Molecular cloning, functional characterization and expression of the β-amyrin synthase gene involved in saikosaponin biosynthesis in Bupleurum chinense DC.

Yanping Mao1,3 · Hua Chen1 · Jun Zhao1 · Yuchan Li1 · Liang Feng1 · Yuping Yang1 · Yiguan Zhang2 · Ping Wei2 · Dabin Hou1

Received: 24 February 2022 / Accepted: 23 July 2022 / Published online: 16 September 2022
© The Author(s), under exclusive licence to Society for Plant Biochemistry and Biotechnology 2022

Abstract

Bupleurum chinense DC. is a commonly used plant in traditional Chinese medicine, and saikosaponins (SSs) are the main active oleanane-type triterpene saponins in B. chinense. β-Amyrin synthase (β-AS) is an important enzyme in oleanane-type triterpenoid saponin synthesis, but its role in saikosaponin synthesis has rarely been studied. Here, the putative β-AS gene BcBAS1 (Accession No. ON890382) selected according to metabolomic and transcriptomic analyses was cloned and functionally characterized by heterologous expression in Escherichia coli and Pichia pastoris, and its subcellular localization and expression patterns were examined. The molecular weight of the BcBAS1 recombinant protein was approximately 87 kDa, and this protein could catalyse the production of β-amyrin, the precursor of SSs. Furthermore, BcBAS1 was located in the cytosol, and relative expression in four tissues of the four genotypes was positively correlated with SSA and SSD contents. Our results indicate that BcBAS1 is a β-AS gene and may play an important role in saikosaponin biosynthesis and regulation. This study sheds light on the role of β-AS genes in the synthesis of SSs and provides insights for the metabolic engineering of SSs.

Keywords Bupleurum chinense DC. · β-AS · Saikosaponin · Gene cloning and expression · Characterization and function

Introduction

Radix Bupleuri, the dried root of Bupleurum chinense DC. and B. scorzonerifolium Willd., is an important medicinal herb in Eurasia and North Africa for the treatment of fever, pain, inflammation (Ma et al. 2016), chronic hepatitis (Zhao et al. 2012), depression (Li et al. 2017), and cancer (Zhao et al. 2019). B. chinense DC. and B. scorzonerifolium Willd. are officially listed in the Chinese Pharmacopoeia (Chinese Pharmacopoeia Commission 2020), and B. chinense DC. is considered the best source of Radix Bupleuri due to its long and thick roots, thin root skin and few branching roots (Jiang et al. 2020). Radix Bupleuri is also widely used in traditional Chinese medicine and is included in more than 1900 prescriptions. In particular, a lung cleansing and detoxifying decoction (Qing Fei Pai Du Tang) with Radix Bupleuri as the main component has been proven effective in patients with mild and ordinary COVID-19 (Cao et al. 2020). The active components of Radix Bupleuri are complex and include essential oils, flavonoids, coumarins, lignans, triterpene saponins, alkaloids and polyacetylenes (Ashour and Wink 2011).

SSs, including SSA, SSb, SSc and SSD (Huang et al. 2009), are the main active components of Radix Bupleuri and are unique to plants in the genus Bupleurum (Hirai et al. 2004). The total content of SSs in the dry roots of Bupleurum plants can reach 7% (Ashour and Wink 2011). SSs are penta cyclic...
triterpenoid saponins, and triterpenoids, as plant secondary metabolites, are among the most abundant natural products in many plants. The biosynthesis of triterpenoids depends on the isoprenoid pathway (Moses et al. 2014). The biosynthesis of triterpenoids begins with the production of 2,3-oxidosqualene and 2,3-oxidosqualene cyclized to different forms of triterpenoid skeletons catalyzed by oxidosqualenecyclases (OSCs). The triterpenoid skeletons are hydroxylated or oxidized by cytochrome P450 enzymes (P450s) and finally glycosylated by uridine diphosphate glycosyltransferases (UGTs), leading to different triterpene saponins (Sui et al. 2021).

OSC s are widely considered to play crucial roles in the production of different types of triterpenoid saponins, and they are also defined as important branch points regulating the synthesis of different triterpenoid saponins (Basyuni et al. 2007). β-AS is a member of the OSC family, which is responsible for the production of oleanolic acid saponin skeletons. Currently, more than 1600 sequences of β-AS have been registered in the GenBank database, and many of them have been cloned and functionally characterized in many plant species (Ma et al. 2018b). However, only a few studies have been performed on β-AS genes in plants of the genus Bupleurum. There are only two β-AS genes, called bcAS1 (Gao et al. 2015) and BcBAS (Li et al. 2020), with fully cloned sequences from B. chinense DC., and BcBAS has been further expressed in Saccharomyces cerevisiae to analyze its function. In addition, the subcellular localization and expression patterns of β-AS in different genotypes of B. chinense DC. are still unclear. Therefore, it is necessary to investigate the cloning, characterization, subcellular localization and expression patterns of β-AS genes in Bupleurum plants.

In our previous study, 24 β-AS unigenes were identified by integrative analysis of the metabolome and transcriptome. These 24 β-AS unigenes potentially represented four nonredundant β-ASs (BcBAS1, BcBAS2, BcBAS3, and BcBAS4) that were selected by BLAST, and Bc61215 (BcBAS1) was significantly correlated with SSd and SSB (He et al. 2021). In this study, we obtained the full-length sequence of BcBAS1 and expressed it in E. coli DH5α. The subcellular localization in protoplasts of Arabidopsis leaves and the function inferred using a heterologous expression system in Pichia pastoris GS115 were also investigated. We further explored the expression patterns of BcBAS1 in four genotypes of B. chinense DC. for the first time and analyzed the correlation between the expression of this gene and the contents of SSS. The results indicated that BcBAS1 is a β-AS gene and catalyzes only the formation of β-amyrin. BcBAS1 may be the key enzyme in the synthesis of SSS, and our research may provide insights for the metabolic engineering of SSSs.

Materials and methods

Plant materials

The plant materials in this study were four genotypes of B. chinense DC., namely, CBC1, CBC3, RX and FS. Seeds of these genotypes were sown at the beginning of March 2020 and planted in the experimental field of the Medicinal Plant Research Laboratory, Southwest University of Science and Technology in Mianyang (31°32´ N and 104°42´ W), Sichuan Province of China. Fresh roots, stems, leaves and flowers of B. chinense DC. were collected at the full blooming stage in July 2020. All samples were frozen immediately with liquid nitrogen and stored at -80°C until further use.

RNA isolation, cDNA synthesis and gene cloning

Total RNA was extracted from different tissues of the four genotypes of B. chinense DC. using an RNA Prep Pure Plant Kit (Tiangen, China) according to the manufacturer’s instructions. Each RNA extraction was performed for mixed samples of 3 individual plants. First-strand cDNA was synthesized from 1000 ng of total RNA using All-in-One First-Strand cDNA Synthesis SuperMix (TransScript, China) following the manufacturer’s instructions. Full-length BcBAS1 sequences were taken from the B. chinense DC. transcriptome library, including from four different tissues (root, leaf, stem and flower), for a total of 93486 unigenes (He et al. 2021). The open reading frame (ORF) of BcBAS1 was located using ORF Finder, and specific primers for amplification of BcBAS1 were designed by Primer 5.0 software according to its full-length sequence. Taking cDNA as the template, PCR amplification of BcBAS1 was carried out by using a FastPfu DNA Polymerase Kit (TransScript, China). The PCR conditions were 95°C for 2 min, followed by 35 cycles of 95°C for 20 s, 45°C for 20 s and 72°C for 70 s and then 72°C for 5 min. The PCR products were examined by agarose gel electrophoresis, purified and ligated into the pEASY-T1 Simple Cloning Vector, and transformed into E. coli DH5α cells. Finally, positive colonies were selected for sequencing. The primers used in this study are listed in Table 1.

Sequence and phylogenetic analyses

The full-length sequence of BcBAS1 was analyzed using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) along with the encoded amino acid sequences and homology. The amino acid multiple sequence alignment was determined by DNAMAN software. The predicted amino acid sequences of BcBAS1 - BcBAS4 and those of other plants obtained from the GenBank database (https://www.
**Table 1** List and sequences of primers used in this study

| Primer code | Primer Sequence (5’→3’) | Usage |
|-------------|-------------------------|-------|
| BCβAS1-F    | CGTGAAGGTTGTGGTGAAGAAAAAA | cDNA amplification |
| BCβAS1-R    | TTGTGTAATCCGCTGTTGAAGAAAATTCG | |
| BCβAS1-F(SL) | GAGGATCTCGAGGGGATTGATGGATTGCAAGGATGAAGGTAAGAATGG | Subcellular localization |
| BCβAS1-R(SL) | TTCTCTTTTACCAATGATGGATGGAAGGAATCG | |
| BCβAS1-F(RP) | GGACGAGAAATCTGCACTGATGGTAAGGAGGATTGCAAGGATGAAGGTAAGAATGG | Recombinant proteins |
| BCβAS1-R(RP) | GTGCCGCGGCTGTTAGGATGTTGAAGGATG | |
| BCβAS1-F(FE) | CTGCTGAATGCGGCGGCTGTTAGGATGTTGAAGGATG | Function expression |
| BCβAS1-R(FE) | GACCGAAGAAGAGGACTATGGGATGTAAGGAGGATTGCAAGGATGAAGGTAAGAATGG | |
| 5’AOX1 | GACTGTCGTCAATGACAGAGAC | qRT–PCR |
| 3’AOX1 | GCAAAGGCGATTCGACATCC | |
| ADF5(F) | TGGGTTAGGCGACCTTG | |
| ADF5(R) | GACAACATCTACAAAGTGCCTCAT | |
| ADF5(SL) | CGAGACGCCCTGCAGTGCAG | |
| ADF5(RP) | TAGAGGTCTACACTGCTGAGT | |
| BCβAS1-F(RT) | ATGAAAATCGGTGAAGGCTCT | |
| BCβAS1-R(RT) | TTCTCCTTTACCAATGATGGATGGAAGGAATCG | |

The fragment was inserted into the pBI221-GFP vector to generate the pBI221-GFP-βAS fusion proteins. A pBI221-PM-mCherry vector containing an Arabidopsis H+-ATPase plasma membrane signal protein fragment fused with the red fluorescent protein mCherry ORF was used as a membrane localization control marker(Nelson et al. 2007). All fluorescence images, including GFP and mCherry signals, were visualized using a laser scanning confocal microscope (Leica TCS SP, Germany). The excitation and emission wave lengths were 488 nm and 510–550 nm for GFP and 587 nm and 610–650 nm for mCherry, respectively.

**qRT–PCR**

To compare the expression levels of BcBAS1 in different tissues of the four genotypes of B. chinense DC., quantitative reverse transcription PCR (qRT–PCR) was performed using the CFX96 Real-Time PCR System (Bio-Rad, USA) with Top Green qPCR SuperMix (TransScript, China). RNA was extracted and first-strand cDNA was synthesized as described previously. Two reference genes, ADF5 and ADF7, which we selected before, were used for normalization(Yu et al. 2019). The qRT–PCR primers are shown in Table 1. The qRT–PCR conditions were 94°C for 30 s, followed by 40 cycles of 94°C for 5 s and 60°C for 30 s. The expression level of each gene was calculated from the Cq values using the EΔΔCt method(Long et al. 2016). The fold change in transcript levels is presented as the mean ± standard error (SE) of transcript expression levels obtained from at least three independent experiments.

**Prokaryotic expression**

For the expression of BcBAS1 in E. coli, the ORF of BcBAS1 was amplified by homologous recombination and subcloned into the PET28a expression vector, and the recombinant vector PET28a-βAS was introduced into E. coli BL21(DE3) cultivated at 37°C in liquid LB medium with shaking at 220 rpm for expression. The empty PET28a vector was also transformed as a control. When the cells of the culture reached 0.4–0.6 at OD600 nm, 1 mM isopropyl β-D-thiogalactoside (IPTG) was added, and the cultures were incubated for 4 h at 37°C, 8 h at 25°C and 16 h at 16°C. The cells were sonicated for 8 min after culturing. Cell lysates and culture supernatants were collected by centrifugation at 12000 r/min for 10 min at 4°C. The protein distribution was detected by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS–PAGE).

**Eukaryotic expression**

The PCR product for functional expression was ligated with the pPIC9k vector by homologous recombination using a ClonExpress II One Step Cloning Kit (Vazyme, China). P. pastoris strain GS115 was transformed with the linearized pPIC9k-βAS vector by the restriction enzyme SacI using a P. pastoris positive clone assay kit (Coolaber, China), and the linearized empty pPIC9k vector was transformed into GS115 as a negative control. The above suspension culture was incubated on MD medium plates, and then single
colonies were selected on YPD plates with a high concentration of G418.

Genomic DNA was extracted from yeast cells using a yeast genomic DNA extraction kit (Tiangen, China). Positive clones were generated by PCR using the universal primers 5’AOX1 and 3’AOX1. The confirmed recombinant P. pastoris strain GS115 was grown in BMGY medium at 28°C until the culture reached OD600 = 2.0–6.0. After centrifugation, the resuspended cell pellet was grown in BMMY medium for 5 days, and 100% methanol was added to a final concentration of 0.5% every 24 h.

**GC–MS analyses**

To analyse the products of fermentation, transgenic yeast cells were harvested by centrifugation, mixed with 20% KOH (containing 50% ethanol), and lysed in boiling water for 15 min. After cooling at room temperature, the mixture was extracted with 10 ml of N-hexane 3 times, and the extract was concentrated to 1 ml using a rotary evaporator (Rongya RE-52A, China) for GC–MS analysis. An aliquot (1 µl) with split injection was taken for analysis using a Shimadzu GC-2010 Plus gas chromatography system equipped with an SH-Rxi-5Sil MS column (30 m × 0.25 mm × 0.25 µm). The injection temperature was 280°C, and the column temperature programme was as follows: 180°C for 1 minute, followed by an increase in temperature with a heating rate of 20°C per minute up to 280°C, with a final increase to 300°C at a rate of 2°C per minute and holding for 2 minutes (Yin et al. 2018).

**Statistical analysis**

All data were analysed with one-way analysis of variance (ANOVA) and Duncan’s multiple range tests (p < 0.05) by SPSS Statistics 20.0 software, and p < 0.05 was considered statistically significant.

**Results**

**Sequence and phylogenetic analyses**

To understand the evolutionary relationships of BcBASs with other plant OSCs, we compared their amino acid sequences to those of other reported OSCs with known functions (Table S1). The results of the phylogenetic analysis (Fig. 1) divided these genes into group I and group II based on their sequences. The BcBASs were all clustered in group I, which is the monofunctional β-AS gene group, and β-amyrin was the major compound. Group II comprised multifunctional enzymes producing α-amyrin and β-amyrin, together with other compounds, such as δ-amyrin, lupeol, taraxasterol, Ψ-taraxasterol, dammarenediol-II. Furthermore, BcBAS3 was clustered with the CrAS and OEA genes, which can produce α-amyrin and β-amyrin. BcBAS2 clustered most closely with PNY1. BcBAS4 was grouped with AaBAS, TkASC6 and AsOXA1, and these genes were characterized as monofunctional β-ASs that could produce β-amyrin only. BcBAS1 was different from other BcBASs; it clustered in group I but was not close to the others. These results suggested that BcBAS2 and BcBAS4 may be monofunctional β-ASs, BcBAS3 may be a multiproduct synthase, and BcBAS1 may be a monofunctional or multifunctional β-AS.

We designed specific primers according to the transcriptome sequence, and the 2498 bp full-length cDNA of BcBAS1 obtained had an ORF of 2289 bp that encoded a protein of 762 amino acids with a molecular weight of approximately 87 kDa, and the theoretical isoelectric point (pI) was 5.92. We found that BcBAS1 showed high similarity with β-ASs; it shared 98.60% sequence similarity with a β-AS of B. chinense DC. (HQ166837.1) and had 97.86%, 85.66%, and 84.00% similarities with β-ASs in Bupleurum kaoi (AY514455.1), Daucus carota var. sativus (XM_017372332.1) and Panax quinquefolium (JX262290.1), respectively. Moreover, alignment analyses (Fig. 2) of conserved regions of the BcBAS1 amino acid sequences among Bupleurum (Li et al. 2020), Arabidopsis (Kolesnikova et al. 2007), Ginseng (Kushiro et al. 1998) and Glycyrrhiza (Hayashi et al. 2001) showed that BcBAS1 exhibits the typical conserved regions of OSC genes, similar to others. BcBAS1 not only contains a DCTAE motif, which is considered a substrate-binding motif and is highly conserved in eukaryotic OSCs (Vishwakarma et al. 2013), but also has a MWCYCR motif, which is a characteristic motif of β-ASs. In addition, it possesses 4 QW motifs that may play a role in strengthening the structure of the enzyme and stabilizing its carbocation intermediates (Poralla et al. 1994).

**Subcellular localization of BcBAS1**

Online prediction of protein subcellular localization showed that BcBAS1 may be present in the cytosol. To confirm the predicted localization of BcBAS1, we constructed a pBII221-GFP-βAS vector and transiently expressed plasmids in Arabidopsis protoplasts, together with the H+-ATPase protein fused with mCherry. As shown in Fig. 3, there was significant green fluorescence around the edges of the protoplasts. The BcBAS1-GFP signal was similar to that of GFP alone and did not overlap with the red fluorescent protein mCherry signals, indicating that BcBAS1 was...
analysis showed that 8 h at 25°C and 16 h at 16°C resulted in successful expression of the fusion protein with the expected size of approximately 87 kDa, which was not shown in the negative control (Fig. 4). The recombinant proteins were expressed as inclusion bodies, and 16 h at 16°C was the best condition for fusion protein induction.

To further define the function of BcBAS1, we constructed a β-amyrin biosynthetic pathway in \textit{P. pastoris}. BcBAS1 was integrated into the genome of \textit{P. pastoris} GS115, and the empty pPIC9k plasmid was used as a control. The results showed that BcBAS1 was successfully expressed indeed a cytosol-localized gene, which was consistent with the prediction.

**Heterologous expression and functional characterization of BcBAS1**

To express the fusion protein of BcBAS1, the recombinant vector pET28a-βAS was constructed and confirmed by sequencing. The recombinant vector pET28a-βAS was transformed into \textit{E. coli} BL21(DE3) and induced with IPTG for 4 h at 37°C, 8 h at 25°C and 16 h at 16°C. SDS–PAGE analysis showed that 8 h at 25°C and 16 h at 16°C resulted in successful expression of the fusion protein with the expected size of approximately 87 kDa, which was not shown in the negative control (Fig. 4). The recombinant proteins were expressed as inclusion bodies, and 16 h at 16°C was the best condition for fusion protein induction.

To further define the function of BcBAS1, we constructed a β-amyrin biosynthetic pathway in \textit{P. pastoris}. BcBAS1 was integrated into the genome of \textit{P. pastoris} GS115, and the empty pPIC9k plasmid was used as a control. The results showed that BcBAS1 was successfully expressed
Expression analysis of BcBAS1

To gain insight into the expression patterns of the BcBAS1 gene, qRT‒PCR was conducted on four genotypes of B. chinense DC., including four tissues, the root, leaf, flower and stem. As shown in Fig. 6, BcBAS1 expression was detected in all tissues of the four genotypes, and its expression was significantly higher in the roots than in the other tissues of the four genotypes, with the lowest expression level in CBC1, CBC3 and RX stems, while the flower showed the lowest expression level in FS. The BcBAS1 transcript level

in P. pastoris GS115 (Fig. S1), and the correct colonies were cultivated and induced for 5 days. GC–MS analysis of cell extracts showed a clear new peak with a retention time of 17.535 min (Fig. 5b), which perfectly matched that of β-amyrin, and it was not detected in the control. There were three ion fragments of cell extracts and β-amyrin at m/z 189.0, 203.0 and 218.0 (Fig. 5a). The results indicated that BcBAS1 cDNA encoded a monofunctional β-AS and could catalyse the production of β-amyrin only, which validated the function of BcBAS1 predicted in the phylogenetic analysis.

Fig. 2 Comparison of deduced amino acid sequences of BcBAS1 with β-amyrin synthases from other plants, including Bupleurum, Arabidopsis, ginseng and Glycyrrhiza. Sequences were aligned using DNAMAN software. The highly conserved DCTAE is shown in the yellow box, QW motifs are boxed in green, and the MWCYCR motif is boxed in red.
SPSS Statistics 20.0. As shown in Table 2, BcBAS1 expression was positively correlated with SSa and SSd in the four genotypes of B. chinense DC. and showed an extremely significant positive correlation with SSa and SSd in the leaves of CBC3 (p < 0.01). BcBAS1 expression also showed a significant positive correlation with SSd in the leaves of FS (p < 0.05), and the other correlations were not statistically significant (p < 0.05). Based on the analysis, we speculated that the expression levels of the BcBAS1 gene influenced the contents of SSs. The BcBAS1 gene may play an important role in the synthesis of SSs, and an increase in the expression of the BcBAS1 gene may lead to elevated production of SSs and their precursors.

in the rootswas 121.7-fold higher than that in the stems of CBC3, while it was 92.3-, 53.6- and 12.6-fold higher in RX, FS and CBC1, respectively. The highest transcript level in roots indicated that roots might be the main location of saikosaponin synthesis in B. chinense DC. Moreover, the expression levels and patterns of BcBAS1 showed large differences among the four genotypes, which may result from differences in the genotypes themselves.

**Correlation between BcBAS1 gene expression and saikosaponin content**

To investigate the relationship between BcBAS1 gene expression levels and the contents of SSs (He et al. 2021), Pearson correlation analysis was performed using IBM SPSS Statistics 20.0. As shown in Table 2, BcBAS1 expression was positively correlated with SSa and SSd in the four genotypes of B. chinense DC. and showed an extremely significant positive correlation with SSa and SSd in the leaves of CBC3 (p < 0.01). BcBAS1 expression also showed a significant positive correlation with SSd in the leaves of FS (p < 0.05), and the other correlations were not statistically significant (p < 0.05). Based on the analysis, we speculated that the expression levels of the BcBAS1 gene influenced the contents of SSs. The BcBAS1 gene may play an important role in the synthesis of SSs, and an increase in the expression of the BcBAS1 gene may lead to elevated production of SSs and their precursors.
Triterpenoid saponins are one of the largest groups of natural products that function in increasing plant resistance (Sawai and Saito 2011) and are widely used in the food, pharmaceutical and cosmetics industries (Augustin et al. 2011; Osbourn et al. 2011). SSs are oleanane-type and ursane-type triterpenoid saponins, and more than 120 SSs have already been isolated from the genus *Bupleurum* (Ashour and Wink 2011; Lin et al. 2013). β-Amyrin synthetase is considered a key enzyme that catalyses the first step in the synthesis of triterpenoid saponins. After 2,3-oxidosqualene is synthesized through upstream processes of the mevalonate (MVA) and methylerythritol phosphate (MEP) pathways, β-amyrin synthetase is responsible for the cyclization of 2,3-oxidosqualene to the triterpene backbone β-amyrin, and then β-amyrin undergoes hydroxylation and oxidation by P450 and UGT to oleanane-type triterpenoid saponins in downstream processes. Here, we focused on the role of β-AS in saikosaponin synthesis. We selected four nonredundant β-AS genes (BcBAS1-4) from the transcriptome of *B. chinense* DC., and the four genes phylogenetically belonged to β-AS group I. We isolated BcBAS1, and the sequence alignments showed that BcBAS1 had typical conserved motifs for substrate binding and that the structure strengthened and stabilized the carbocation intermediate. BcBAS1 retained the necessary functional domains during evolution, consistent with

### Discussion

![Fig. 5](image-url) Identification of β-amyrin by GC–MS analysis of fermentation extracts. a: GC chromatograms of the β-amyrin standard and products extracted from yeast expressing pPIC9k-βAS. b: The blank and pink lines represent extracts of yeast expressing pPIC9k-βAS that showed a β-amyrin signal at a retention time of 17.535 min, and the blue line represents extracts of the control yeast strain expressing pPIC9k

![Fig. 6](image-url) The expression of BcBAS1 in different tissues of four genotypes of *B. chinense* DC. a: The expression of BcBAS1 in CBC1. b: The expression of BcBAS1 in CBC3. c: The expression of BcBAS1 in RX. d: The expression of BcBAS1 in FS. The different letters indicate significant differences (p < 0.05)

### Table 2

| Saikosaponin | CBC1 root | CBC1 leaf | CBC3 root | CBC3 leaf | FS root | FS leaf | RX root | RX leaf |
|--------------|-----------|-----------|-----------|-----------|---------|---------|---------|---------|
| SSa          | 0.881     | 0.929     | 0.861     | 1**       | 0.847   | 0.947   | 0.872   | 0.978   |
| SSd          | 0.922     | 0.984     | 0.939     | 1**       | 0.947   | 1*      | 0.890   | 0.876   |
other β-AS genes from near-source plant species(Hayashi et al. 2001; Kolesnikova et al. 2007; Kushiro et al. 1998).

To confirm the contribution of BcBAS1 to the biosynthesis of SSs, we examined its function using the P. pastoris expression system, which contains an inherent MVA pathway that can supply the precursor 2,3-oxidosqualene of triterpenoid saponin and is also tolerant to the use of heterologous proteins of other plants(Guo et al. 2020; Ro et al. 2006). We clearly detected β-amyrin among the metabolic products of the engineered yeasts, and β-amyrin was also the only product of fermentation, suggesting that BcBAS1 was a monofunctional β-AS. The function of BcBAS1 was similar to that of BCBAS(Li et al. 2020), another β-amyrin synthetase gene functionally validated in B. chinense DC.. The amino acid sequences from the coding regions of BcBAS1 and BCBAS were 73.28% similar, and both catalyse the cyclization of 2,3-oxidosqualene into β-amyrin only. The genes HMG-CoA reductase(HMGR), squalene synthase(SS), squalene epoxidase(SE), UGTand P450 were previously reported to play roles in saikosaponin biosynthesis(Sui et al. 2021). The triterpenoid pathways of ginsenoside(Dai et al. 2014; Liu et al. 2015; Wang et al. 2015), mogroside(Dai et al. 2015; Itkin et al. 2016) and hederagenin(Ma et al. 2018a)have been successfully reconstructed in yeast, and triterpenoids or their precursors were detected in fermentation products. Based on these studies, it could be speculated that saikosaponin production may be achieved by assembling biosynthesis genes encoding a heterologous saikosaponin pathway in yeast.

Subcellular localization is critical to the function of proteins, and protein localization and function are interdependent. The subcellular localization of BcBAS1 indicated that it is present in the cytosol, and β-amyrin was detected in the cell lysates. The MVA pathway, which is the major source of the precursors for triterpenoid saponins, is also carried out in the cytosol(Rodriguez-Concepcion and Boronat 2002). Thus, it could be speculated that BcBAS1 directly catalysed 2,3-oxidosqualene to β-amyrin in the cytosol of B. chinense DC.. In Osmanthus, the OiBETA-AS responsible for the synthesis of triterpenoids was localized in the nucleus and membrane(Yang et al. 2018). Another amyrin synthetase gene,AgAS, in Ampelopsis grossedentatawas localized primarily in the nucleus and membrane, with a minor amount in the cytosol(Zhang et al. 2017). The subcellular localization of these three proteins is not similar, and the reasons may be related to the differences in sequences, cellular environments, specific functions, or variants between species. We also succeeded in expressing the proteins of BcBAS1, which had a molecular weight of approximately 87 kDa, similar to that of other β-ASs, using an E. coli expression system(Liu et al. 2009; Wu et al. 2012).

The expression levels of BcBAS1 and the contents of SSs in different tissues among the four genotypes of B. chinense DC. were correlated in this study for the first time. BcBAS1 was detected in all tissues, and the highest expression was observed in roots. The expression differences may be attributed to the tissue specificity of BcBAS1. The phenomenon that the β-AS genes were specifically expressed in some specific tissues was also reported in other plants(Dhar et al. 2014; Zheng et al. 2015), which might also be one of the reasons for the inequitable distribution of SSs or other triterpenoid saponins. The relationship between the expression levels of BcBAS1 and the content of SSs was also investigated in this study. Previous studies in other plants showed that the transcription levels of genes in the saponin synthesis pathway were correlated with the content of saponins. The expression levels of squalene synthase(SSQ), squalene monoxygenase(SQE) and β-AS genes were highly correlated with the accumulation of triterpenoid saponins in Radix Polygalae(Zhang et al. 2014). P450 and UGT transcripts correlated with the contents of saponins in the roots of Panax plants(Wei et al. 2020). In our studies, the strong positive correlation between the expression of BcBAS1 and the contents of SSs in the roots and leaves of B. chinense DC. showed that BcBAS1 was indeed involved in the synthesis of SSs. On the other hand, some studies have shown that the correlation between β-AS and SSs might also be changed when genes in other pathways change or plants are under stress(Yang et al. 2019), indicating that the synthesis and regulation of SSs are not the results of single genes but dependent on the combination of all genes in the pathway. Environmental factors also affect the contents. Therefore, in future studies dealing with the production of SSs, the relationship among the pathway genes in metabolic engineering, the effect of the growth environment during the cultivation process, and optimal conditions to increase the accumulation of SSs should be taken into account.

Conclusions

In this study, a SSs pathway gene named BcBAS1 from B. chinense DC. was successfully cloned and characterized using heterologous expression systems in E. coli and P. pastoris. Our results showed that the recombinant protein size of BcBAS1 was approximately 87 kDa. BcBAS1 is a β-AS gene, and β-amyrin is the sole fermentation product of P. pastoris expressing BcBAS1. BcBAS1 was localized in the cytoplasm, which was consistent with our predictions. Moreover, the expression levels of BcBAS1 were strongly associated with the contents of SSs. The highest BcBAS1 expression levels were found in the roots of the four genotypes, and BcBAS1 may be involved in the biosynthesis of triterpenoid saponins.
and regulation of SSs. However, the function of BcBAS1 needs further exploration, and the coregulation of P450s and UGTs in the biosynthesis of SSs also needs to be investigated in the future. The results of this work further reveal the biosynthesis pathway of SSs and might be beneficial for further investigation of the biosynthesis and regulation of SSs, providing a foundation for the metabolic engineering of SSs in the future.

Abbreviations

SSs  Saikosaponins
β-AS  β-Amyrin synthase
OSCs  Oxidosqualenecyclases
P450s  Cytochrome P450 enzymes
UGTs  Uridine diphosphate glycosyltransferases
GFP  Green fluorescent protein
qRT–PCR  Quantitative reverse transcription PCR
SDS–PAGE  Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
MVA  Mevalonate

Authors’ contributions

Conceptualization, Dabin Hou; methodology, Dabin Hou, Hua Chen and Yiguang Zhang; software, Yanping Mao; Ping Wei and Yuping Yang; investigation, Yanping Mao, Jun Zhao, Yuchan Li and Liang Feng; resources, Dabin Hou; writing—original draft preparation, Yanping Mao; writing—review and editing, Dabin Hou and Yanping Mao; visualization, Yanping Mao; funding acquisition, Dabin Hou. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the National Transgenic Major Project of China (2019ZX08010004-005), the Crop and Livestock Breeding Project in Sichuan (2021YFYZ0012-13), and the Major Horizontal Project (20zh0234).

Data availability

All the data were presented in the main manuscript and additional supporting files.

Declarations

Conflict of interest

The authors declare that they have no conflicts of interest.

References

Ashour ML, Wink M (2011) Genus Bupleurum: a review of its phytochemistry, pharmacology and modes of action. J Pharm Pharmacol 63:305–321. https://doi.org/10.1111/j.2042-7158.2010.01170.x
Augustin JM, Kuzina V, Andersen SB, Bak S (2011) Molecular activities, biosynthesis and evolution of triterpenoid saponins. Phytochemistry 72:435–457. https://doi.org/10.1016/j.phytochem.2011.01.015
Basuymi M, Oku H, Tsujimoto E, Kinjo K, Baba S, Takara K (2007) Triterpene synthases from the Okinawan mangrove tribe, Rhizophoraceae. FEBS J 274:5028–5042. https://doi.org/10.1111/j.1742-4658.2007.06025.x
Cao P, Wu S, Wu T, Deng Y, Zhang Q, Wang K, Zhang Y (2020) The important role of polysaccharides from a traditional Chinese medicine-Lung Cleansing and Detoxifying Decoction against the COVID-19 pandemic. Carbohydr Polym 240:116346. https://doi.org/10.1016/j.carbpol.2020.116346
Chinesse Pharmacopoeia Commission (2020) Pharmacopoeia of the people’s republic of China. China Medical Sciences and Technology Press, Beijing
Dai L, Liu C, Zhu Y, Zhang J, Men Y, Zeng Y, Sun Y (2015) Functional characterization of cucurbitadienol synthase and triterpene glycosyltransferase involved in biosynthesis of mogrosides from Siraitia grosvenorii. Plant Cell Physiol 56:1172–1182. https://doi.org/10.1093/pcp/pcv043
Dai Z, Wang B, Liu Y, Shi M, Wang D, Zhang X, Liu T, Huang L, Zhang X (2014) Producing aglycons of ginsenosides in bakers’ yeast. Sci Rep 4:3698. https://doi.org/10.1038/srep03698
Dhar N, Rana S, Razdan S, Bhat WW, Hussain A, Dhar RS, Vaishnavi S, Hamid A, Vishwakarma R, Tattoo SK (2014) Cloning and functional characterization of three branch point oxidosqualene cyclases from Withania somnifera (L.) dunal. J Biol Chem 289:17249–17267. https://doi.org/10.1074/jbc.M114.571919
Gao K, Wu SR, Wang L, Xu YH, Wei JH, Siu C (2015) Cloning and analysis of β-amyrin synthase gene in Bupleurum chinense. Genes Genom 37:767–774. https://doi.org/10.1007/s12358-015-0307-0
Guo H, Wang H, Huo YY (2020) Engineering critical enzymes and pathways for improved triterpenoid biosynthesis in yeast. ACS Synth Biol 9:2214–2227. https://doi.org/10.1021/acssynthbio.0c00124
Hayashi H, Huang P, Kirakosyan A, Inoue K, Hiraoka N, I chesshiro Y, Kushiro T, Shibusya M, Ebizuka Y (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. https://doi.org/10.1248/bpb.24.912
He Y, Chen H, Zhao J, Yang Y, Yang B, Feng L, Zhang Y, Wei P, Hou D, Zhao J, Yu M (2021) Transcriptome and metabolome analysis to reveal major genes of saikosaponin biosynthesis in Bupleurum chinense. BMC Genomics 22:839. https://doi.org/10.1186/s12864-021-08144-6
Hirai MY, Yano M, Goodenowe DB, Kanaya S, Kimura T, Awazuhara M, Arita M, Fujitani Y, Saito K (2004) Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana. Proc Natl Acad Sci U S A 101:10205–10210. https://doi.org/10.1073/pnas.0403218101
Huang HQ, Zhang X, Xu ZQ, Xu Y, Yan SK, Zhang WD (2009) Fast determination of saikosaponins in Bupleurum by rapid resolution liquid chromatography with evaporative light scattering detection. J Pharm Biomed Anal 49:1048–1055. https://doi.org/10.1016/j.jpba.2009.01.011
Itilin M, Davidovich-Rikanati R, Cohen S et al (2016) The biosynthetic pathway of the nonsugar, high-intensity sweetener mogroside V from Siraitia grosvenorii. Proc Natl Acad Sci U S A 113:E7619–E7628. https://doi.org/10.1073/pnas.1604828113
Jiang H, Yang L, Hou A, Zhang J, Wang S, Man W, Zheng S, Yu H, Wang X, Yang B, Wang Q, Kuang H (2020) Botany, traditional uses, phytochemistry, analytical methods, processing, pharmacology and pharmacokinetics of Bupleuri Radix: a systematic review. Biomed Pharmacother 124:106293. https://doi.org/10.1016/j.biopha.2020.110679
Kolesnikova MD, Wilson WK, Lynch DA, Obermeyer AC, Matuda S (2007) Arabidopsis camellia C synthase evolved from enzymes that make pentacycles. Org Lett 9:5223–5226. https://doi.org/10.1021/ol702399g
Kushiro T, Shibusya M, 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 256:238–244. https://doi.org/10.1046/j.1342-1327.1998.2560238.x
Li HY, Zhao YH, Zeng MJ, Fang F, Li M, Qin TT, Ye LY, Li HW, Qu R, Ma SP (2017) Saikosaponin D relieves unpredictable chronic mild stress induced depressive-like behavior in rats: involvement of HPA axis and hippocampal neurogenesis. Psychopharmacology 234:3385–3394. https://doi.org/10.1007/s00213-017-4720-8

Li JC, Wang C, Qi WT, Liu CL (2020) Cloning and functional characterization of the β-amyrin synthase gene from Bupleurum chinense. Biol Plant 64:314–319. https://doi.org/10.32615/bp.2020.008

Lin TY, Chiou CY, Chiou SJ (2013) Putative genes involved in saikosaponin biosynthesis in Bupleurum species. Int J Mol Sci 14:12806–12826. https://doi.org/10.3390/ijms140612806

Liu XB, Liu M, Tao XY, Zhang ZX, Wang FQ, Wei DZ (2015) Metabolic engineering of *Pichia pastoris* for the production of dammarenediol-II. J Biotechnol 216:47–55. https://doi.org/10.1016/j.jbiotec.2015.10.005

Liu Y, Cai Y, Zhao Z, Wang J, Li J, Xin W, Xia G, Xiang F (2009) Cloning and Functional Analysis of a beta-amyrin synthase gene associated with oleane acid biosynthesis in Gentiana straminea MAXIM. Biol Pharm Bull 32:818–824. https://doi.org/10.1248/bpb.32.818

Long X, He B, Fang Y, Tang C (2016) Identification and characterization of the glucose-6-phosphate dehydrogenase gene family in the para rubber tree, *Hevea brasiliensis*. Front Plant Sci 7:215. https://doi.org/10.3389/fpls.2016.00215

Ma XL, Li WX, Wang D, Lu FP, Dai ZB, Zhang XL (2018a) Biosynthesis analysis of hederagenin pathway and its construction in yeast cells. Zhongguo Zhong Yao Za Zhi 43:1844–1850. https://doi.org/10.1016/j.nh.2017.10.015

Ma Y, Bao Y, Wang S, Li T, Chang X, Yang G, Meng X (2016) Anti-inflammation effects and potential mechanism of saikosaponins by regulating nicotinate and nicotinamide metabolism and arachidonic acid metabolism. Inflammation 39:1453–1461. https://doi.org/10.1007/s00213-016-0377-4

Ma YS, Yang R, Zhou S, Yin YC, Zhang XD, Liu Y (2018b) B-amyrin synthase, one of the most important key enzymes for triterpene skeleton formation in higher plants. Pak J Bot 50:231–243

Moses T, Papadopoulou KK, Osbourn A (2014) Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. Crit Rev Biochem Mol Biol 49:439–462. https://doi.org/10.3109/10409238.2014.953628

Nelson BK, Cai X, Nebenfuhr A (2007) A multicolored set of organelle markers for co-localization studies in Arabidopsis and other plants. Plant J 51:1126–1136. https://doi.org/10.1111/j.1365-313X.2007.03212.x

Osbourn A, Goss RJ, Field RA (2011) The saponins: polar isoprenoids with important and diverse biological activities. Nat Prod Rep 28:1261–1268. https://doi.org/10.1039/c1np00011b

Poralla K, Hewelt A, Prestwich GD, Abe I, Reipen I, Sprenger G (1994) A specific amino acid repeat in squalene and oxidosqualene cyclases. Trends Biochem Sci 19:157–158. https://doi.org/10.1016/0968-0004(94)90276-3

Ro DK, Paradise EM, Onelet M et al (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440:940–943. https://doi.org/10.1038/nature04640

Rodriguez-Concepcion M, Boronat A (2002) Elucidation of the methylenedioxy phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiol 130:1079–1089. https://doi.org/10.1104/pp.007138

Sawai S, Saito K (2011) Triterpenoid biosynthesis and engineering in plants. Front Plant Sci 2:25. https://doi.org/10.3389/fpls.2011.00025

Sui C, Han WJ, Zhu CR, Wei JH (2021) Recent progress in saikosaponin biosynthesis in bupleurum. Curr Pharm Biotechnol 22:329–340. https://doi.org/10.2174/138920121999200918101248

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197

Vishwakarma RK, Sonawane P, Singh S, Kumari U, Ruby, Khan BM (2013) Molecular characterization and differential expression studies of an oxidosugale cyclase (OSC) gene of Brahami (*Bacopa monniera*). Physiol Mol Biol Plants 19:547–553. https://doi.org/10.1016/j.smp.2012.01.015-9

Wang P, Wei Y, Fan Y, Liu Q, Wei W, Yang C, Zhang L, Zhao G, Yue J, Yan X, Zhou Z (2015) Production of bioactive ginsenosides Rh2 and Rg3 by metabolically engineered yeasts. Metab Eng 29:97–105. https://doi.org/10.1016/j.meb.2015.03.003

Wei G, Yang F, Wei F et al (2020) Metabolomes and transcriptomes revelead the saponin distribution in root tissues of *Panax quinquefolius* and *Panax notoginseng*. J Ginseng Res 44:757–769. https://doi.org/10.1007/j.101.19.2015.05.009

Wu Y, Zou HD, Cheng H, Zhao CY, Sun LF, Su SZ, Li SP, Yuan YP (2012) Cloning and characterization of a beta-amyrin synthase gene from the medicinal tree *Aralia elata* (Araliaceae). Genet Mol Res 11:2301–2314. https://doi.org/10.4238/2012.August.13.4

Yang L, Zhao Y, Zhang Q, Cheng L, Han M, Ren Y, Yang L (2019) Effects of drought-re-watering-drought on the photosynthesis physiology and secondary metabolite production of *Bupleurum chinense* DC. Plant Cell Rep 38:1181–1197. https://doi.org/10.1007/s00209-019-02436-8

Yang X, Ding W, Yue Y, Xu C, Wang X, Wang L (2018) Cloning and expression analysis of three critical triterpenoid pathway genes in *Osmanthus fragrans*. Electron J Biotechnol 36:1–8. https://doi.org/10.1016/j.ejbt.2018.08.007

Yin Y, Gao L, Zhang X, Gao W (2018) A cytochrome P450 monooxygenase responsible for the C-22 hydroxylation step in the Paris polyphylla steroidal saponin biosynthesis pathway. Phytochemistry 156:116–123. https://doi.org/10.1016/j.phytochem.2018.09.005

Yu M, Liu D, Li YC, Sui C, Chen GD, Tang ZK, Yang CM, Hou DB, Wei JH (2019) Validation of reference genes for expression analysis in three *Bupleurum* species. Biotechnol Biotechnol Equip 33:154–161. https://doi.org/10.1007/s10522-018-1557556

Zhang F, Li X, Li Z, Xu X, Peng B, Qin X, Du G (2014) UPLC-Q/TOF MS-based metabolomics and qRT-PCR in enzyme gene screening with key role in triterpenoid saponin biosynthesis of *Polyspora temufoila*. PLoS ONE 9:e105765. https://doi.org/10.1371/journal.pone.0105765

Zhang YW, Yi HJ, Zhao S, Zheng JG, Xu M (2017) Cloning and expression analysis of amyrin synthase gene coding sequence in *Ampelopsis grossedentata*. Acta Bot Boreali Occident Sin 37:428–435

Zhao W, Li JJ, Yue SQ, Zhang LY, Dou KF (2012) Antioxidant activity and hepatoprotective effect of a polysaccharide from Bei Chaihu (*Bupleurum chinense* DC). Carbohydr Polym 89:448–452. https://doi.org/10.1016/j.carbpol.2012.03.027

Zhao X, Liu J, Ge S, Chen C, Li S, Wu X, Feng X, Wang Y, Cai D (2019) Saikosaponin A inhibits breast cancer by regulating Th1/Th2 balance. Front Pharmacol 10:624. https://doi.org/10.3389/fphar.2019.00624

Zheng X, Luo X, Ye G, Chen Y, Ji X, Wen L, Xu Y, Xu H, Zhan R, Chen W (2015) Characterisation of two oxidosugale cyclases responsible for triterpenoid biosynthesis in *Ilex asprella*. Int J Mol Sci 16:3564–3578. https://doi.org/10.3390/ijms16023564

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s);
author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.