B-AMYRIN SYNTHASE, ONE OF THE MOST IMPORTANT KEY ENZYMES FOR TRITERPENE SKELETON FORMATION IN HIGHER PLANTS

YONG-SHENG MA*, RUI YANG†, SHAN ZHOU, YAN-CHAO YIN, XIAO-DONG ZHANG AND YING LIU*

School of Life Science, Beijing University of Chinese Medicine

*Corresponding author’s email: liuyliwd@sina.com Phone: +8601084738646; Fax: +8601084738611

Abstract

β-amyrin synthase (β-AS) is one of the most important key enzymes involved in mevalonic acid (MVA) pathway. It is a cyclase responsible for cyclization of 2, 3-oxidosqualene into β-amyrin, which is defined as an important branch point between primary and secondary metabolism. It has been found in 37 higher plant species. In this paper, we obtained 475 DNA sequences, 220 mRNA sequences, and 99 amino acid sequences of β-AS registered in NCBI by Oct, 2016, and analyzed conserved domains and the evolutionary relationships between different species with DNAMAN 6.0.3.99 and MEGA 5.0. In order to get the latest and comprehensive information of β-AS, more than 300 papers were searched and 80 of them were reviewed. Pub Med, Web of Science, Science Direct, and Research Gate, were information sources through the search terms of “β-amyrin synthase”, “biosynthesis”, “oxidosqualene cyclases” and their combinations, mainly from year 2010 to 2016. Studies were selected from Science Citation Index journals. All of the references linked to the registered DNA and mRNA sequences in NCBI database were also reviewed. The full-length of β-AS DNA sequence ranges from 3900 bp to 8800 bp, and β-AS mRNA sequence ranges from 2100 bp to 2900 bp. The bioinformatic analysis and a lot of papers show that Gln-Trp (QW) motifs, Asp-Cys-Thr-Ala-Glu (DCTAE) motif, and Met-Trp-Cys-Tyr-Cys-Arg (MWCYCR) motif, are mainly responsible for its catalytic function. So far, the function of β-AS has been verified in 30 species. This paper will lay a foundation for further studies of β-AS and other oxidosqualene cyclases for triterpene skeleton formation in higher plants.

Key words: β-amyrin synthase, Triterpene, MVA pathway, Bioinformatic analysis, Oxidosqualene cyclase.

Introduction

Terpenoids are very important secondary metabolites in higher plants, including monoterpene, sesquiterpenes, diterpene, triterpene, and polyterpene. Many of them are important active ingredients (Parveen et al., 2010; Shi et al., 2015; Shi et al., 2016; Basyuni et al., 2007), and possess remarkable pharmacological properties, such as antitumor (Petronelli et al., 2009), anti-HIV (Wei et al., 2008; Kongkum et al., 2013; Callies et al., 2015; Kuo et al., 2009), anti-inflammatory (Aziz et al., 2015; Li et al., 2014; Chen et al., 2014; Liaw et al., 2015; Jung et al., 2005), antibacterial (He et al., 2011; Shai et al., 2008; Wang et al., 2013; Harizion et al., 2015; Lopez et al., 2011), antiplatelet (Shi 2013; Yang et al., 2009; Li et al., 2013), hypocholesterolemic (Wang, 2016; Rezaei-Golmisheh et al., 2015), immune adjuvant (Qiao, 2015), and other activities (Song et al., 2011; Jiang et al., 2008; Wang et al., 2009).

Currently, two biosynthetic pathways of terpenoids have been basically clarified in higher plants, they are mevalonic acid (MVA) pathway in the cytoplasm and 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway in the plastid (Kim et al., 2014; Sando et al., 2008). MEP pathway is responsible for monoterpene and diterpene formation (Dudareva et al., 2005), and MVA pathway is responsible for sterol, sesquiterpenes, and triterpene formation (Yang et al., 2012). In MVA pathway, 2, 3-oxidosqualene is the precursor of terpenes, which can be catalyzed into different terpenes by different 2, 3-oxidosqualene cyclases (OSCs). Plant OSCs provide a number of options for their catalytic control. A number of OSCs have been cloned and their functions have been confirmed, including β-amyrin synthase (β-AS) from Panax ginseng(Yang et al., 2012), Medicago truncatula (Iturbe-Ormaetxe et al., 2003), Glycyrrhiza glabra (Hayashi et al., 2001), and Pisum sativum (Morita et al., 2000), lupeol synthase (LUS) from Arabidopsis thaliana (Herrera et al., 1998) and Withania somnifera (Dhar et al., 2014), cycloartenol synthase (CAS) from G. glabra (Hayashi et al., 2000) and Costus speciosus (Kawano et al., 2002), Dammarenediol-II synthase (DS) from P. ginseng (Tansakul et al., 2006), lanosterol synthase (LS) from Siraitia grosvenorii (Dai et al., 2015), and multifunctional triterpene synthases from P. sativum, which catalyze the formation of α-amyrin, β-amyrin, and lupeol (Morita et al., 2000). These OSCs, including β-AS, α-AS, LUS, DS, LS, and CAS, finally lead to the production of different triterpene saponins, such as oleanane-type triterpene saponin, bearberry hexane-type triterpene saponin, lupinane-type triterpene saponin, dammarane-type triterpene saponin, lanostane-type triterpene saponin, and sterol (Fig. 1).

β-AS is responsible for the production of oleanane-type triterpene saponin, which is widely present in Leguminosae (Ali et al., 2016), Araliaceae (Gao et al., 2015), Umbelliferae (Wu et al., 2012), and other families (Wang et al., 2011). Among these families, there are many famous medicinal plants, such as Glycyrrhiza uralensis, Astragalus membranaceus, Isatis indigotica, P. ginseng, and Panax quinquefolium. Oleanane-type triterpene saponin is extensively regarded as the marker active compound in these medicinal plants. For example, glycyrrhizic acid (GA), the
marker compound in *G. uralensis*, possesses antitumor (Li et al., 2014), anti-inflammatory (Pang et al., 2016), antiviral (Baltina et al., 2015), and immune-regulation activities (Asl et al., 2008). Ginsenosides, the most important active components in *P. ginseng* and *P. quinquefolium*, possess immune regulation (Zhang et al., 2015), antiinflammatory (Chen et al., 2008), antiaging (Lee et al., 2012), and antiviral (Song et al., 2014) activities. Oleanolic acid, the most important compound in *Svertyia lucidae* and *Ligustrum lucidum Ait.*, possesses liver protection (Kim et al., 2004), antitumor (Yan et al., 2010), and anti-inflammatory (Bednarczyk-Cwynar et al., 2016) activities. β-AS is a key enzyme for the production of these important triterpene saponins. Studies of β-AS can help us to deeply parse their biosynthetic pathway and improve their accumulation.

Attracted by the key position of β-AS, we conducted a series of studies about β-AS gene in *G. uralensis*, one of the most frequently used Chinese herbs. We cloned a 4109 bp β-AS full-length DNA sequence (Chen et al., 2013) and a 2289 bp β-AS full-length cDNA sequence from *G. uralensis* (Shen et al., 2009), and we also verified its function in *Saccharomyces cerevisiae*. We investigated the influences of co-expression of β-AS gene and another functional gene involved in MV A pathway, squalene synthase (*SQS*) gene, in *S. cerevisiae* and found that the co-expression enhanced the accumulation of β-amin (Liu et al., 2014). And we also revealed the temporal and spatial specificity of the expression of β-AS gene in *G. uralensis*, and found that the root tip was the suitable plant material and May, June, August and September were the right acquisition time (Liu et al., 2012). However, after finishing the above researches about *G. uralensis* β-AS gene, we find there are still so many questions and puzzles to be resolved. For example, single nucleotide polymorphisms (SNPs), insertion-deletion length polymorphism (InDel) and copy number variations (CNVs) are present in β-AS genes from *G. uralensis*, but we are not clear how they influence the GA accumulation. And there is a pair of GA epimer, differed only in the C18-H, the formation of which are influenced by β-AS gene, but we don’t know how it works. So we decide to deeply and comprehensively analyze and recognize the β-AS genes in higher plants, and which is the original cause of this paper.

In this paper, we obtained 475 DNA sequences, 220 cDNA sequences, and 99 amino acid sequences of β-AS registered in NCBI database by Oct, 2016. We investigated typical β-AS DNA sequences in 15 different species from 9 families, mRNA sequences in 37 species from 20 families, and the corresponding amino acid sequences using all kinds of bioinformatics online tools and softwares such as ExPASy Proteomic tools, DNAMAN 6.0.3.99, and MEGA 5.0. The sequence alignment, physicochemical properties, functional domains, and evolutionary relationship were investigated. To get the latest and comprehensive information of β-AS, more than 300 papers were searched and 80 of them were reviewed. Pub Med, Web of Science, Science Direct, and Research Gate, were information sources through the search terms of “β-amyrisynthase”, “biosynthesis”, “OSC”, and their combinations, mainly from year 2010 to 2016. Studies were selected from Science Citation Index journals. All the references linked to the registered DNA and cDNA sequences in NCBI database were also reviewed. Furthermore, the latest technology and the applications of β-AS in focused plants were also summarized and discussed. Hopefully, we wish this paper could lay a foundation for further studies of β-AS and other OSCs.

**Bioinformatics analysis of β-AS**:
The full-length DNA, mRNA, and amino acid sequences of β-AS were analyzed using online bioinformatics tools (http://www.ncbi.nlm.nih.gov) with the time restriction of Oct. 2016. The deduction of the amino acid sequences, calculation of theoretical molecular mass and pl, were performed with ExPASy Proteomic tools provided at http://www.expasy.ch/tools/. Conserved domains in β-AS were detected using Conserved Domain Database search tool (CDD) on NCBI server (http://www.ncbi.nlm.nih.gov/structure/cdd/ wrpsb.cgi). A multiple alignment of amino acid sequences was performed with DNAMAN 6.0.3.99. Phylogenetic tree was constructed using MEGA 5.0. by Neighbor-Joining (N-J) method and reliability of nodes has been tested with 1000 bootstrap replicates.

**Bioinformatic analysis of β-AS DNA sequences**:β-AS DNA sequences of 15 different species from 9 families have been recorded in NCBI. The length of these β-AS DNA sequences ranges from 3900 bp to 8800 bp. Using DNAMAN 6.0.3.99 the consistency of them is determined to be 35.86%. The phylogenetic tree (Fig. 2) shows that Fragaria vesca L., Prunus mume Sieb. Pyrus brentschieleri Rehder, Malus pumila Mill. from family Rosaceae, and Camelina sativa (L.) Crantz, Arabidopsis thaliana (L.) Heynh. from family Cruciferae are clustered into one branch. Lycopersicon esculentum Mill, Nicotiana tomentosiformis, and Nicotiana sylvestris from family Solanaceae, are clustered into one branch. And Glycine max (L.) Merr. and Cicer arietinum L. from family Leguminosae, are clustered into one branch. The N-J tree analysis results are basically in line with genetic relationship, β-AS DNA sequences from different families are separated into different branches, respectively.

**Bioinformatic analysis of β-AS mRNA sequences**:β-AS mRNA sequences of 37 species from 20 families have been recorded in NCBI. The length of these β-AS mRNA sequences ranges from 2100 bp to 2900 bp. The consistency of them is 63.38%. As the N-J tree (Fig. 3) shows, seven species from family Leguminosae including *G. glabra*, *G. uralensis*, *Lotus corniculatus* L. var. japonicus Regel, *Vigna radiata* (L.) Wilczek, *G. max*, *Cicer arietinum* L., and *Pisum sativum* L. are aggregated. *Jatropha curcas* L. and *Euphorbia tirucalli* L. both from family Euphorbiaceae are clustered together. *F. vesca*, *P. mume*, *P. brentschieleri*, and *P. pumila* from family Rosaceae are clustered into one branch. Furthermore, *Aralia elata*, *P. ginseng*, *P. quinquefolium*, and *Panax japonicas* from family Araliaceae are gathered together. *Bupleurum chinense* DC. from family Umbelliferae is also clustered into this branch, which indicates that *B. chinense* DC. has a close relationship with Araliaceae plants. In addition, *Barbara vulgaris* R. Br., *C. sativa*, and *A. thaliana* from family Cruciferae are clustered together. *Centella asiatica* (L.) Urban from family Umbelliferae and *Bacopa monnieri* (L.) Wettst. from family Scrophulariaceae are gathered together. Except individual examples, the N-J tree analysis results are basically in line with genetic relationship, β-AS mRNA sequences from different families are also clustered into different branches.
**Fig. 1.** The MVA pathway for triterpene biosynthesis ($\beta$-AS is marked in red).
Fig. 2. The cluster tree based on 15 β-AS full-length DNA sequences.

Fig. 3. The cluster tree based on 37 β-AS full-length mRNA sequences.
AMYRIN SYNTHASE, ONE OF THE MOST IMPORTANT KEY ENZYMES FOR FORMATION IN PLANTS.

Fig. 4. The N-J tree of β-AS amino acid sequences. (Light green represents Leguminosae, pink represents Asteraceae, red represents Araliaceae, yellow represents Cruciferae, and other colors represent different families, respectively.)

**Bioinformatics analysis of β-AS amino acid sequences:** β-AS amino acid sequences of 32 species from 29 genera, 20 families have been recorded in NCBI. The similarity of these sequences is 81.62%. The N-J tree is showed in Fig. 4, species from the same family are marked in the same colour. Seven species from family Leguminosae including G. uralensis, G. glabra, G. soja, G. max, P. sativum, M. truncatula, and L. japonicas (marked in green) are clustered together. K. septemlobus, A. elata, P. ginseng, P. quinquefolius, and P. japonicas (marked in red) from family Araliaceae are clustered into the same branch. Artemisia annua and Aster sedifolius (marked in pink) from family Composite are clustered together. A. longiligumis from family gramineae, C. borivilianum from family Liliaceae and N. sativa from family Ranunculaceae are separated into different branches. N-J tree is consistent with botanical classification status, and accorded with genetic evolution rule.

**Physicochemical properties analysis of β-AS:** The basic physicochemical properties of β-AS, including number of amino acid residues, molecular weight, isoelectric point, half-life period, instability parameters, and average hydrophilic coefficient are listed in Table 1. β-AS is composed of about 760 amino acid residues with the molecular mass of 87 kDa or so. The isoelectric points range from 5.70 to 6.30. Theoretically, the half-life period is consistent, it is 30 hours in mammalian reticulocytes In vitro, and exceed 20 hours in yeast and 10 hours in E. coli In vivo. The average coefficient of hydrophilic indicates that these β-AS are consistently hydrophobic.

**Sequence alignment analysis of β-AS amino acid sequences:** To have a better understanding of the amino acid sequence characteristics of β-AS, 12 representative β-AS amino acid sequences from family Leguminosae, Araliaceae, Umbelliferae, Composite, Brassicaceae, and Polygalaceae have been selected for further alignment analysis using DNAMAN 6.0.3.99, since these selected species have attracted much attention depending on their high-level triterpene contents and remarkable pharmacological activities (Chinese Pharmacopoeia Commission, 2015). The consistency of these sequences is 81.57%. Fig. 5 shows the similarity of the 12 sequences. 100%, 75%, 50%-75%, and less than 30% similarity are marked in black, pink, blue, and white, respectively. The similarity between P. quinquefolius and P. japonicas is highest (98.69%), while the similarity between A. longiligumis and P. quinquefolius is lowest (49.15%). Using SOPMA, the secondary structure of the above three β-AS sequences, P. quinquefolius, P. japonicas, and A. longiligumis was investigated. The results are given in Fig. 6. They are all composed of approximate 40% α-helices, 17% extended strand, 33% random coil, and 10% β-turns.
Fig. 5. Analysis of 12 β-AS amino acid sequences. (Black represents the similarity of the amino acid sequences is 100%, pink represents 75%, blue represents 50~75%, and white represents less than 30%. The functional area of β-AS is marked in red. The blue frame shows Gln-Trp (QW) motif, the red frame shows Asp-Cys-Thr-Ala-Glu (DCTAE) motif, and the yellow frame shows Met-Trp-Cys-Tyr-Cys-Arg (MWCYCR), respectively.)
B-AMYRIN SYNTHASE, ONE OF THE MOST IMPORTANT KEY ENZYMES FOR FORMATION IN PLANTS.

Fig. 6. The prediction of β-AS secondary structure. (a, b, and c show the β-AS secondary structure of P. quinquefolius, P. japonicas, and A. longiglumis, respectively. Blue represents α-helices, red represents extended strand, green represents β-turn, and purple represents random coil.)

Domain structure prediction: The highly conserved motifs of β-AS play the key role in substrate binding and protonation. Asp-Cys-Thr-Ala-Glu (DCTAE) implicated in substrate binding motif has been reported to be associated with β-amyrin specificity (Vishwakarma et al., 2013). Recently, mutation study of Euphorbia tirucalli OSC revealed that DCTAE was a putative initiation site for the polycyclization reaction (Ito et al., 2013). Kushiro et al. (Kushiro et al., 1998) illustrated that the tryptophan residue in Met-Trp-Cys-Arg (MWCYCR) and Asp-Cys-Thr-Ala-Glu (DCTAE) motifs of β-AS played a significant role in formation of cyclic backbone. Furthermore, Gln-Trp (QW) motifs are indicated by solid bars, which properties of the OSC superfamily (Poralla et al., 1994), and may strengthen the structure of the enzyme and stabilize the carbocation intermediates during cyclization. Using ProtParam and NCBI conserved domain search tool, it is determined that the fragments from 125th to 735th amino acid residue is the active site cavity of β-AS (Fig. 7a), and the fragment from 604th to 618th residues (DGSWYGNWGVCFTYG) is the conserved domain (Fig. 7b).

In Figure 5, DCTAE motif is framed in green, QW motifs are framed in blue, MWCYCR motif is framed in yellow, and the 604th-618th residues are framed in red. DCTAE motif from 491-495 and QW motif from 163-170 are 100% conserved, which confirmed the key positions of these two motifs. In the comparison among P. quinquefolius, P. japonica, and A. longiglumis from 604th to 618th residues, there are three mutation sites, 611th, 614th, and 617th. There are S, I, and A in A. longiglumis, while N, V, and T in P. quinquefolius and P. japonica, respectively. The MWCYCR motif of P. quinquefolius and P. japonica are same with each other while A. longiglumis are quite different. In the QW motifs from 728 to 752, A. longiglumis shows the most obvious difference with other species, which is correlated with the genetic distance. Since A. longiglumis belong to monocotyledon, while the other 11 species are from dicotyledon.
Table 1. The basic physicochemical properties of β-AS amino acid sequences.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession number</th>
<th>Number of amino acids</th>
<th>Molecular weight</th>
<th>Isoelectric point</th>
<th>Instability parameters</th>
<th>Hydrophilic coefficient</th>
<th>Half-life period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine max</td>
<td>AAM21264.1</td>
<td>739</td>
<td>84603.6</td>
<td>6.10</td>
<td>48.60</td>
<td>-0.342</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Glycyrrhiza uralensis</td>
<td>ADE88148.1</td>
<td>762</td>
<td>87072.7</td>
<td>6.19</td>
<td>47.77</td>
<td>-0.286</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Avena longiglumis</td>
<td>AAT38895.1</td>
<td>757</td>
<td>86858.9</td>
<td>6.12</td>
<td>45.03</td>
<td>-0.336</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Euphorbia tirucalli</td>
<td>BAEE3642.1</td>
<td>762</td>
<td>87589.5</td>
<td>6.78</td>
<td>47.35</td>
<td>-0.320</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Artemisia annua</td>
<td>ACA13386.1</td>
<td>761</td>
<td>87503.2</td>
<td>5.87</td>
<td>49.24</td>
<td>-0.331</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Bupleurum chinense</td>
<td>ABY90140.2</td>
<td>764</td>
<td>87791.2</td>
<td>6.02</td>
<td>49.99</td>
<td>-0.374</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Panax ginseng</td>
<td>AGC09939.1</td>
<td>761</td>
<td>87775.2</td>
<td>5.92</td>
<td>58.44</td>
<td>-0.368</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Panax japonicus</td>
<td>AKN23431.1</td>
<td>761</td>
<td>87900.4</td>
<td>5.84</td>
<td>49.38</td>
<td>-0.369</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Vaccaria hispanica</td>
<td>ABK76265.1</td>
<td>760</td>
<td>87521.3</td>
<td>5.89</td>
<td>45.91</td>
<td>-0.312</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Barbarea vulgaris</td>
<td>AFF27506.1</td>
<td>762</td>
<td>87503.4</td>
<td>6.29</td>
<td>47.91</td>
<td>-0.271</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
<tr>
<td>Polygala tenuifolia</td>
<td>ABL07607.1</td>
<td>762</td>
<td>87343.0</td>
<td>6.21</td>
<td>46.50</td>
<td>-0.308</td>
<td>&gt;10h (E, in vivo); 30m (E, in vitro); &gt;20y (E, in vivo)</td>
</tr>
</tbody>
</table>

Table 2. The registered β-amyrin synthetase sequence information in GenBank.

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>GenBank accession number</th>
<th>Sequence length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liliaceae</td>
<td>Chlorophytum borivilianum</td>
<td>KM245582.1</td>
<td>2277 bp</td>
</tr>
<tr>
<td>Euphorbiaceae</td>
<td>Jatropha curcas</td>
<td>XM_012232059.1</td>
<td>2307 bp</td>
</tr>
<tr>
<td>Liliaceae</td>
<td>Lotus japonicus</td>
<td>AF478455.1</td>
<td>2458 bp</td>
</tr>
<tr>
<td>Leguminosae</td>
<td>Glycyrhiza glabra</td>
<td>AB037203.1</td>
<td>2671 bp</td>
</tr>
<tr>
<td>Leguminosae</td>
<td>Glycyrrhiza uralensis</td>
<td>FJ627179.1</td>
<td>2289 bp</td>
</tr>
<tr>
<td>Leguminosae</td>
<td>Vigna radiata</td>
<td>XM014667164.1</td>
<td>2647 bp</td>
</tr>
<tr>
<td>Leguminosae</td>
<td>Cicer aritinum</td>
<td>XM_004488858.2</td>
<td>2672 bp</td>
</tr>
<tr>
<td>Poaceae</td>
<td>Oryza sativa Japonica</td>
<td>KC416147.1</td>
<td>2262 bp</td>
</tr>
<tr>
<td>Poaceae</td>
<td>Braguiera gymnorrhiza</td>
<td>AB289585.1</td>
<td>2280 bp</td>
</tr>
<tr>
<td>Poaceae</td>
<td>Sesamum orientale Linn</td>
<td>XM_011095493.1</td>
<td>2573 bp</td>
</tr>
<tr>
<td>Betulaceae</td>
<td>Betula platyphylla</td>
<td>AB055512.1</td>
<td>2519 bp</td>
</tr>
<tr>
<td>Malvaceae</td>
<td>Gossypium raimondii</td>
<td>XM_012592029.1</td>
<td>2712 bp</td>
</tr>
<tr>
<td>Asteraceae</td>
<td>Aster ticiatars</td>
<td>EU330197.1</td>
<td>2286 bp</td>
</tr>
<tr>
<td>Gentianaceae</td>
<td>Gentiana straminea</td>
<td>FJ790411.1</td>
<td>2286 bp</td>
</tr>
<tr>
<td>Ranunculaceae</td>
<td>Nigella sativa</td>
<td>FJ013228.1</td>
<td>2430 bp</td>
</tr>
<tr>
<td>Ranunculaceae</td>
<td>Fragaria × ananassa Duch.</td>
<td>XM_004305745.2</td>
<td>2900 bp</td>
</tr>
<tr>
<td>Solanaceae</td>
<td>Solanum lycopersicum</td>
<td>NM_001247675.1</td>
<td>2535 bp</td>
</tr>
<tr>
<td>Solanaceae</td>
<td>Withania somnifera</td>
<td>HQ266579.1</td>
<td>2289 bp</td>
</tr>
<tr>
<td>Solanaceae</td>
<td>Bupleurum chinense</td>
<td>HQ166837.1</td>
<td>2602 bp</td>
</tr>
<tr>
<td>Umbellifera</td>
<td>Centella asiatica</td>
<td>AY520818.1</td>
<td>2562 bp</td>
</tr>
<tr>
<td>Umbellifera</td>
<td>Arabidopsis thaliana</td>
<td>AB374428.1</td>
<td>2280 bp</td>
</tr>
<tr>
<td>Umbellifera</td>
<td>Barbarea vulgaris</td>
<td>JQ172795.1</td>
<td>2289 bp</td>
</tr>
<tr>
<td>Crucifera</td>
<td>Camelina sativa</td>
<td>XM_010431459.1</td>
<td>2392 bp</td>
</tr>
<tr>
<td>Caryophyllace</td>
<td>Vaccaria hispanica</td>
<td>DQ915167.1</td>
<td>2511 bp</td>
</tr>
<tr>
<td>Sterculiaceae</td>
<td>Theobroma cacao</td>
<td>XM_007023302.1</td>
<td>2909 bp</td>
</tr>
<tr>
<td>Sterculiaceae</td>
<td>Chinese aralis</td>
<td>HM219225.1</td>
<td>2292 bp</td>
</tr>
<tr>
<td>Araliaceae</td>
<td>Panax ginseng</td>
<td>AB009030.1</td>
<td>2589 bp</td>
</tr>
<tr>
<td>Scrophulariace</td>
<td>Bacopa monnieri</td>
<td>HM769762.1</td>
<td>2765 bp</td>
</tr>
<tr>
<td>Salicaeae</td>
<td>Populus trichocarpa</td>
<td>XM_002310313.2</td>
<td>2682 bp</td>
</tr>
<tr>
<td>Polygalaceae</td>
<td>Polygala tenuifolia</td>
<td>EBI07623.1</td>
<td>2934 bp</td>
</tr>
</tbody>
</table>
Transmembrane structure and signal peptide prediction: The transmembrane domain location analysis indicates that P. quinquefolius and P. japonicas have no transmembrane region, while A. longiligum has two transmembrane regions, 117th-141th and 606th-634th amino acid residues. It shows that β-AS of P. japonicas and P. quinquefolius do not cross the membrane, while β-AS of A. longiligum is "anchored" to a specific site in the cytoplasmic matrix to perform the catalytic function. Signal peptide analysis indicates that the C-score, S-score and Y-score are the same in P. quinquefolius and P. japonicas. The highest score of cleavage site is 0.111 at 47th amino acid residue, the highest score of signal peptide site is 0.134 at 39th amino acid residue, and the highest score of combined cleavage site is 0.111 at 47th amino acid residue. In A. longiligum, the highest score of original cleavage site is 0.111 at 46th amino acid residue, the highest score of signal peptide site is 0.236 at 1st amino acid residue, and the highest score of combined cleavage site is 0.139 at 11th amino acid residue. The above demonstrates that β-AS of P. quinquefolius, P. japonicas, and A. longiligum all have no signal peptides, and they exert their activities in cytoplasm. This inference is consistent with the action site of MVA pathway in cell.

The three-dimensional protein model analysis of β-AS: The three-dimensional protein models of P. quinquefolius, P. japonicas, and A. longiligum were also determined. As shown in Fig. 8, the model of P. quinquefolius and P. japonicas are quite similar, while A. longiligum showed moderate diversity.

The research progress of genetic engineering

Gene cloning and functional verification of β-AS: Up to date, 60 β-AS cDNAs have been cloned from 37 species in 21 families as listed in Table 2. In most studies (Iтуре-Ормаэткс и др., 2003; Kim et al., 2005), the conserved regions of the β-AS synthase were used to design specific primers and amplify the target sequences through RT-PCR or RACE methods. The amplified products were cloned into pGEM-T Easy vector and transformed into disarmed DH5α E. coli cells. For B. gynorrhiza, B. platyphilla, L. esculentum, P. tenuifolia, and G. glabra, the full-length cDNAs were cloned into yeast expression vector pYES2 under the control of the GAL10 promoter, and then obtained plasmids were introduced into a triterpenoid synthase-deficient yeast mutant GIL77, which led to the production of β-amyrin (Vishwakarma et al., 2013; Kirby et al., 2008; Hayashi et al., 2001). For A. apiacea, B. vulgaris, S. vaccaria, and A. thaliana, the full-length cDNAs were introduced into the high-copy yeast expression vector pESC-URA under control of the GAL10 promoter, and expressed in S. cerevisiae (Sun et al., 2013; Wang et al., 2011).

β-AS expression studies: RNA interference was used to analyze the function of β-AS involved in ginsenoside biosynthesis, it was found that down-regulation of β-AS expression resulted in reducing levels of β-amyrin and oleanane-type ginsenoside and increasing level of dammarane-type ginsenoside. Since there are many OSCs led to different triterpene formation, the regulation of related cyclase is also important. Depending on a research conducted by Zhang F (Zhang et al., 2014), the expression levels of SQS, squalene monooxygenase (SQE), and β-AS were highly correlated, which suggested that overexpression of coded gene, such as SQS, SQE, and β-AS increased the production of triterpene saponin. In another research, two key enzymes involved in sterol pathway in S. cerevisiae, HMGR and lanoster synthase, were manipulated to increase triterpene production. It was found that β-amyrin production was improved by 50%, and squalene level had a 12-fold increase, which indicated that a high expression level of LUS and CYS increased the pathway into lupeol and phytosterol synthesis, rather than oleanolic acid synthesis. Therefore, it appeared that an increase of oleanolic acid production required an elevated level of β-AS, SQS and SE expression, and a suppression of LUS (Mangas et al., 2008).

Furthermore, the temporal and spatial specificity of the expression of β-AS has also been investigated. Hуrе-Ормаэткс I et al. found the expression pattern of β-AS differed in different plant tissues. β-AS gene in different tissues of M. truncatula was analyzed, the highest transcript levels was found in the shoot meristem and stem tissue (Iтуре-Ормаэткс et al., 2003). And we also revealed the temporal and spatial specificity of the expression of β-AS gene in G. uralensis, and found that the root was the suitable plant material and May, June, August and September were the right acquisition time (Liu et al., 2012).
Exogenous stimulator regulation: MeJA (methyl jasmonate MJ) was a commonly used chemical inducers for promoting plant cell secondary metabolite biosynthesis. It induced the defense responses and plant protection element (mainly flavonoids and terpenoids) alone or synergistically in plants. As one of the important substances in cell signal transduction system, MeJA is widely involved in the process of plant defense signal transduction and amplification, which induces the expression of anti-reaction products (Mitra & Baldwin, 2014; Wu et al., 2008; Schlogl et al., 2008; Li et al., 2014).

Hayashi H et al. (Hayashi et al., 2004) found in the cultured cells, the addition of MeJA up-regulated β-AS mRNA expression level and soyasaponin biosynthesis, but down-regulated lupeol synthase expression level. While, the addition of gibberellins down-regulated β-ASm RNA expression level. Suzuki & Dixon (2005) explored cell suspension cultures of M. truncatula to MeJA, which led to a 50-fold induction of β-AS. Their work established Medicago cell suspension cultures as a well model for future genomics approaches to explore the regulation of legume secondary metabolism. Also, Confalonieri M et al. (Confalonieri et al., 2009) found that β-ASmRNA expression level was up-regulated and the accumulation of oleanolic acid increased in G. scabra by treatment with MeJA over a period from 6 hours to 10 days (Nasrollahi et al., 2014).

The effect of drought stress on β-AS gene expression in G. glabra has been studied, and the results showed that drought was conducive to β-AS expression, thus contributing to the accumulation of glycyrrhizin (Nasrollahi et al., 2014). Basyuni M et al. (Knott & Reynolds, 1990) found salt was beneficial to β-AS expression in mangrove plants Kandelia candel and B. gymnorrhiza. These results enhanced triterpenoids in the adaptation of mangroves to endure salt or water stress.

Transgenic applications: So far, M. truncatula can be genetically transformed and regenerated, and which provides a possibility to regulate the content of triterpene saponin in transgenic plants by ectopic expression, overexpression or gene silencing. Agrobacterium-mediated transformation was used to introduce a novel Aster sedifolius β-AS in M. truncatula with the cauliflower mosaic virus 35S promoter, the transgenic plants exhibited greater amounts of triterpenic compounds than control plants. P. japonicus is a rare Chinese herb containing ginsenosides as its main active ingredient. Plants cannot produce ginsenosides because it lacks a key rate-limiting enzyme, β-AS. However, it can produce a secondary metabolite, 2, 3-oxidosqualene, which is a precursor for ginsenoside biosynthesis. β-AS gene from P. japonicus was transformed into rice cultivar using an Agrobacterium-mediated approach, and 68 rice transgenic plants was got. Real-time PCR and Western blotting analyses showed that β-AS gene overexpressed in rice. HPLC analysis showed that the concentration of oleanane-type sapogenin oleanolic acid in transgenic rice was 8.3~11.5 mg/100 g dw (Huang et al., 2015).

Discussion

This paper reviewed the research progress of β-AS, a key enzyme for triterpene skeleton formation. All of the registered DNA, cDNA, and amino acid sequences of β-AS in GenBank have been gathered, the physicochemical, hydrophilic and hydrophobic properties, secondary structures, and functional domains have been investigated. The evolutionary relationships between different species have also been analyzed in addition, the experimental methods to obtain β-AS gene sequences and verify their functions have been summarized. We found that OSCs play a crucial role in terpene formation. There are several kinds of OSCs, such as β-AS, α-AS, LUS, DS, LS, and CS, which suggest that a high expression level of one OSCs may down-regulate others. Therefore, multi-gen control may be a promising destination for the certain product. It is helpful to improve the content of target products by leading metabolites to the certain direction we need. What’s more, both the inside (mRNA sequences) and outside factors (host and production system/conditions, drought stress, and exogenous stimulation) can affect the content of secondary metabolites.

Triterpenes are extensively distributed isoprenoids found in many different kinds of organisms and form the one of main categories of natural products. They have multiple functions in biosystem. Because of diversity of triterpenes constituted in higher plants, it is interesting to confirm how different triterpene synthases controls the product specificity during cyclization of oxidosqualene. In addition, studies have revealed mechanisms that lead to product specificity by comparing the cloning and sequence of several different triterpene synthase cDNAs, such as the folded geometry of the substrate and the end point of the resulting cation.

Furthermore, the improvement of qualitative and quantitative techniques of medicinal plants is the key to the stable supply of crude drugs. An excessive production of triterpenoid saponins in medicinal plants is a solution to this purpose. The introduction and overexpression of the triterpene synthase gene in a suitable host plant should be the primary task to this target.

Competing interests: All authors declare that they have no competing interests.

Acknowledgement

This work was supported by National Science Foundation of China (81503181).

References


B. AMYRIN SYNTHASE, ONE OF THE MOST IMPORTANT KEY ENZYMES FOR FORMATION IN PLANTS...


Liaw, C.C., H.C. Huang, P.C. Hsiao, L.J. Zhang and Z.H. Lin. 2015. 5beta, 19-nortriterpenoids with tricyclo [6.3.0.0 (2,11)]undecane skeleton from Xanthium strumarium L. Molecules, 19(9): 12898-12908.


AMYRIN SYNTHASE, ONE OF THE MOST IMPORTANT KEY ENZYMES FOR FORMATION IN PLANTS


