR analysis with *AeAS* specific primer (5'-TACTATTGCCAGCATTGCTGC-3') and 3'AOX1 primer (5'-GCAAATGGCATTCTGACATCC-3').

Heterologous expression of *AeAS* in *P. pastoris*

The screened recombinant strain (His⁺ Mut⁺) integrated with pPIC9K-*His₆-AeAS* was cultured in 20 mL BMGY medium with a 100-mL flask at 30°C, until an OD₆₀₀ value of 2 to 6 was attained. Yeast cells were harvested by centrifugation at 3000 g at room temperature for 5 min and then resuspended in 100 mL BMMY medium with a 250-mL flask. Yeast cells were grown at 30°C for 4 days in a shaking incubator at 200 rpm. To induce expression of the recombinant protein, the methanol at a final concentration of 0.5% (v/v) was added to the culture every 24 h. Supernatants of the culture at 0, 24, 48, and 72 h were collected for the extraction

and analysis of the recombinant protein. The procedure of SDS-PAGE and Western blot was the same as described above.

HPLC determination of β -amyrin in *P. pastoris*

After methanol induction for 3 days, the yeast cells were collected and baked until completely dry at 80°C. The dry sample (0.2 g) was extracted with 25 mL HCl/alcohol (2:8, v/v) solution at 95°C for 4 h and concentrated to dryness under reduced temperature. The dry material was dissolved with 30 mL sterile deionized H₂O and extracted by ether for 4 times (15 mL each). The mixed extraction solution was concentrated to dryness at room temperature. The crude sample was dissolved with 10 mL methanol and filtered with a 0.45- μ m filter HPLC analysis. The HPLC equipment was carried out on a Shimadzu LC-20A System (Shimadzu, Kyoto, Japan) including two LC-20AT automatic rinsing pumps, an SPD-M20A UV-VIS detector, a manual sample injector 7725 valve with a 20- μ L loop, and a DGU-20A₃ degasser. All separations were performed using a C₁₈ column (250 x 4.6 mm, 5 μ m particle size) from Phenomenex (Torrance, CA, USA). The mobile phase was methanol/water (98:2, v/v). The injection volume was 10 μ L, the effluent was monitored at 254 nm, the flow rate was 1.0 mL/min and the column temperature was 40°C. A standard calibration curve assembled from a range of concentrations (25, 125, 250, 500, 1000 μ g/mL) of β -amyrin (Sigma) was used as standard for HPLC analysis. Data acquisition and processing were accomplished with the Shimadzu LC solution software.

Expression of *AeAS*, *AeFPS*, *AeSS*, and *AeSE* in *A. elata*

Total RNA extraction and cDNA synthesis from young leaves, young stems, young later roots, and flowers of *A. elata* were the same as described above. Based on the mRNA sequence of *AeFPS* (HM219226), *AeSS* (GU354313), *AeSE* (GU354314), *AeAS* (HM219225), and *Panax notoginseng* glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*, DQ186631) genes in GenBank, we designed specific primers for quantitative real-time RT-PCR as follows: primers for *AeFPS* (product length of 260 bp, forward primer: 5'-CTTATTGAAATGGGAACCTA-3', reward primer: 5'-TCATACTCGGCAAATACATC-3'), primers for *AeSS* (product length of 274 bp, forward primer: 5'-AGTCTTCAGAGGGGTAGTG-3', reward primer: 5'-GAAGATAATAGCAATCAGGGC-3'), for *AeSE* (product length of 258 bp, forward primer: 5'-GGAAGAGAGAGAAAATAACG-3', reward primer: 5'-CATTACACAATCCTCCAG-3'), for *AeAS* (product length of 197 bp, forward primer: 5'-TTCCCAAGTGAAGGTAG-3', reward primer: 5'-GAACACTGTATCAAGATGCC-3'), for *AeGAPDH* (product length of 133 bp, forward primer: 5'-TTCCTCTACTGACCAC-3', reward primer: 5'-GGAAGCACTTTACCAACAG-3'). RT-PCR was carried out in a reaction mixture of 20 μ L, containing 2 μ L 1:5 (v/v) dilution of synthesized cDNA, primers to a final concentration of 0.4 μ M each, 8 μ L 2.5X RealMasterMix, 1 μ L 20X SYBR solution (Takara). The standard cycling conditions were 95°C 30 s, followed by 40 cycles of 95°C for 5 s, 60°C 30 s, 72°C 30 s with the ABI 7500 system (Applied Biosystems Inc., Foster City, CA, USA). The data were analyzed with the 7500 Fast System SDS version 1.3.1 software (Applied Biosystems). All reactions were performed in triplicate, and relative gene-expression determinations were made with the comparative delta-delta C_T method (2^{- $\Delta\Delta$ C_T}; Livak and Schmittgen, 2001) by using the housekeeping *GAPDH* gene as the internal control.

RESULTS AND DISCUSSION

Cloning of a cDNA encoding *AeAS*

In order to isolate an *AeAS* gene, specific primers were designed from the sequence of *P. ginseng* PNY1 (Kushiro et al., 1998a). Both *A. elata* and *P. ginseng* belong to the Araliaceae family and they are highly homology. Finally, a full-length cDNA encoding for β -AS was isolated from *A. elata* young leaves. The cDNA contained an ORF of 2292 bp and was deduced to code for a 763 amino acid protein (Figure 1). The *AeAS* amino acid sequence showed the high identity of 97% to PNY1 and 87% to PNY2, as would be expected, since all species belong to the Araliaceae. So we may design species-specific primers to clone unknown genes from known genes of highly homologous plants, which are the same family without degenerate primers PCR, 5'-rapid amplification of cDNA ends (RACE), and 3'-RACE. Based on this strategy, we cloned *AeFPS* (HM219226), *AeSS* (GU354313), *AeSE* (GU354314) successfully. *AeAS* possesses the amino acid motif DCTAE, thought to form part of the active site of OSCs (Abe et al., 1993; Abe and Prestwich 1995) and the four QW (QXXXGXW) motifs, which are thought to be involved in the stabilization of carbocationic intermediates formed during the cyclization of OSCs (Poralla et al., 1994). In addition, *AeAS* amino acid sequence contains the Trp residue in the MWCYCR motif that plays an important role in the formation of β -amyrin in *P. ginseng* (Kushiro et al., 2000). The second Lys in the FIKKSQ motif of *AeAS* is strictly conserved in monofunctional β -AS and may favor E-ring expansion to produce oleanane or ursane type triterpene saponins (Basyuni et al., 2007).

The DCTAE motif is underlined, the MWCYCR motif is double underlined, the FIKKSQ motif is underlined with a discontinued line, and the QW motifs are boxed.

Sequence and phylogenetic analysis of *AeAS* amino acid sequence

Alignment results showed that *AeAS* amino acid sequence had identity of 97% to PSN1, 87% to PSN2, 84% to BPY, 84% to BgbAS, 83% to EtAS, 82% to GgbAS1, 82% to LjOSC1, 81% to MtAMY1, 81% to PSY, 79% to SvBS, 79% to AsOXA1, 76% to AtBAS, 76% to GsAS1, 72% to NsbAS1, and 49% to AsbAS1. Fifteen dicotyledonous monofunctional β -ASs showed high similarities (72 to 97%) to each other and displayed lower similarities (48%) to AsbAS1 from the monocotyledonous plant *A. strigosa*. A phylogenetic tree was constructed from the deduced amino acid sequences of β -ASs (Figure 2). In the tree, AsbAS1 is clearly distinct from other β -ASs of dicotyledonous plants. The phylogenetic tree shows that *AeAS*, PNY1, together with PNY2, form one branch of the β -AS cluster. MtAMY1, PSY, LjOSC1, and GgbAS1 form the other branch, because all species belong to the Leguminosae family. Figure 3 shows that crucial motifs of β -AS are well conserved in these dicot plants, except that Gly257 instead of Trp257 in MWCYCR motif of GsAS1, and Ala448 instead of Lys448, Asn448 instead of Lys448, Asn450 instead of Lys450, and Gln452 instead of Lys452 in FIKKSQ motif of AsOXA1, GsAS1, AtBAS, and NsbAS1, respectively. Although their crucial residues are changed from other plants, they are also product specific β -AS by function establishment in *P. pastoris* (Cammareri et al., 2008; Shibuya et al., 2009; Liu et al., 2009; Scholz et al., 2009). Therefore, these observations suggested that the presence of additional protein domains act to control the reaction product of β -ASs. From the high level of amino acid sequence identity and comparison of critical motifs with β -AS of other plants, the *AeAS* cDNA may code for the product specific β -AS.

```

1  ATGTGGAAGCTTAAGATAGCGGAAGGGAATAAGAATGACCCGATTTTGTACAGTACCAATAATTTGTAGGGCGGCAAACGTGGGAGTTC
1  M W K L K I A E G N K N D P Y L Y S T N N F V G R Q T W E F
91  GACCCGGATTATGTGGCTAGTCCCGGAGAGCTAGAGGAGGTGGAACAAGCTCGACGGCAGTTTGGGAGAACCAGTATCAGGTCAAGCCT
31  D P D Y V A S P G E L E E V E Q A R R Q F W E N R Y Q V K P
181 AGTGGTGATCCTCTGGCCTATGCAGTTCCTGAGAGAGAAGAATTTCAACAAACAATCCCCAAGTGAAGTAGGAGATGACGAGGCA
61  S G D L L W R M Q F L R E K N F K Q T I P Q V K V G D D E A
271 GTTACTTATGAGGCCGCCACTACCACACTCCGAAGGGCGTCCACTTTTTCTCAGCTTTGCAGGCCAGCGACGGTCATTGGCCCTGCAGAG
91  V T Y E A A T T T L R R A V H F F S A L Q A S D G H W P A E
361 AATGCCGACCTCTCTTTTCTCCCGCCCTTGGTGATGTGTATATATCACAGGCATCTTGATACAGTGTCCAGCAGAACATCGA
121 N A G P L F F L P P L V M C V Y I T G H L D T V F P A E H R
451 AAAGAAATCTTCGCTACATATATTGTACATCAGAATGAAGATGGCGGGTGGGATTCATATTGAGGGGCATAGCACCATGTTCTGCACA
151 K E I L R Y I Y C H Q N E D G G W G F H I E G H S T M F C T
541 ACTCTAGCTACATTTGTATGCGTATACCTGGAGAGGGCCGAGTGGTGGTGAACAATCGTGTGTGAGGGCCGAAATGGATCCTT
181 T L S Y I C M R I L G E G P D G G V N N A C V R G R K W I L
631 GACCATGGCAGTGAACCGCTATACCTTCATGGGGCAAGACTTGGCTTCGATACTTGGTGTATATGAATGGACAGGAACAACCAATG
211 D H G S V T A I P S W G K T W L S I L G V Y E W T G S N P M
721 CCCCAGAAATCTGGATTCTCCCTCTTCTTCTCCTATGCACCCAGCTAAATGTGGTGTATTGCCGGATGGTTTACATGCCAATGTCA
241 P P E F W I L P S F L P M H P A K M W C Y C R M V Y M P M S
811 TATTTATATGGGAAGAGTTTGTGGTCCAATAACTCCTCTTATTTTACAATTAAGAGAAGAACTTTATGCTCAACCCCTACAATGAAATC
271 Y L Y G K R F V G P I T P L I L Q L R E E L Y A Q P Y N E I
901 AATGGAGAAAACACGCTCATGTGTGCAAGGAGGACATCTACTATCCTCATCCTTTAATAACAAGACCTGCTCTGGGATAGTCTCTAT
301 N W R K T R H V C A K E D I Y Y P H P L I Q D L L W D S L Y
991 GTATTAAGTGAACCTCTTTTAACTCGTTGGCCATTTAACAATGAGAGAGAAAGCTCTGCAGACTACCATGAAACACATTCACATGAA
331 V L T E P L L T R W P F N K L R E K A L Q T T M K H I H Y E
1081 GATGAGAACAGTTCGATATATACCATTGGATGTGTGGAAAGGTTTGTGTATGCTTGCCTTGTGGGTGAGAGCTCAAATGGAGATTAT
361 D E N S R Y I T I G C V E K V L C M L A C W V E D P N G D Y
1171 TTCAGAAACACCTTGCAAGGATCCCAGATTATATATGGTGTGCAAGTGAAGTGAAGATGCAGAGTTTGGTAGTCAGGATGGGAT
391 F K K H L A R I P D Y I W V A E D G M K M Q S F G S Q E W D
1261 ACAGTTTCGGCATTCAAGCATTGTTGGCTAGTACTAATCATGAAATGGACCTACTCTTATGAAAGGACGACTTCATCAAAAAG
421 T G F G I Q A L L A S D L T H E I G P T L M K G H D -----
1351 TCCCAGTCAAGGATAATCCTTCTGGTGACTTTAAAGGCATGTATCGCCACATTTCTAAAGGATCGTGGACTTTTTTTCAGATCAAGATCAC
451 S Q V K D N P S G D F K G M Y R H I S K G S W T F S D Q D H
1441 GGATGGCAAGTTTCTGATTGCACCTGCAGAGGACTAAAGTGTTCGCTTATTTTCTCAACGATGCCAGAGGAAATAGTTGGCAAGAAATG
481 G W Q V S D C T A E G L K C C L I F S T M P E E I V G K K G M
1531 GAACCCAGAAAGACTGTATGATTCTGTTAATGTTCTGCTTTCCTCCAGAGCAAAAATGGTGGCTTAGCAGCATGGAGCCTGACGAGGCT
511 E P E R L Y D S V N V L L S L Q S K N G G L A A W E P A G A
1621 CAGGAATGGTTAGAGTGTCTCAATCTACGGAATCTTTGCGAGCATGTGCATTTGAGCAGGATGTAGAGTGCACCTTCATCGGCAATC
541 Q E W L E L L N P T E F F A D I V I E H E Y V E C T S S A I
1711 CAAGCTCTGGTTCGTTTAAAAAGTTATATCTGGACACCCGAAAGAAGGAGATAGATAATTTTATATGAATGCTGCCGTTACCTTGAA
571 Q A L V L F K K L Y P G H R K K E I D N F I M N A V R Y L E
1801 GACATACAAATGCCTGATGGTTTCATGGTATGGAACTGGGGTGTGTGCTTTACTTATGTTAGCTGGTTTGTCTTGGGGGCTAGCGCA
601 D I Q M P D G S W Y G N W G V C F T Y G S W F A L G G L A A
1891 GCTGGAAGACATACTACAATTTGTGACGCTGTCGTAAGCTGTGATTTTCTACTCAAATCACAGATGGATGACGGCGGTTGGGGAGAA
631 A G K T Y Y N C A A V R K A V D F L L K S Q M D D G G W G E
1981 AGTACCTTTCTTCCCGAAAAGGTATATGTACCATTAGAAGGAAACCGCTCAAATTTGGTACATACTGGATGGCCCTTGATGGGACTC
661 S Y L S C P K K V Y V P L E G N R S N L V H T G W A L M G L
2071 ATTCATTCGGAGCAGGCTGAGAGAGACCCACACCTCTTCCCGTGGCCCAAGTTATTGATCAATTTCTCAGATGGAAGATGGTATTTT
691 I H S E Q A E R D P T P L H R A A K L L I N S Q M E D G D F
2161 CCCCACAGGAAATAACTGGAGTTTTTATGAAGAATGCATGTTGCACATGCAGCTTACCAGAACATATACCATTGTGGGCTCTAGCA
721 P Q Q E I T G V F M K N C M L H Y A A Y R N I Y P L W A L A
2251 GAGTATCGGAGGCGGGTCCCATACCGTCCCTAGGCACCTAA
751 E Y R R R V P L P S L G T *

```

Figure 1. Nucleotide and deduced amino acid sequences of *AeAS* from *Aralia elata* (GenBank accession Nos. HM219225 and ADK12003). The DCTAE motif is underlined, the MWCYCR motif is double underlined, the FIKKSQ motif is underlined with a discontinued line, and the QW motifs are boxed.

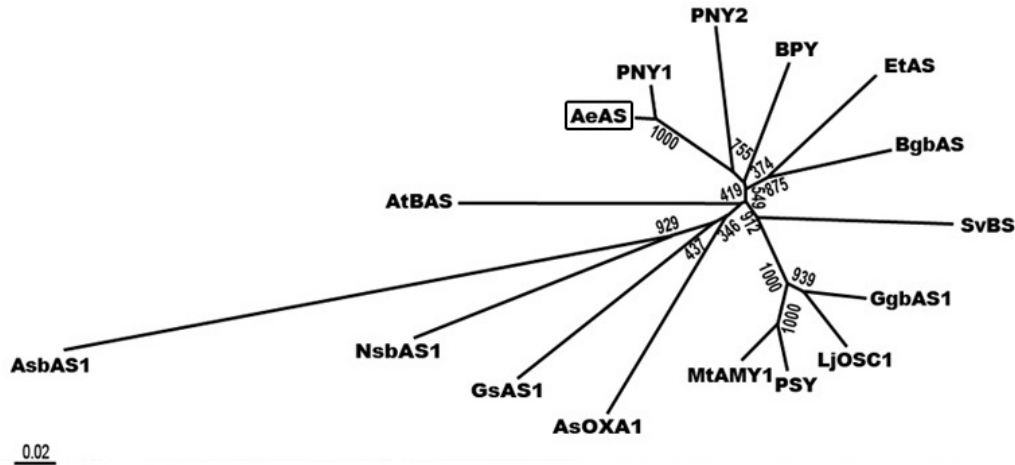


Figure 2. Phylogenetic tree of the deduced peptide sequences of *AeAS* and other β -amyrin synthase plants. Numbers indicate bootstrap values from 1000 replicates. *AeAS* is boxed.

	QXXGXXW	QXXGXXW	MWCYCR	FIKKSQ	DCTAE	QXXGXXW	QXXGXXW
AeAS	111 QASDGHW 117 161 QNEDGGW 167 258 MWCYCR 263 447 FIKKSQ 452 486 DCTAE 490 603 MPFGSW 609 652 QMDGGW 658						
PNY1	111 QASDGHW 117 161 QNEDGGW 167 258 MWCYCR 263 447 FIKKSQ 452 486 DCTAE 490 603 MPFGSW 609 652 QMDGGW 658						
PNY2	109 QADSGHW 115 159 QNDDGGW 165 256 MWCYCR 261 445 FIKKSQ 450 484 DCTAE 488 601 MPFGSW 607 650 QRSIDGGW 656						
BPY	109 QASDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 446 FIKKSQ 451 485 DCTAE 489 602 MPFGSW 608 651 QRDNGGW 657						
EtAS	109 QASDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 446 FIKKSQ 451 485 DCTAE 489 602 CTARGSW 608 651 QKQGGW 657						
BgbAS	109 QASDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 446 FIKKSQ 451 485 DCTAE 489 602 QTSFGSW 608 651 QRDNGGW 657						
SvBS	109 QASDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 447 FIKKSQ 452 486 DCTAE 490 603 CYPFGSW 609 652 QKEDGGW 658						
GgbAS1	109 QTSIDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 446 FIKKSQ 451 485 DCTAE 489 602 CTARGSW 608 651 QREDGGW 657						
LjOSC1	109 QTSIDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 446 FIKKSQ 451 485 DCTAE 489 602 CTARGSW 608 651 QRGDGGW 657						
PSY	109 QTSIDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 446 FIKKSQ 451 485 DCTAE 489 602 QTEIDGGW 608 651 QREDGGW 657						
MtAMY1	109 QTSIDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 446 FIKKSQ 451 485 DCTAE 489 602 CTARGSW 608 651 QREDGGW 657						
AsOXA1	109 QADSGHW 115 159 QNEDGGW 165 256 MWCYCR 261 445 FIKKSQ 451 484 DCTAE 488 601 MPFGSW 607 650 QKPDGGW 656						
GsAS1	109 QASDGHW 115 159 QNEDGGW 165 256 MWCYCR 261 445 FIKKSQ 450 484 DCTAE 488 601 MPFGSW 607 650 QKQDGGW 656						
AtBAS	110 QASDGHW 116 160 QKEDGGW 166 257 MWSYCR 262 447 FIKKSQ 452 486 DCTAE 490 603 QTRIDGGW 609 652 QKDNGGW 658						
NsbAS1	113 QTSIDGHW 119 163 QNEDGGW 169 260 MWCYCR 265 449 YVRCSSQ 454 488 DCTCE 492 605 MPFGSW 611 654 QLESQGW 660						
AsbAS1	109 QADSGHW 115 159 QNEDGGW 165 256 FNCHTS 261 446 FIKKSQ 451 484 DCTAE 488 601 QNKFGSW 607 650 QQTIDGGW 656						

Figure 3. Alignment of amino acid sequences of high conserved regions in *AeAS* and other β -amyrin synthase plants. The high conserved regions, four QW motifs, the MWCYCR motif, FIKKSQ motif, and DCTAE motif are compared. The critical residues of these motifs are black. The identical amino acid residues are indicated with black background and the different amino acid residues with white background. Gray shade indicates 70% or more conservation among all the aligned sequences.

Heterologous expression and function characterization of *AeAS*

To further investigate the function of *AeAS*, the full-length cDNA of *AeAS* was heterologous expressed in *E. coli* and *P. pastoris*, and product from yeast expressing the *AeAS* gene induced by methanol was analyzed by HPLC. SDS-PAGE analysis identified a protein of molecular mass about 88 kDa expressed by the *E. coli* strain BL21 when induced by IPTG (Figure 4A). Western blot of this separation revealed that this product was a His-Tag fusion protein (Figure 4B). Similarly, SDS-PAGE analysis identified a recombination protein of molecular mass about 88 kDa expressed by the *P. pastoris* GS115 when induced by methanol (Figure 5A). The

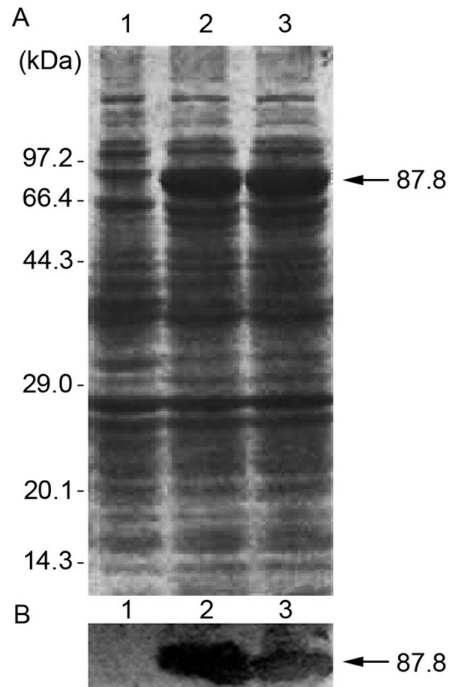


Figure 4. SDS-PAGE analysis and Western blot of *AeAS* expressed in *Escherichia coli*. **A.** SDS-PAGE; **B.** Western blot. Lane 1 = Empty pET28a vector; lanes 2 and 3 = pET28a containing *AeAS* induced by isopropyl thio- β -D-galactoside for 5 h.

recombination protein was also detected by Western blot (Figure 5B). The results showed that *AeAS* is successfully expressed in *E. coli* and *P. pastoris*. The product specificity of *AeAS* was tested by expression of the *AeAS* cDNA in yeast induced by methanol. HPLC analysis indicated that the product was not present in extracts of negative control cells carrying the empty pPIC9K vector, and the yeast cells expressing *AeAS* accumulated a product with a retention time of 11.7 min that was identical to that of standard β -amyrin (Figure 6). The results showed that the *AeAS* cDNA codes for a product specific β -AS, similar to other plants monofunctional β -AS. Furthermore, yeast appears to be an excellent model system not only for the functional characterization of product-specific β -AS, but also for the production of target products such as β -amyrin and other oleanane type triterpene saponins (Phillips et al., 2006; Lodeiro et al., 2007).

Expression of *AeAS* and other genes involving in triterpene saponin synthesis in *A. elata*

AeFPS, *AeSS*, *AeSE*, and *AeAS* are triterpene saponin pathway genes in *A. elata*. The melting curve of the reference gene *AeGAPDH* showed a single peak, indicating that a single PCR product was present. The Ct value of the reference gene *AeGAPDH* in different organs of *A. elata* was relatively stable, indicating that it can be used as internal control (data not shown). Therefore, we used the average Ct values of *AeGAPDH* as internal standard to nor-

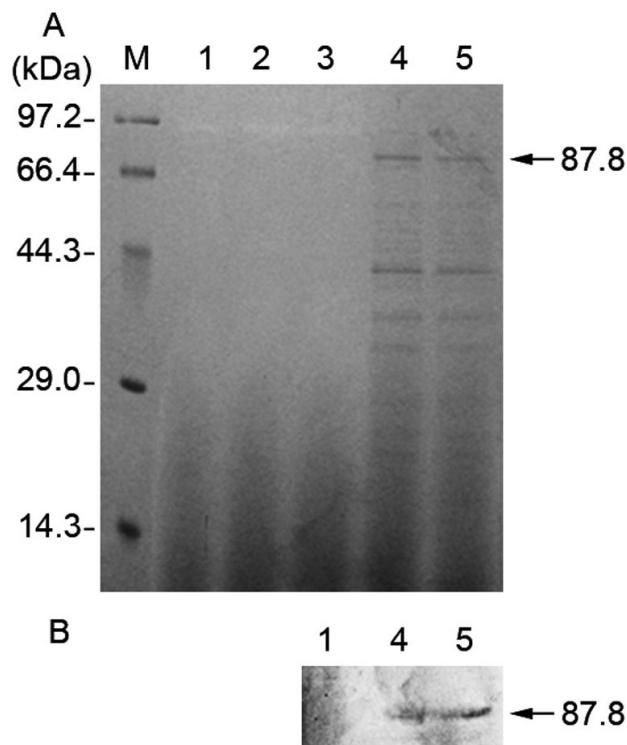


Figure 5. SDS-PAGE analysis and Western blot of *AeAS* expressed in *Pichia pastoris*. **A.** SDS-PAGE; **B.** Western blot. Lane M = Protein molecular weight marker (low); lane 1 = empty pPIC9K vector; lane 2 = pPIC9K containing *AeAS* without methanol; lane 3 = pPIC9K containing *AeAS* induced by methanol for 24 h; lane 4 = pPIC9K containing *AeAS* induced by methanol for 48 h; lane 5 = pPIC9K containing *AeAS* induced by methanol for 72 h.

malize the quantification of mRNA expression. Quantitative PCR analysis showed the relative expression of *AeFPS*, *AeSS*, *AeSE*, and *AeAS* in stems, leaves, later roots, and flowers of *A. elata* (Figure 7). All genes could be detected in all tissues, but at different levels. The expression of *AeFPS* was high in stems, moderate in leaves and flowers, and very low in later roots. The expression of *AeSS* was high in stems, moderate in leaves and flowers, and very low in later roots. The expression of *AeSE* was high in stems and leaves, and very low in later roots and flowers. The expression of *AeAS* was high in stems and leaves, moderate in flowers, and very low in later roots. We found that the expression of all genes is very low in later roots and the result is apparently conflicting with biochemical results that showed a high content of triterpenoid saponins in *A. elata* roots (Song et al., 2001). A possible explanation of the contradictory results is that saponins are synthesized in green parts of the plant, and then transported to the roots just like *Calendula officinalis* (Hostettmann and Marston, 1995).

The housekeeping gene *AeGAPDH* was used to normalize the amount of sample cDNA added to the reaction. Bars indicated standard deviation from triplicate amplification.

In conclusion, we successfully cloned and characterized the β -AS gene associated with oleanolic acid biosynthesis in *Aralia elata* (Miq.) Seem. for the first time.

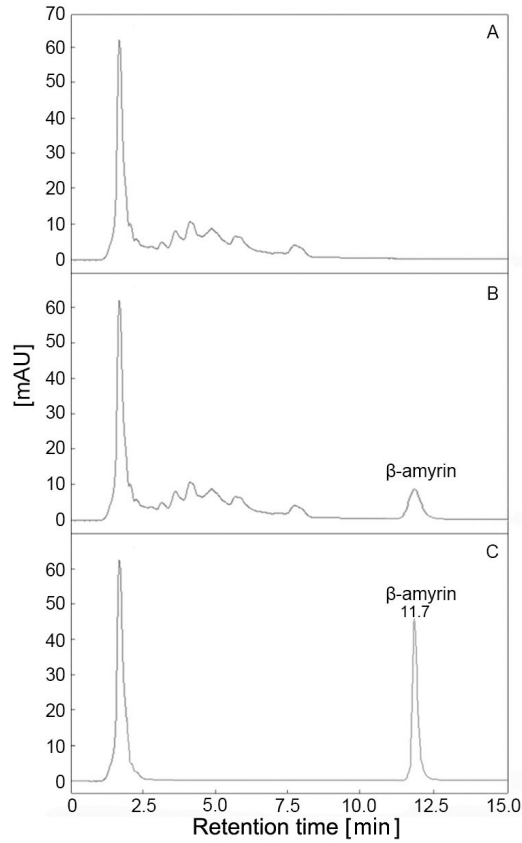


Figure 6. HPLC analysis of yeast cell extracts. **A.** Empty pPIC9K vector. **B.** pPIC9K containing *AeAS*. **C.** Standard β -amyrin. The retention time of standard β -amyrin is 11.7 min.

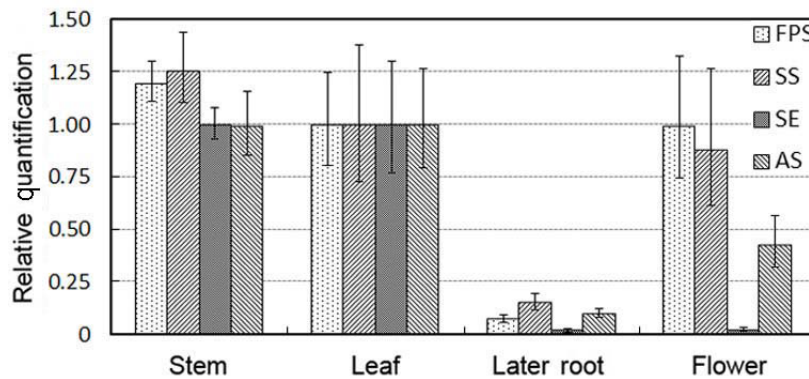


Figure 7. Relative quantification of expression levels for *AeFPS* (FPS), *AeSS* (SS), *AeSE* (SE), and *AeAS* (AS) genes in stems, leaves, later roots, and flowers of *Aralia elata*, respectively.

ACKNOWLEDGMENTS

Research supported by the Natural Science Foundation Program (#20101587) of Jilin Province, China, and the Application Technology Research Program of Changchun, China (#08YJ29).

REFERENCES

- Abe I and Prestwich GD (1995). Identification of the active site of vertebrate oxidosqualene cyclase. *Lipids* 30: 231-234.
- Abe I, Rohmer M and Prestwich GD (1993). Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. *Chem. Rev.* 93: 2189-2206.
- Basyuni M, Oku H, Tsujimoto E, Kinjo K, et al. (2007). Triterpene synthases from the Okinawan mangrove tribe, Rhizophoraceae. *FEBS J.* 274: 5028-5042.
- Camareri M, Consiglio MF, Pecchia P, Corea G, et al. (2008). Molecular characterization of β -amyrin synthase from *Aster sedifolius* L. and triterpenoid saponin analysis. *Plant Sci.* 175: 255-261.
- Chung CK and Jung ME (2003). Ethanol fraction of *Aralia elata* Seemann enhances antioxidant activity and lowers serum lipids in rats when administered with benzo(a)pyrene. *Biol. Pharm. Bull.* 26: 1502-1504.
- Haralampidis K, Bryan G, Qi X, Papadopoulou K, et al. (2001). A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots. *Proc. Natl. Acad. Sci. U. S. A.* 98: 13431-13436.
- Hayashi H, Huang P, Kirakosyan A, Inoue K, et al. (2001). Cloning and characterization of a cDNA encoding beta-amyrin synthase involved in glycyrrhizin and soyasaponin biosyntheses in licorice. *Biol. Pharm. Bull.* 24: 912-916.
- Hostettmann K and Marston A (1995). Saponins. Cambridge University Press, Cambridge.
- Iturbe-Ormaetxe I, Haralampidis K, Papadopoulou K and Osbourn AE (2003). Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus*. *Plant Mol. Biol.* 51: 731-743.
- Kajikawa M, Yamato KT, Fukuzawa H, Sakai Y, et al. (2005). Cloning and characterization of a cDNA encoding beta-amyrin synthase from petroleum plant *Euphorbia tirucalli* L. *Phytochemistry* 66: 1759-1766.
- Kim JS, Shim SH, Chae S, Han SJ, et al. (2005). Saponins and other constituents from the leaves of *Aralia elata*. *Chem. Pharm. Bull.* 53: 696-700.
- Kim OK, Lee EB and Kang SS (1993). Antihyperglycemic constituent of *Aralia elata* root bark. (II). Isolation and action of the constituents. *Saengyak Hakhoechi* 24: 219-222.
- Kushiro T, Shibuya M and Ebizuka Y (1998a). Beta-amyrin synthase-cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. *Eur. J. Biochem.* 256: 238-244.
- Kushiro T, Shibuya M and Ebizuka Y (1998b). Towards Natural Medicine Research in the 21st Century. In: Excerpta Medica International Congress Series (Ageta H, Aimi N, Ebizuka Y and Honda G, eds.). Elsevier Science, Amsterdam, 421-428.
- Kushiro T, Shibuya M, Masuda K and Ebizuka Y (2000). Mutational studies on triterpene synthases: engineering lupeol synthase into β -amyrin synthase. *J. Am. Chem. Soc.* 122: 6816-6824.
- Lee JH, Ha YW, Jeong CS, Kim YS, et al. (2009). Isolation and tandem mass fragmentations of an anti-inflammatory compound from *Aralia elata*. *Arch. Pharm. Res.* 32: 831-840.
- Li L, Song SJ, Li LZ, Liang ZX, et al. (2006). Chemical constituents of the buds of *Aralia elata* (Miq.) Seem. (III). *J. Shenyang Pharm. Univ.* 23: 495-498.
- Li L, Song SJ, Liang ZX and Xu SX (2007). A new triterpenoidal saponin from the buds of *Aralia elata* (Miq.) Seem. *Asian. J. Tradit. Med.* 2: 1-4.
- Liu Y, Cai Y, Zhao Z, Wang J, et al. (2009). Cloning and Functional Analysis of a β -amyrin synthase gene associated with oleanolic acid biosynthesis in *Gentiana straminea* MAXIM. *Biol. Pharm. Bull.* 32: 818-824.
- Livak KJ and Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402-408.
- Lodeiro S, Xiong Q, Wilson WK, Kolesnikova MD, et al. (2007). An oxidosqualene cyclase makes numerous products by diverse mechanisms: a challenge to prevailing concepts of triterpene biosynthesis. *J. Am. Chem. Soc.* 129: 11213-11222.
- Meesapyodsuk D, Balsevich J, Reed DW and Covello PS (2007). Saponin biosynthesis in *Saponaria vaccaria*. cDNAs encoding β -amyrin synthase and a triterpene carboxylic acid glucosyltransferase. *Plant Physiol.* 143: 959-969.
- Morita M, Shibuya M, Kushiro T, Masuda K, et al. (2000). Molecular cloning and functional expression of triterpene synthases from pea (*Pisum sativum*) new alpha-amyrin-producing enzyme is a multifunctional triterpene synthase. *Eur. J. Biochem.* 267: 3453-3460.

- New Medical College of Jiangsu (1977). Dictionary of Chinese Materia Medica. Shanghai Scientific and Technological Publishing, Shanghai.
- Nhiem NX, Lim HY, Kiem PV, Minh CV, et al. (2011). Oleanane-type triterpene saponins from the bark of *Aralia elata* and their NF-kappaB inhibition and PPAR activation signal pathway. *Bioorg. Med. Chem. Lett.* 21: 6143-6147.
- Page RD (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357-358.
- Phillips DR, Rasbery JM, Bartel B and Matsuda SP (2006). Biosynthetic diversity in plant triterpene cyclization. *Curr. Opin. Plant Biol.* 9: 305-314.
- Poralla K, Hewelt A, Prestwich GD, Abe I, et al. (1994). A specific amino acid repeat in squalene and oxidosqualene cyclases. *Trends Biochem. Sci.* 19: 157-158.
- Saito S, Ebashi J, Sumita S, Furumoto T, et al. (1993). Comparison of cytoprotective effects of saponins isolated from leaves of *Aralia elata* Seem. (Araliaceae) with synthesized bisdesmosides of oleanolic acid and hederagenin on carbon tetrachloride-induced hepatic injury. *Chem. Pharm. Bull.* 41: 1395-1401.
- Sawai S, Shindo T, Sato S, Kaneko T, et al. (2006). Functional and structural analysis of genes encoding oxidosqualene cyclases of *Lotus japonicus*. *Plant Sci.* 170: 247-257.
- Scholz M, Lipinski M, Leupold M, Luftmann H, et al. (2009). Methyl jasmonate induced accumulation of kalopanaxsaponin I in *Nigella sativa*. *Phytochemistry* 70: 517-522.
- Shibuya M, Katsube Y, Otsuka M, Zhang H, et al. (2009). Identification of a product specific β -amyrin synthase from *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 47: 26-30.
- Song SJ, Nakamura N, Ma CM, Hattori M, et al. (2001). Five saponins from the root bark of *Aralia elata*. *Phytochemistry* 56: 491-497.
- Thompson JD, Higgins DG and Gibson TJ (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.
- Yendo AC, de Costa F, Gosmann G and Fett-Neto AG (2010). Production of plant bioactive triterpenoid saponins: elicitation strategies and target genes to improve yields. *Mol. Biotechnol.* 46: 94-104.
- Yoshikawa M, Yoshizumi S, Ueno T, Matsuda H, et al. (1995). Medicinal foodstuffs. I. Hypoglycemic constituents from a garnish foodstuff "taranome", the young shoot of *Aralia elata* SEEM.: elatosides G, H, I, J, and K. *Chem. Pharm. Bull.* 43: 1878-1882.
- Yoshikawa M, Murakami T, Harada E, Murakami N, et al. (1996a). Bioactive saponins and glycosides. VI. Elatosides A and B, potent inhibitors of ethanol absorption, from the bark of *Aralia elata* SEEM. (Araliaceae): the structure-requirement in oleanolic acid glucuronide-saponins for the inhibitory activity. *Chem. Pharm. Bull.* 44: 1915-1922.
- Yoshikawa M, Murakami T, Harada E, Murakami N, et al. (1996b). Bioactive saponins and glycosides. VII. On the hypoglycemic principles from the root cortex of *Aralia elata* Seem.: structure related hypoglycemic activity of oleanolic acid oligoglycoside. *Chem. Pharm. Bull.* 44: 1923-1927.
- Zhang H, Shibuya M, Yokota S and Ebizuka Y (2003). Oxidosqualene cyclases from cell suspension cultures of *Betula platyphylla* var. *japonica*: molecular evolution of oxidosqualene cyclases in higher plants. *Biol. Pharm. Bull.* 26: 642-650.
- Zhang M, Liu G, Tang S, Song S, et al. (2006). Effect of five triterpenoid compounds from the buds of *Aralia elata* on stimulus-induced superoxide generation, tyrosyl phosphorylation and translocation of cytosolic compounds to the cell membrane in human neutrophils. *Planta Med.* 72: 1216-1222.