Electronic Journal of Biotechnology 36 (2018) 1–8

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

Research article

Cloning and expression analysis of three critical triterpenoid pathway genes in *Osmanthus fragrans*

Xiulian Yang, Wenjie Ding, Yuanzheng Yue, Chen Xu, Xi Wang, Lianggui Wang *

College of Landscape Architecture of Nanjing Forestry University, No. 159 Longpan Road, Nanjing 210037, Jiangsu, China
Key Laboratory of Landscape Architecture of Jiangsu Province, No. 159 Longpan Road, Nanjing 210037, Jiangsu, China

A R T I C L E I N F O

Article history:
Received 15 November 2017
Accepted 24 August 2018
Available online 31 August 2018

Keywords:
Aroma
Essential oils
Flower gene expression
Ornamental tree
Osmanthus fragrans
Terpenoid metabolic pathway
Squalene epoxidaes
Squalene synthase
Terpenes
Triterpenoid biosynthesis

A B S T R A C T

Background: Osmanthus fragrans is an important ornamental tree and has been widely planted in China because of its pleasant aroma, which is mainly due to terpenes. The monoterpenoid and sesquiterpenoid metabolic pathways of sweet osmanthus have been well studied. However, these studies were mainly focused on volatile small molecule compounds. The molecular regulation mechanism of synthesis of large molecule compounds (triterpenoids) remains unclear. Squalene synthase (SQS), squalene epoxidase (SQE), and beta-amyrin synthase (BETA-AS) are three critical enzymes of the triterpenoid biosynthesis pathway.

Results: In this study, the full-length cDNA and gDNA sequences of *OfSQS*, *OfSQE*, and *OfBETA-AS* were isolated from sweet osmanthus. Phylogenetic analysis suggested that *OfSQS* and *OfSQE* had the closest relationship with *Sesamum indicum*, and *OfBETA-AS* sequence shared the highest similarity of 99% with that of *Olea europaea*. The qRT-PCR analysis revealed that the three genes were highly expressed in flowers, especially *OfSQE* and *OfBETA-AS*, which were predominantly expressed in the flowers of both "Boye" and "Rixiang" cultivars, suggesting that they might play important roles in the accumulation of triterpenoids in flowers of *O. fragrans*. Furthermore, the expression of *OfBETA-AS* in the two cultivars was significantly different during all the five flowering stages; this suggested that *OfBETA-AS* may be the critical gene for the differences in the accumulation of triterpenoids.

Conclusion: The evidence indicates that *OfBETA-AS* could be the key gene in the triterpenoid synthesis pathway, and it could also be used as a critical gene resource in the synthesis of essential oils by using bioengineered bacteria.

How to cite: Yang X, Ding W, Yue Y, et al. Cloning and expression analysis of three critical triterpenoids pathway genes in *Osmanthus fragrans*. Electron J Biotechnol 2018;36. https://doi.org/10.1016/j.ejbt.2018.08.007.

© 2018 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. All rights reserved. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Osmanthus fragrans, commonly known as sweet osmanthus, is an evergreen shrub or small tree belonging to Oleaceae [1]. It is distributed in China and other Asian countries, such as South Korea, India, and Thailand [2,3,4]. Presently, it has been widely cultivated as an urban ornamental tree in China [2]. Its fresh flowers have a pleasant fruity and sweet aroma. The extracts from *O. fragrans* flowers are also important sources of fragrance in the perfume and cosmetic industries [5,6]; the extracts can also be used to reduce inflammation, resist oxidation, and prevent aging [7].

As a group of secondary metabolites, triterpenoids are synthesized by the mevalonate pathway (MVA pathway) and 2-Cmethyl-D-erythritol-4-phosphate pathway (MEP pathway) [8]. Although these pathways have been studied in some plants, the downstream genes of the MEP pathway remain unclear, especially those related to the synthesis of triterpenoids. Squalene synthase (SQS, EC 2.5.1.21), which catalyzes the first enzymatic step that shifts carbon pool away from the central isoprenoid pathway toward the biosynthesis of terpenes, could determine the yield of subsequent products such as triterpenes, sterols, and cholesterol [9]. It has been reported that *SQS* genes play an important role in regulating the biosynthetic pathway of triterpenes, and the overexpression of *SQS* genes in *Panax ginseng* leads to enhanced accumulation of triterpenes [10]. This regulation increases the mRNA accumulation of downstream genes such as squalene epoxidase (SQE, EC 1.14.99.7) and increases the production of phytosterols and triterpenes [11]. Squalene epoxidase, a downstream gene of *SQS*, catalyzes the formation of 2,3-oxidosqualene [12]. The precursor 2,3-oxidosqualene is catalyzed by beta-amyrin synthase (BETA-AS, EC 5.4.99.39) and cycloartenol synthase (*CAS*) to produce beta-amyrin and cycloartenol, respectively, which are further modified to form triterpenoid and phytosterol, respectively (Fig. 1) [13]. *BETA-AS*

* Corresponding author.
E-mail address: wlg@njfu.com.cn (L. Wang).
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

https://doi.org/10.1016/j.ejbt.2018.08.007
0717-3458/© 2018 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. All rights reserved. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

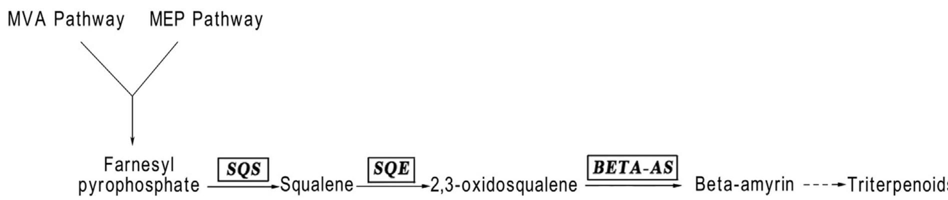

Fig. 1. Simplified representation of the triterpenoid biosynthetic pathway. MVA pathway: Mevalonate pathway, MEP pathway: 2-Cmethyl-D-erythritol-4-phosphate pathway, *SQS*: Squalene synthase gene, *SQE*: Squalene epoxidase gene, *BETA-AS*: β-Amyrin synthase gene. The three genes in the box are the genes investigated in this study.

was defined as an important branch point along the metabolic pathway of triterpenoids, and it plays a regulatory role in the biosynthesis of triterpenoids [14].

Triterpenoids isolated from *O. fragrans* showed hypolipidemic and antioxidative activity [15]. However, the process of biosynthesis of these compounds is unclear, and the molecular mechanism of synthesis of triterpenoids in sweet osmanthus remains to be clarified. In the present study, the full-length cDNA and gDNA sequences of three genes, namely *OfSQS*, *OfSQE*, and *OfBETA-AS*, were successfully isolated and further identified. Then, the expression patterns of the three genes in the different flower development stages of two *Osmanthus* cultivars were analyzed by qRT-PCR.

2. Materials and methods

2.1. Plant materials

Two cultivars of *O. fragrans*, “Boye Jingui” (with strong aroma) and “Rixiang Gui” (with light aroma) have been grown under natural conditions since 2005 in the campus of Nanjing Forestry University, Jiangsu Province, China (112.32′E, 156.32′W). The flowers of “Rixiang Gui” at the full blooming stage (S4) collected in 2015 were used for gene cloning. Different organs, including root, stem, leaf, and flower at the full blooming stage were collected in October 2016 for tissue-specific expression study. From September to October 2016, flowers were harvested during five flowering stages including bud-pedicel stage (S1), bud-eye stage (S2), primary blooming stage (S3), full blooming stage (S4), and flower fading stage (S5), respectively. Samples were immediately frozen in liquid nitrogen and stored at -80°C for subsequent use.

2.2. RNA extraction and first-strand cDNA synthesis

Total RNA was extracted using the RNAprep pure Kit (Tiangen, China) following the manufacturer's instructions. The quality of the extracted RNA was determined by NanoDrop 2000 Spectrophotometer (Thermo Scientific, USA), and RNA integrity was evaluated by agarose gel electrophoresis. Then, the first strand cDNA was synthesized with the Revert Aid First Strand cDNA Synthesis Kit (Thermo Scientific) according to the manufacturer's protocol.

2.3. Cloning of full-length cDNAs

According to the EST (expressed sequence tags) sequences of the *O. fragrans* transcriptomic databases, specific primers (Table S1) were designed to obtain the core sequences of *OfSQS*, *OfSQE*, and *OfBETA-AS*. Next, by using the 3′-Full RACE Core Set with PrimeScript™ RTase (Takara Biotechnology) and SMARTer™ RACE cDNA Amplification Kit (Clontech, USA), full-length sequences of these three genes were obtained. The specific primers for 3′ RACE and 5′ RACE (Table S2) were designed using Oligo 6.0 software. Finally, the PCR products of these three genes were purified and cloned into pEASY®-T1 vector (Transgen Biotech, China) to confirm by sequencing.

2.4. Isolation of genomic DNA sequences of OfSQS, OfSQE, and OfBETA-AS genes

The genomic DNA was extracted using the Plant Genomic DNA Kit (Tiangen, China) following the manufacturer's instructions. The genomic DNA extracted from “Rixiang Gui” petals was used as a template to obtain the genomic DNA sequences of *OfSQS*, *OfSQE*, and *OfBETA-AS* genes with the specific primers (Table S3). After PCR amplifications, the purified products were directly sequenced by Genescrip Inc. (Genescrip, China).

2.5. Bioinformatics analysis

Open reading frame and protein prediction were made by NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Nucleotide sequences were identified using the NCBI Blast program (http://www.ncbi.nlm.nih.gov/BLAST). Physical and chemical parameters of proteins were determined by the ProtParam tool (http://web.expasy.org/protparam). The conservative domain was predicted by the PFAM tool (http://smart.embl-heidelberg.de/). The signal peptide was predicted by SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/). Transmembrane topology prediction was made by TMHMM Server version 2.0 Server (http://www.cbs.dtu.dk/services/TMHMM/). The structure of genomic organization was established by Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/). The homology analysis was conducted using BLAST of GenBank (http://www.ncbi.nlm.nih.gov/BLAST/). Finally, the phylogenetic trees were constructed by ClustalX 2.1 and MEGA 5.0 with 1000 bootstraps.

2.6. Gene expression analysis

Quantitative real-time RT-PCR (qRT-PCR) was performed using ABI StepOnePlus Systems (Applied Biosystems, USA) with SYBR Premix Ex Taq (Takara Biotechnology). The RNA samples were quantified by NanoDrop 2000 Spectrophotometer. cDNA was synthesized from 5 μg total RNA and diluted 10-fold for the gene expression experiment. *OfRAN* and *OfRPB2* were considered as the reference genes for different organs and different flowering stages, respectively [16]. The specific primers used in the experiment to detect the gene expression levels were designed by Primer Premier 5.0 software (Table S4). The thermal cycle conditions used were as follows: 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, 60°C for 30 s, 95°C for 15 s, 60°C for 1 min, and 95°C for 15 s [17]. The relative expression levels were calculated by the $2^{-\triangle\triangle CT}$ method. Data were presented as mean values with error bars indicating standard error. Different letters indicate significant differences at the 0.01 level according to Tukey's test.

2.7. Subcellular localization assay

The coding region (without stop codon) of *OfBETA-AS* was amplified from cDNA template of “Rixiang Gui” petals with gene-specific primers (Table S5). Vector (*35s::GFP-OfBETA-AS*) constructed by subcloning the coding region of *OfBETA-AS* into the Super1300::GFP vector (provided by Dr. Yuanzheng Yue) was used for the subcellular localization assay.

The *35s::GFP-OfBETA-AS* and Super1300::GFP control vector were electroporated into *Agrobacterium tumefaciens* strain EHA105 by using Eppendorf Eporator®4309 (Eppendorf, Germany). Then, the tobacco (*Nicotiana benthamiana*) leaves were infiltrated by the vector-containing *Agrobacterium*. After 2 d of incubation at 21°C and 14 h photoperiod, the infiltrated plants were used to observe the GFP fluorescence signal on a laser canning confocal microscope Zeiss LSM 710 (Zeiss, Germany).

3. Results

3.1. Cloning and sequence analysis of cDNAs

To achieve the full-length cDNAs of *OfSQS*, *OfSQE*, and *OfBETA-AS*, the RACE technology was used to obtain the 3′ region and 5′ region of these genes. The full-length sequence of *SQS* cDNA was 1672 bp and contained an open reading frame (ORF) of 1245 bp encoding a protein with 414 amino acids, a 5′-untranslated region (5'-UTR) of 90 bp, and a 3'-UTR of 337 bp. The cDNA of *SQE* was 1841 bp with an ORF of 957 bp encoding a protein with 318 amino acids, a 5'-UTR of 494 bp, and a 3'-UTR of 390 bp. The cDNA of *BETA-AS* was 2737 bp with an ORF of 2289 bp encoding a protein with 762 amino acids, a 5'-UTR of 185 bp, and a 3'-UTR of 263 bp. The sequences of *OfSQS*, *OfSQE*, and *OfBETA-AS* have been submitted to GenBank with the accession numbers of KY992860, KY992861, and KY992862, respectively.

The theoretical isoelectric points (pI) of OfSQS, OfSQE, and OfBETA-AS were 7.57, 9.01, and 5.97, respectively, and the putative molecular weights of these genes were 47.67, 34.51, and 87.21 kDa, respectively. The instability index showed that OfSQS, OfSQE, and OfBETA-AS belonged to the group of unstable proteins. In amino acid sequences of the three genes, no signal peptide sequence was identified. Transmembrane region analysis of the amino acid sequence indicated that the presence of strong transmembrane helixes from 281 to 303 aa and 387 to 409 aa in OfSQS, 251 to 270 aa and 277 to 299 aa in OfSQE, and 609 to 631 aa in OfBETA-AS. The conserved domains of OfSQS and OfSQE proteins were from 44 to 316 aa and 4 to 277 aa, respectively, and for the OfBETA-AS protein, the conserved domains were from 100 to 406 and 415 to 754 aa. These conserved domains evidenced that these genes could be categorized as a member of the SQS family, SQE family, and SQHop cyclase C and SQHop cyclase N superfamily (Fig. 2a), respectively. Alignment result revealed that 6 conservative domains, namely I (58–76), II (77–90), III (166–188), IV (206–224), V (310–324), and VI (391–417), were found in OfSQS. The putative flavin adenine dinucleotide (FAD) binding domain was found in the amino acid sequence of OfSQE. Motif analysis revealed that OfBETA-AS contained the four QW motifs (111–117, 161–167, 603–609, and 652–658), the DCTAE motif (485–489), and the MWCYCR (256–262) motif among the highly conserved regions.

3.2. Cloning and sequence analysis of genomic DNA

To analyze the exons/introns of the three genes, genomic sequences of *OfSQS*, *OfSQE*, and *OfBETA-AS* were obtained by PCR amplification with genomic DNA. Sequence analysis indicated that the sizes of *OfSQS*, *OfSQE*, and *OfBETA-AS* genomic DNA were 7496 bp, 1950 bp, and 5606 bp, respectively. By homologous alignment of the cDNA and genomic DNA sequences, the genomic organizations of the three genes were elucidated and positions of introns/exons were determined. Our results showed that *OfSQS, OfSQE, and OfBETA-AS* consisted of 12, 4, and 17 introns, respectively (Fig. 2b). All these identified introns began

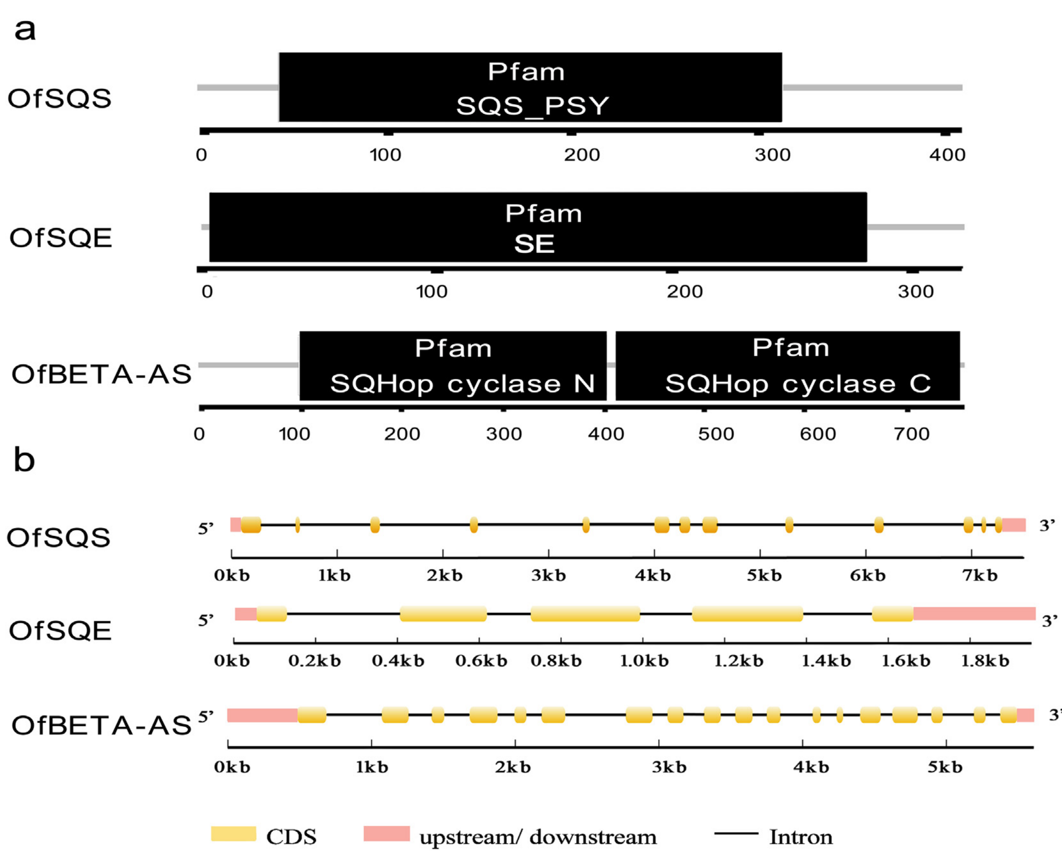

Fig. 2. Structure of OfSQS, OfSQE, and OfBETA-AS. (a) Conserved domains of OfSQS, OfSQE, and OfBETA-AS. The conserved domain analysis of the three proteins was performed using the Pfam Conserved Domain Database. The result indicated that OfSQS, OfSQE, and OfBETA-AS were categorized as a member of SQS family, SQE family, and SQHop cyclase C and SQHop cyclase N superfamily, respectively. (b) Genomic organization of the OfSQS, OfSQE, and OfBETA-AS genes. Exons are represented as yellow lines, introns as fine lines, and UTRs as blue lines. The structure of genomic organization was determined using Gene Structure Display Server.

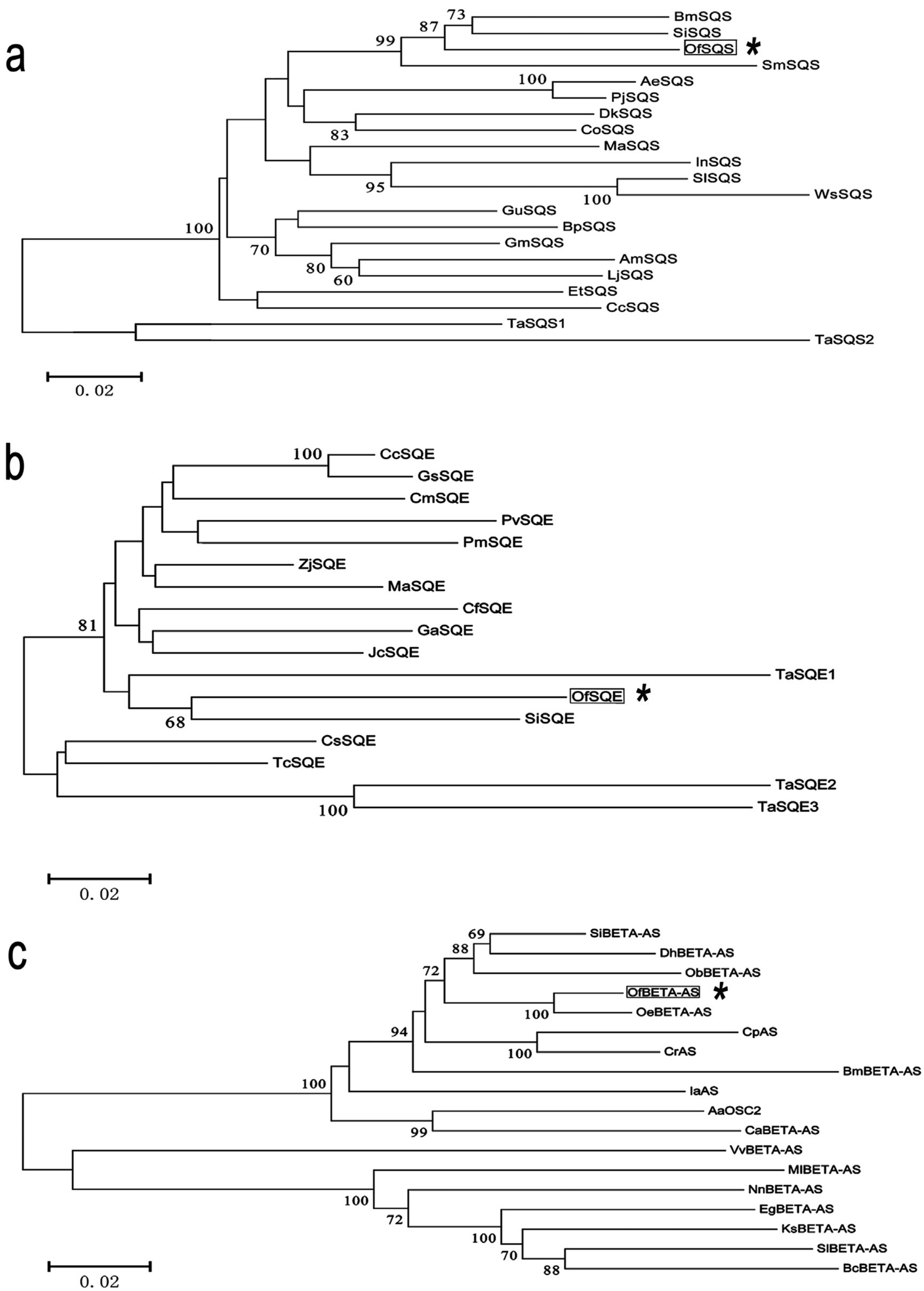

Fig. 3. Phylogenetic trees of *OfSQS*, *OfSQE*, and *OfBETA-AS* from *O. fragrans* and other plants using the deduced amino acids. Phylogenetic trees were constructed by ClustalX 2.1 and MEGA 5.0 with the 1000 bootstrap method. (a) Phylogenetic tree of *OfSQS* with other plants such as *S. indicum* (XP_011092839.1), *B. monnieri* (ADX01171.1), *C. oleifera* (AGB05603.1), *S. miltiorrhiza* (ACR57219.1), *Diospyros kaki* (ACN69082.1), *E. tirucalli* (BAH23428.1), *Solanum lycopersicum* (NP_001234716.2), *Panax japonicus* (ALB38664.1), *Glycine max* (NP_001236365.1), *Aralia elata* (ADC32654.1), *Lotus japonicus* (BAC56854.1), *Astragalus membranaceus* (ADW27427.1), *Ipomoea nil* (XP_019169346.1), *Betula platyphylla* (AKR76253.1), *Glycyrrhiza uralensis* (ACS66750.1), *Morus alba* (AOV62782.1), *Withania somnifera* (ADW78251.1), *Corchorus capsularis* (OMO74384.1), and *A. thaliana* (AtSQS1-AT4G34640.1; AtSQS2-AT4G34650.1); (b) Phylogenetic tree of *OfSQE* with other plants such as *Z. jujuba* (XP_015885090.1), *C. melo* (XP_016901349.1), *C. cajan* (KYP34787.1), *C. sativus* (XP_004141303.1), *S. indicum* (XP_011092466.1), *M. alba* (AOV62783.1), *Glycine soja* (KHN17192.1), *P. mume* (XP_008220536.1), *Citrus sinensis* (XP_006466420.1), *Panax vietnamensis* var. fuscidiscus (AIK23028.1), *Cephalotus follicularis* (GAV57995.1), *Theobroma cacao* (XP_007047610.2), *Gossypium arboreum* (KHG06672.1), *Citrus sinensis* (KDO78855.1), *Jatropha curcas* (NP_001295656.1), and *A. thaliana* (AtSQE2-AT2G22830.1; AtSQE3-AT4G37760.1; AtSQE1-AT1G58440.1); (c) Phylogenetic tree of *OfBETA-AS* with other plants such as *O. europaea* (BAF63702.1), *S. indicum* (XP_011096562.1), *Coffea canephora* (CDP12624.1), *D. hygrometricum* (KZV51042.1), *Catharanthus roseus* (AEX99665.1), *Ocimum basilicum* (AFH53506.1), *Calotropis procera* (AND78515.1), *Artemisia annua* (AHF22084.1), *Ilex asprella* var. Asprella (AIS39793.1), *Bacopa monnieri* (ADM86392.1), *Centella asiatica* (AAS01523.1), *Erythranthe guttata* (XP_012842229.1), *Daucus carota subsp. sativus* (XP_017230789.1), *Nelumbo nucifera* (XP_010245756.1), *Maesa lanceolata* (AHF49822.1), *Vitis vinifera* (XP_002270051.2), *Solanum lycopersicum* (NP_001234604.1), *Bupleurum chinense* (ABY90140.2), *Eucalyptus grandis* (XP_010063692.1), and *Kalopanax septemlobus* (ALO23119.1).

with the sequence GT and ended with AG, conforming to the GT/AG rules. Although the three genes contained multiple introns, no alternative splicing phenomenon was found in different organs and different flower development stages (Fig. S1).

3.3. Phylogenetic analysis of OfSQS, OfSQE, and OfBETA-AS

The BLAST tool was used to compare the obtained sequences with known sequences in the GenBank database. The results of the BLASTp analysis of the three amino acid sequences demonstrated that the proteins shared high similarity with those of other plants. For OfSQS, the protein sequence similarity with those of other plants was higher than 85%, such as 90% with *Sesamum indicum*, 89% with *Bacopa monnieri*, and 86% with *Camellia oleifera*, *Salvia miltiorrhiza*, and *Euphorbia tirucalli*. For OfSQE, the sequence similarity to those of other plants was 87% with *Ziziphus jujuba*, 86% with *Cucumis melo*, *Cajanus cajan and Prunus mume*, and 85% with *Cucumis sativus* and *Morus alba*. For OfBETA-AS, the sequence similarity was 99% with *O. europaea*, 91% with *S. indicum*, and 88% with *Dorcoceras hygrometricum*.

Cluster analysis of OfSQS with 20 SQS proteins from other plants showed that OfSQS was closest to the SQS of *S. indicum* and *B. monnieri*, and was also quite distant from SQS1 and SQS2 of *Arabidopsis thaliana* (Fig. 3a). For OfSQE, phylogenetic analysis of the deduced amino acid sequences revealed a closer evolutionary relationship with SiSQE and AtSQE1 (Fig. 3b). Moreover, OfBETA-AS shared the closest relationship with *O. europaea* (Fig. 3c).

3.4. Expression analysis in different organs

To clarify the tissue-specific expression patterns of *OfSQS*, *OfSQE*, and *OfBETA-AS* genes in the two cultivars of *O. fragrans*, i.e., "Boye" and "'Rixiang," qRT-PCR experiment was performed with different organs including roots, stems, leaves, and flowers (Fig. 4). The tissue-specific expression patterns of *OfSQS*, *OfSQE*, and *OfBETA-AS* among the examined tissues were predominant in the flowers and then in the leaves and stems, while the roots had the lowest expression level. The *OfSQS* transcript level in the opening flower was 8-fold higher than that in the "Rixiang" root, while it was only 2.4-fold higher in "Boye." In "Rixiang" and "Boye," the transcript levels of *OfSQE* in the opening

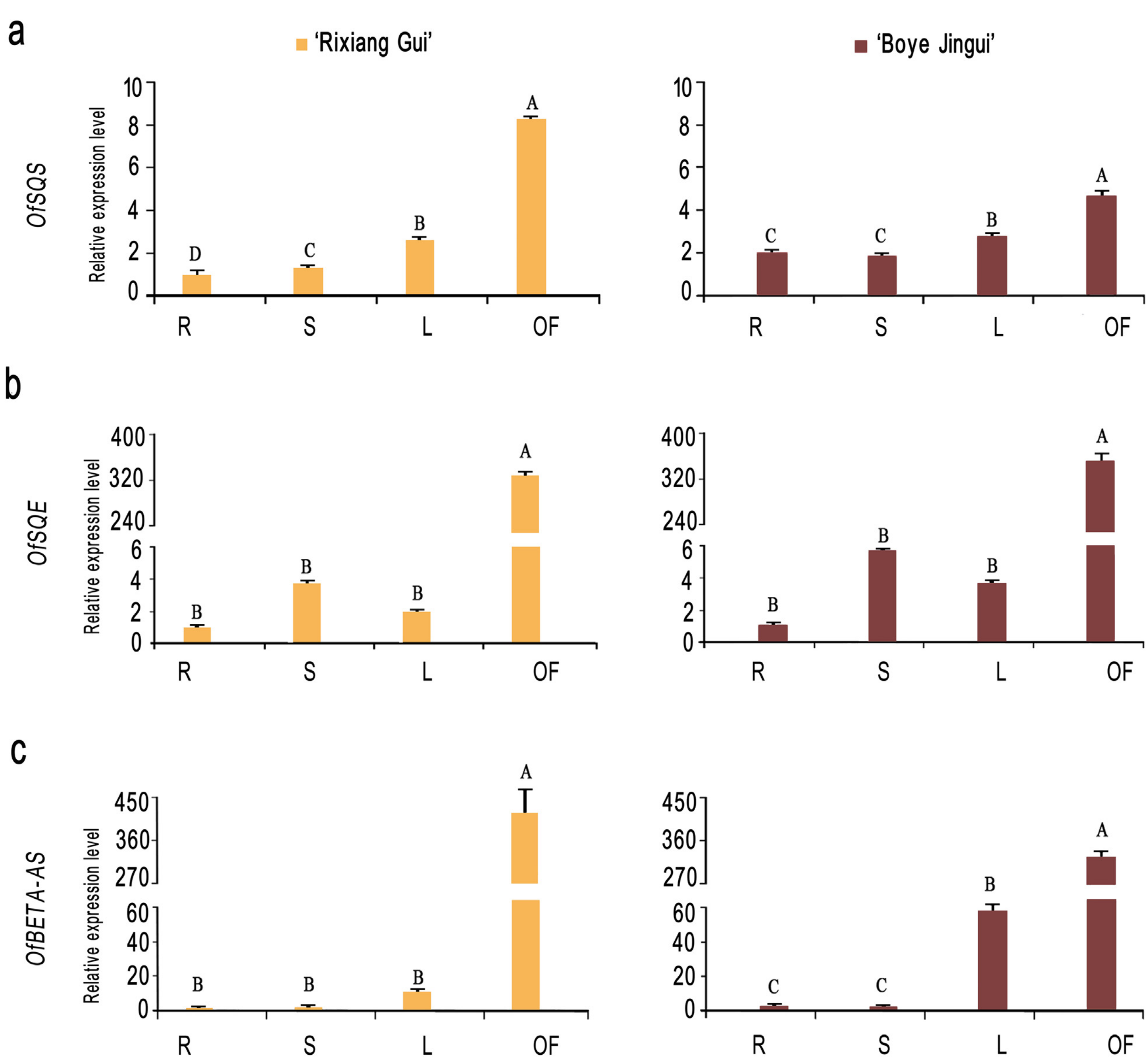

Fig. 4. Expression patterns of (a) *OfSQS*, (b) *OfSQE*, and (c) *OfBETA-AS* genes in four different organs of *O. fragrans*. R: root, S: stem, L: Leaf, and OF: opening flower. Data were presented as mean values with error bars indicating standard error. *Different letters* denote significant differences at the 0.01 level according to Tukey's test.

flowers were 328-fold and 346-fold higher than in the root, respectively. For the *OfBETA-AS* gene, the transcript level of "Rixiang" and "Boye" in the opening flower was 397-fold and 303-fold higher than that in the root, respectively.

3.5. Expression analysis during flower development

To analyze the expression patterns of *OfSQS*, *OfSQE*, and *OfBETA-AS* genes during flower development, qRT-PCR was performed at the five flowering stages: bud-pedicel stage (S1), bud-eye stage (S2), primary blooming stage (S3), full blooming stage (S4), and flower fading stage (S5) (Fig. 5). In "Rixiang," the expression level of the *OfSQS* gene did not show a significant change at the five stages, and the transcript level of *OfSQE* remained stable from S1 to S3 and then declined from S3 to S5. For *OfBETA-AS*, the expression level showed a regular downregulated trend at the five flowering stages. However, *OfSQS*, *OfSQE*, and *OfBETA-AS* showed the same expression trend in "Boye," which increased from S1 to S2 and decreased from S3 to S5.

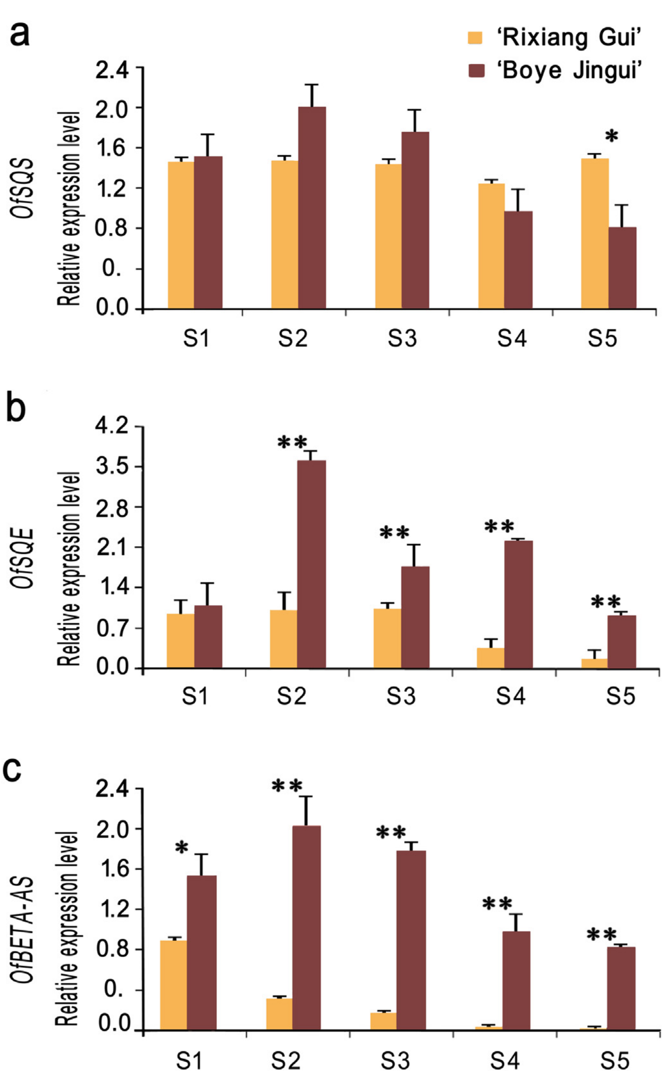

Fig. 5. Expression patterns of (a) *OfSQS*, (b) *OfSQE*, and (c) *OfBETA-AS* genes at the five different flowering stages of *O. fragrans*. These cDNA templates were isolated from bud-pedicel stage (S1), bud-eye stage (S2), primary blooming stage (S3), full blooming stage (S4) and flower fading stage (S5). Data were presented with error bars indicating standard error of three independent biological samples with three technical repeats. *Asterisks* (*) indicate a significant difference at $P < 0.05$ and *asterisks* (**) indicate a significant difference at $P < 0.01$ according to Tukey's test.

3.6. Subcellular localization

To confirm the subcellular localization of the OfBETA-AS protein, the *35s::GFP-OfBETA-AS* vector was constructed and infiltrated into the tobacco leaves. A significant fluorescence signal was detected in the cell nucleus and cell membrane (Fig. 6), which indicated that *OfBETA-AS* was a nucleus and membrane localized gene.

4. Discussion

Triterpenes are natural compounds that are mainly extracted from higher plants [18]. Genes involved in the biosynthesis of triterpenes have been identified and studied in numerous plants, including *A. thaliana* [19], *P. ginseng* [21], *Glycyrrhiza uralensis* [21], and *Rleutherococcus senticosus* [22]. However, little research on this topic has been performed on sweet osmanthus.

Analysis of amino acid sequence alignment revealed that among the 6 conserved domains of OfSQS, domains III, IV, and V were more conservative [23]. Moreover, a large amount of aspartic acid (DXXDD) was present in domain II and IV, which would influence the combination of FPP and Mg^{2+} [20,24]. Domain V was the binding domain of NADPH that controlled the transition to format squalene [24]. Domain VI was considered as the anchoring signal of the biological membrane [23]. The amino acid sequence (135–165) of OfSQE was indicated to be the putative flavin adenine dinucleotide (FAD) binding domains, which played an important role in the biosynthesis of triterpenoids [25,26]. Sequence analysis revealed that OfBETA-AS contained four QW motifs, the DCTAE motif, and the MWCYCR motif among the highly conserved regions, and it was reported that the QW and DCTAE motifs were present in all β-amyrin synthase and OSC superfamily [14,28]. The DCTAE motif could protonate the squalene epoxide ring, and it was thought to be responsible for initiating the cyclization reaction. The MWCYCR motif is related to the specific formation of β-amyrin [27].

Homology analysis showed that the deduced protein sequences shared high similarity with proteins of other plants. OfSQS, OfSQE, and OfBETA-AS showed 91%, 86%, and 91% identity with known proteins from *S. indicum*, respectively. This suggests a relationship of evolutionary conservatism between *O. fragrans* and *S. indicum*, both of which belong to Tubiflorae. Moreover, OfBETA-AS shared the highest similarity of 99% with *O. europaea*, which also belongs to the Olea family, and the functional expression of ObBETA-AS (*Ocimum basilicum*) led to the production of β-amyrin [28,29].

Analysis of the three genes revealed that the length and position of introns/exons in *OfSQS* and *OfBETA-AS* were consistent with previous reports of *AtSQS1* and *VvBETA-AS* (*Vitis vinifera*), respectively [30,31]. However, there were seven introns in *AtSQE6* [32]. The intron numbers varied depending on the species. No alternative splicing phenomenon was identified in different organs during flower developments, indicating that the functions of these three genes were stable.

Triterpenoids are the downstream products of the MVA and MEP pathways. Upregulation of *SQS* can significantly improve the expression level of *SQE* and *BETA-AS*, suggesting that *SQS* plays a pivotal role in regulating the biosynthesis of triterpenoids [33]. Highest expression levels of *SQS* were reported in the vegetative organs of *Withania somnifera* and *Stevia rebaudiana* [34,35]. In addition, some similar observations found higher *SQS* expression in the roots of other plants, such as *P. ginseng* and *Medicago truncatula* [10,18,20]. However, the transcriptome results of this study showed that *OfSQS* was constitutively expressed in the root, stem, leaf, and flower tissues, with the highest expression in the flowers in both "'Rixiang" and "Boye"; this finding was different from the results obtained for *Aralia elata* flowers that showed lower level of expression [36]. Therefore, the high expression might lead to the accumulation of specific triterpenoids in flowers of *O. fragrans*.

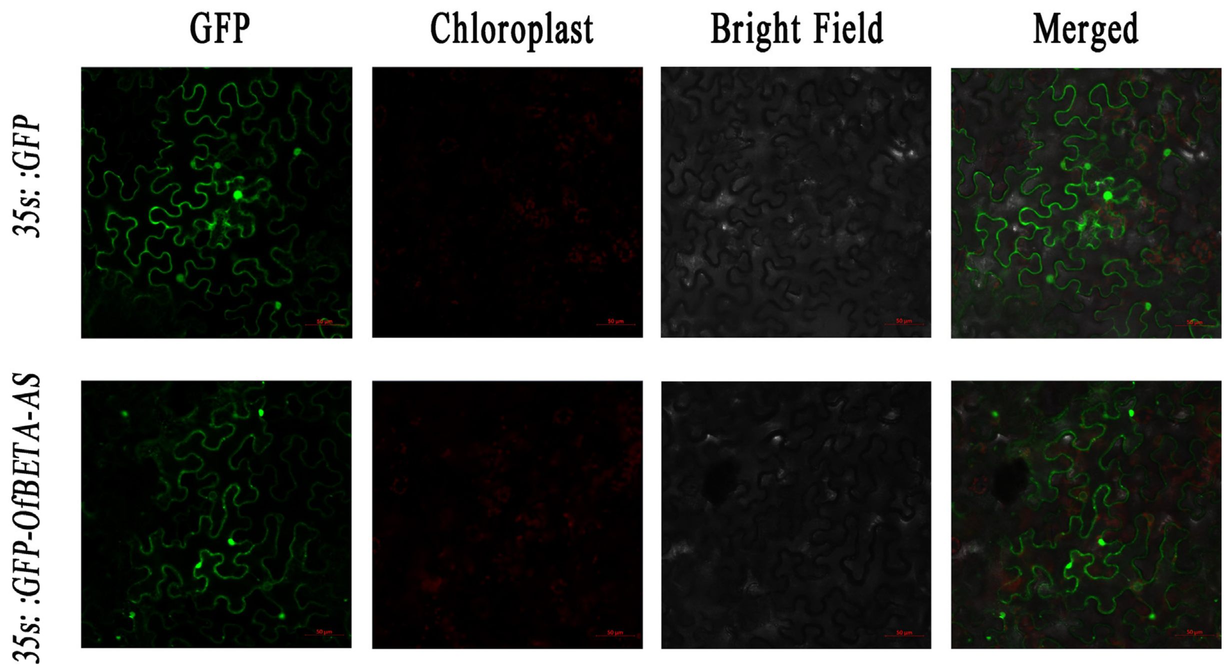

Fig. 6. Subcellular localization of *OfBETA-AS* in leaves of tobacco. *OfBETA-AS* fused with the GFP and GFP positive control were infiltrated into the leaves of tobacco through *Agrobacterium tumefaciens* strain EHA105. After incubation at 21°C with 14 h photoperiod for 2 d, GFP fluorescence signals were observed by a laser scanning confocal microscope. Scale bar 50 μm.

It was reported that the content of essential oil of *O. fragrans* was the highest at the initial flowering stage [7]. The three genes shared the same expression patterns in "Boye," showing considerable variations with a clear peak in the S2 stage, suggesting that the three genes were involved in the synthesis of essential oil in sweet osmanthus. Previous studies showed that the aroma of "Boye" was stronger than that of "Rixiang", and in this study, the expression level of the three genes was higher in "Boye" than in "Rixiang". This indicates that further study of the relationship between these genes and the fragrance of sweet osmanthus is needed. Furthermore, the expression level of *OfSQS* showed no significant difference among the five stages; this result was consistent with that obtained in *W. somnifera* [34]. However, *OfSQE* and *OfBETA-AS* showed a typical trend of change at the five stages. Thus, *OfSQE* and *OfBETA-AS* might play an important role in the formation of triterpenes in sweet osmanthus. In addition, the role of *OfBETA-AS* in the formation of triterpenes is more critical because the difference in the expression level of *OfSQE* at stage S1 was not significant in "Boye" and "Rixiang," while *OfBETA-AS* showed significantly different patterns during the five stages. The OfBETA-AS: GFP fusion protein was detected in the nucleus and membrane, which accorded with the localization characteristics of structural proteins.

5. Conclusions

In this study, three triterpenoid pathway genes, *OfSQS*, *OfSQE*, and *OfBETA-AS*, from *O. fragrans* were cloned, and the structure of the exons/introns of the three genes was further analyzed. Phylogenetic analysis revealed an evolutionarily conserved relationship between genes of *O. fragrans* and other plants. The functions of the three genes were stable, and no alternative splicing phenomenon was found. The qRT-PCR results showed that the *OfSQS*, *OfSQE*, and *OfBETA-AS* genes had a clear flower-specific expression pattern, supporting the hypothesis that the high expression might result in the specific accumulation of triterpenes in the flower of *O. fragrans.* The expression level of the *OfSQS* gene did not show a significant change at the five flowering stages. However, significantly different expression patterns of the *OfSQE* gene in cultivars "Boye" and "Rixiang" was detected from S2 to S5 stages. In addition, the expression patterns of the *OfBETA-AS* gene in cultivars "Boye" and "Rixiang" were significantly different during the whole flowering stages, thus showing that the *OfBETA-AS* gene had an important influence on the production and accumulation of triterpenoids in two *Osmanthus* cultivars from the very beginning. Therefore, *OfBETA-AS* might play a more critical role in the synthesis of triterpenoids. Functional research on the genes will be conducted in our following work.

Supplementary material

https://doi.org/10.1016/j.ejbt.2018.08.007

Conflict of interest

The authors declared that there no conflicts of interest concerning the publication of this paper.

Acknowledgments

This work was funded by the Application Demonstration of the Key Technology and Innovation in Breeding the Six Precious Colored Tree Species of Jiangsu Province (No. BE2017375), the top-notch Academic Programs Project of Jiangsu Higher Education Institutions (TAPP, PPZY2015A063), and the Innovative Plan of Academic Degree Graduate Students in Jiangsu Province.

References

[1] Huang B, Chen HQ, Shao LQ. The ethanol extract of *Osmanthus fragrans* attenuates *Porphyromonas gingivalis* lipopolysaccharide-stimulated inflammatory effect through the nuclear factor erythroid 2-related factor-mediated antioxidant signaling pathway. Arch Oral Biol 2015;60(7):1030–8. https://doi.org/10.1016/j.archoralbio.2015.02.026. PMID: 25912528.

[2] Zhang C, Wang YG, Fu JX, et al. Transcriptomic analysis and carotenogenic gene expression related to petal coloration in *Osmanthus fragrans* 'Yanhong Gui'. Trees 2016; 30(4):1207–23. https://doi.org/10.1007/s00468-016-1359-8.

[3] Baldermann S, Kato M, Kurosawa M, et al. Functional characterization of a carotenoid cleavage dioxygenase 1 and its relation to the carotenoid accumulation and volatile emission during the floral development of *Osmanthus fragrans* Lour. J Exp Bot 2010;618(11):2967–77. https://doi.org/10.1093/jxb/erq123. PMID: 20478967.

[4] Xu C, Li HG, Yang X, et al. Cloning and expression analysis of MEP pathway enzyme-encoding genes in *Osmanthus fragrans*. Genes 2016;7(10):78. https://doi.org/10.3390/genes7100078.

[5] Wang LM, Li MT, Jin WW, et al. Variations in the components of *Osmanthus fragrans* Lour. essential oil at different stages of flowering. Food Chem 2009;114(1):233–6. https://doi.org/10.1016/j.foodchem.2008.09.044. PMID: 27690108.

[6] Xiong LN, Mao SQ, Lu BY, et al. *Osmanthus fragrans* flower extract and acteoside protect against d-galactose-induced aging in an ICR mouse model. J Med Food 2016;19(1):54–61. https://doi.org/10.1089/jmf.2015.3462. PMID: 26181905.

[7] Yao WR, Zhang YZ, Chen Y, et al. Aroma enhancement and characterization of the absolute *Osmanthus fragrans* Lour. J Essent Oil Res 2010;22(2):97–102. https://doi.org/10.1080/10412905.2010.9700272.

[8] Tang Q, Ma XJ, Mo CM, et al. An efficient approach to finding *Siraitia grosvenorii* triterpene biosynthetic genes by RNA-seq and digital gene expression analysis. BMC Genomics 2011;12:343. https://doi.org/10.1186/1471-2164-12-343. PMID: 21729270.

[9] Abe I, Rohmer M, Prestwich GD. Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. Chem Rev 1993;93(6):2189–206. https://doi.org/10.1021/cr00022a009.

[10] Lee MH, Jeong JH, Seo JW, et al. Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* overexpressing squalene synthase gene. Plant Cell Physiol 2004;45(8):976–84. https://doi.org/10.1093/pcp/pch126. PMID: 15356323.

[11] Kim YS, Cho JH, Park S, et al. Gene regulation patterns in triterpene biosynthetic pathway driven by overexpression of squalene synthase and methyl jasmonate elicitation in *Bupleurum falcatum*. Planta 2011;233(2):343–55. https://doi.org/10.1007/s00425-010-1292-9. PMID: 21053012.

[12] Ferrer M, Chernikova TN, Timmis KN, et al. Expression of a temperature-sensitive esterase in a novel chaperone-based *Escherichia coli* strain. Appl Environ Microbiol 2004;70:4499–504. https://doi.org/10.1128/AEM.70.8.4499-4504.2004. PMid 15294778.

[13] Gao K, Wu SR, Wang L, et al. Cloning and analysis of β-amyrin synthase gene in *Bupleurum chinense*. Genes Genomics 2015;37(9):767–74. https://doi.org/10.1007/s13258-015-0307-0.

[14] Basyuni M, Oku H, Tsujimoto E, et al. Triterpene synthases from the *Okinawan mangrove* tribe, Rhizophoraceae. FEBS J 2007;274:5028–42. https://doi.org/10.1111/j.1742-4658.2007.06025.x PMID: 17803686.

[15] Yue YM, Wang JM, Kang WY. *Osmanthus fragrans* triterpenoids and hypolipidemic effect. Chin J Exp Tradit Med Formulae 2013;19:126–8.

[16] Zhang C, Fu JX, Wang YG, et al. Identification of suitable reference genes for gene expression normalization in the quantitative Real-Time PCR analysis of Sweet Osmanthus (*Osmanthus fragrans* Lour.). PLoS One 2015;10:e0136355. https://doi.org/10.1371/journal.pone.0136355. PMID: 26302211.

[17] Yue YZ, Yin CQ, Rui G, et al. An anther-specific gene *PhGRP* is regulated by PhMYC2 and causes male sterility when overexpressed in petunia anthers. Plant Cell Rep 2017;36(9):1401–15. https://doi.org/10.1007/s00299-017-2163-7. PMID: 28597062.

[18] Iturbe-Ormaetxe I, Haralampidis K, Papadopoulou K, et al. Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus*. Plant Mol Biol 2003;51(5):731–43. https://doi.org/10.1023/A:1022519709298. PMID: 12683345.

[19] Mirjalili MH, Moyano E, Bonfill M, et al. Overexpression of the *Arabidopsis thaliana* squalene synthase gene in *Withania coagulans* hairy root cultures. Biol Plant 2011;55(2):357–60. https://doi.org/10.1007/s10535-011-0054-2.

[20] Choi DW, Jung J, Ha YI, et al. Analysis of transcripts in methyl jasmonate-treated ginseng hairy roots to identify genes involved in the biosynthesis of ginsenosides and other secondary metabolites. Plant Cell Rep 2005;23(8):557–66. https://doi.org/10.1007/s00299-004-0845-4. PMID: 15538577.

[21] Liu Y, Zhang N, Chen HH, et al. Cloning and characterization of two cDNA sequences coding squalene synthase involved in glycyrrhizic acid biosynthesis in *Glycyrrhiza uralensis*. Frontier and Future development of Information Technology in Medicine and Education. Lect Notes Elect Eng 2014;269:329–42. https://doi.org/10.1007/978-94-007-7618-0_32.

[22] Seo JW, Jeong JH, Shin CG, et al. Overexpression of squalene synthase in *Eleutherococcus senticosus* increases phytosterol and triterpene accumulation. Phytochemistry 2005;66(8):869–77. https://doi.org/10.1016/j.phytochem.2005.02.016.

[23] Robinson GW, Tsay YH, Kienzle BK, et al. Conservation between human and fungal squalene synthetases: similarities in structure, function, and regulation. Mol Cell Biol 1993;13:2706–17. https://doi.org/10.1128/MCB.13.5.2706. PMid 8474436.

[24] Tansey TR, Shechter I. Squalene synthase: structure and regulation. Prog Nucleic Acid Res Mol Biol 2000;65:157–95. https://doi.org/10.1016/S0079-6603(00)65005-5. PMid 11008488.

[25] Han JY, In JG, Kwon YS, et al. Regulation of ginsenoside and phytosterol biosynthesis by RNA interferences of squalene epoxidase gene in *Panax ginseng*. Phytochemistry 2010;71(1):36–46. https://doi.org/10.1016/j.phytochem.2009.09.031. PMid 19857882.

[26] Abe I, Prestwich GD. Identification of the active site of vertebrate oxidosqualene cyclase. Lipids 1995;30:231–4. https://doi.org/10.1007/BF02537826. PMid 7791531.

[27] Jin ML, Lee DY, Um Y, et al. Isolation and characterization of an oxidosqualene cyclase gene encoding a β-amyrin synthase involved in *Polygala tenuifolia* Willd. saponin biosynthesis. Plant Cell Rep 2014;33(3):511–9. https://doi.org/10.1007/s00299-013-1554-7. PMID: 24420413.

[28] Saimaru H, Orihara Y, Tansakul P, et al. Production of triterpene acids by cell suspension cultures of *Olea europaea*. Chem Pharm Bull 2007;55(5):784–8. https://doi.org/10.1248/cpb.55.784. PMID: 17473469.

[29] Misra RC, Maiti P, Chanotiya CS, et al. Methyl jasmonate-elicited transcriptional responses and pentacyclic triterpene biosynthesis in sweet basil. Plant Physiol 2014;164:1028–44. https://doi.org/10.1104/pp.113.232884. PMid 24367017.

[30] Kribii R, Arró M, Del AA, et al. Cloning and characterization of the *Arabidopsis thaliana* SQS1 gene encoding squalene synthase-involvement of the C-terminal region of the enzyme in the channeling of squalene through the sterol pathway. Eur J Biochem 1997;249(1):61–9. https://doi.org/10.1111/j.1432-1033.1997.00061.x. PMID: 9363754.

[31] Jaillon O, Aury JM, Noel B, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 2007;449(7161):463–7https://doi.org/10.1038/nature06148. PMID: 17721507.

[32] Tabata S, Kaneko T, Nakamura Y, et al. Sequence and analysis of chromosome 5 of the plant *Arabidopsis thaliana*. Nature 2000;408:823–6. https://doi.org/10.1038/35048507. PMID: 11130714.

[33] Braga MV, Urbina JA, Souza WD. Effects of squalene synthase inhibitors on the growth and ultrastructure of *Trypanosoma cruzi*. Int J Antimicrob Agents 2004;24:72–8. https://doi.org/10.1016/j.ijantimicag.2003.12.009. PMID: 15225865.

[34] Bhat WW, Lattoo SK, Razdan S, et al. Molecular cloning, bacterial expression and promoter analysis of squalene synthase from *Withania somnifera* (L.) Dunal. Gene 2012;499:25–36. https://doi.org/10.1016/j.gene.2012.03.004. PMID: 22425978.

[35] Kumar H, Kaul K, Bajpai-Gupta S, et al. A comprehensive analysis of fifteen genes of steviol glycosides biosynthesis pathway in *Stevia rebaudiana* (Bertoni). Gene 2012;492:276–84. https://doi.org/10.1016/j.gene.2011.10.015. PMID: 22037480.

[36] Wu Y, Zou HD, Cheng H, et al. Cloning and characterization of a β-amyrin synthase gene from the medicinal tree *Aralia elata* (Araliaceae). Genet Mol Res 2012;11:2301–14. https://doi.org/10.4238/2012.August.13.4. PMID: 22911600.