Protoplasma
DOI 10.1007/s00709-015-0893-3

ORIGINAL ARTICLE

Molecular cloning and promoter analysis of squalene synthase and squalene epoxidase genes from *Betula platyphylla*

Mengyan Zhang[1] · Siyao Wang[1] · Jing Yin[1,2] · Chunxiao Li[1] · Yaguang Zhan[1,2] · Jialei Xiao[3] · Tian Liang[1] · Xin Li[1]

Received: 24 February 2015 / Accepted: 30 September 2015

Abstract *Betula platyphylla* is a rich repository of pharmacologically active secondary metabolites known as birch triterpenoids (TBP). Here, we cloned the squalene synthase (*SS*) and squalene epoxidase genetic (*SE*) sequences from *B. platyphylla* that encode the key enzymes that are involved in triterpenoid biosynthesis and analyzed the conserved domains and phylogenetics of their corresponding proteins. The full-length sequence of *BpSS* is 1588 bp with a poly-A tail, which contained an open reading frame (ORF) of 1241 bp that encoded a protein of 413 amino acids. Additionally, the *BpSE* full-length sequence of 2040 bp with a poly-A tail was also obtained, which contained an ORF of 1581 bp encoding a protein of 526 amino acids. Their organ-specific expression patterns in 4-week-old tissue culture seedlings of *B. platyphylla* were detected by real-time PCR and showed that they were all highly expressed in leaves, as compared to stem and root tissues. Additionaly, both *BpSS and BpSE* were enhanced following stimulation with ethephon and MeJA. The expression of *BpSS* was enhanced by ABA, whereas *BpSE* was not. The SA treatment did not affect the *BpSS* and *BpSE* transcripts notably. Using a genome walking approach, promoter sequences of 965 and 1193 bp, respectively, for *BpSS and BpSE* were isolated, and they revealed several key *cis*-regulatory elements known to be involved in the response to phytohormone and abiotic plant stress. We also found that the BpSS protein is localized in the cytoplasm. Opening reading frames of *BpSS* and *BpSE* were ligated into yeast expression plasmid pYES2 under control of *GAL1* promoter and introduced into the yeast INVScl1 strain. The transformants were cultured for 12 h, the squalene content of galactose-induced *BpSS* expression yeast cells was 13.2 times of control (empty vector control yeast cells) by high-performance liquid chromatography (HPLC) test method. And, the squalene epoxidase activity of induced *BpSE* expression yeast cell was about 11.8 times of control. These indicated that we cloned birch *BpSS* and *BpSE* that were indeed involved in the synthesis of triteropenoids. This is the first report wherein *SS* and *SE* from *B. platyphylla* were cloned and may be of significant interest to understand the regulatory role of *SS* and *SE* in the triterpenoids biosynthesis of *B. platyphylla*. This is the first report wherein *SS* and *SE* from *B. platyphylla* were cloned and may be of significant interest to understand the regulatory role of SS and SE in the biosynthesis of birch triterpenoids.

Keywords *Betula platyphylla* · Squalene synthase · Squalene epoxidase · Birch triterpenoids · Promoter

Handling Editor: Peter Nick

✉ Jing Yin
yinjing20135@163.com

[1] College of Life Science, Northeast Forestry University, Harbin 150040, China

[2] State Key Laboratory of Tree Genetic Breeding, Northeast Forestry University, Harbin 150040, China

[3] College of Life Science, Northeast Agriculture University, Harbin 150010, China

Introduction

Betula platyphylla contains important secondary metabolites, and one such group of pharmaceutical secondary metabolites is the Birch triterpenoids (TBP), which include betulin, betulinic acid, and oleanolic acid. These TBP have great potential in the development of anticancer and antihuman immunodeficiency virus (HIV) therapeutics (SUN et al. 2002; Patočka 2003; Zuco et al. 2002) and in the treatment of

Published online: 10 October 2015

cardiovascular diseases and novel liver condition therapies (Alakurtti et al. 2006; Zhang et al. 2003; Dang et al. 2012). In recent years, with the growing rise of natural medicine, natural plant resources are becoming increasingly scarce. Indeed, optimal ways to use our natural resources more effectively and rationally are fast becoming a pressing issue. As secondary plant metabolites in their natural state are at a low level, effective ways to upregulate natural plant metabolite production are related to the development of natural medicines, which is a key issue if it can be applied to more effective production. As triterpenoid biosynthesis-related genes continue to be cloned, it is possible to regulate terpene biosynthesis by employing genetic engineering technology and following metabolic engineering principles.

Currently, in terms of the biosynthetic pathway of triterpenoid compounds, Birch triterpenoids are synthesized through the mevalonic acid pathway (MVA) in a series of enzymatic reactions that involves the synthesis of farnesyl pyrophosphate (FPP), which leads to squalene synthesis and then production of squalene 2,3-oxide (Laule et al. 2003; Zhang et al. 2003). Following a series of cyclization, oxidation, hydroxylation, and glycosylation steps, a broad family of triterpenoids are synthesized (Haralampidis et al. 2002). Squalene is the common precursor of cholesterol, sterols, and triterpenoids (Kim et al. 2011a). Squalene synthase (SS) catalyzes the first enzymatic step that commits the carbon pool away from the central isoprenoid pathway toward the biosynthesis of phytosterols, brassinosteroids, and triterpenoids (Busquets et al. 2008). Squalene synthase catalyzes the condensation of two molecules of farnesyl diphosphate to form a linear 30-carbon compound called squalene. Squalene epoxidase (SE) catalyzes the first oxidation reaction through the triterpenoid biosynthesis pathway and the epoxidation of squalene to 2,3-oxidosqualene. The content and activity of SS and SE determine the yield of the subsequent product; thus, SS and SE are the key rate-limiting enzymes in terpenoid biosynthesis. In addition, SS was reported to play an important regulatory role in the phytosterol biosynthetic pathway.

Overexpression of the *SS* gene in *Panax ginseng* led to enhanced accumulation of phytosterols and triterpenes, thus documenting the important regulatory role of squalene synthase (Lee et al. 2004). Moreover, overexpression of the *SS* gene in *Bupleurum falcatum* resulted in enhanced production of both phytosterol and saikosaponins and increased messenger RNA (mRNA) accumulation of downstream genes including squalene epoxidase and cycloartenol synthase (Seo et al. 2005). Inhibitors that are used against SE inhibit sterol formation in fungi, and silencing of one of three isoforms of SE in *P. ginseng* decreases phytosterol formation, indicating that SE plays a pivotal role in the regulation (Han et al. 2010).

Plant squalene synthase (*SS*) and squalene epoxidase genes (*SE*) were characterized from *Glycyrrhiza uralensis* (Lu et al. 2007), *P. ginseng* (Kim et al. 2011a), and some other plant species but were not characterized in *Betula platyphylla*. As part of ongoing efforts to investigate the genes involved in Birch triterpenoid biosynthesis and their regulation, we report for the first time the isolation, cloning, tissue-specific and phytohormone treatment-associated expression of squalene synthase, and squalene epoxidase (designated as *BpSS* and *BpSE*) from *Betula platyphylla*. Promoter sequences of squalene synthase and squalene epoxidase have been isolated and analyzed. The investigation of *BpSS* and *BpSE* also includes their comparative expression in various tissues and under differential phytohormone treatments. We attempted to gain greater understanding of the regulatory role of squalene synthase and squalene epoxidase in Birch triterpenoid biosynthesis. The outcomes of this study have important significance in the context of exploring the molecular mechanism of triterpene biosynthesis and underpin attempts aimed at improving the content of Birch triterpenoids biotechnologically.

Materials and methods

Plant material and phytohormone treatment

Tissue culture seedlings of the birch (*Betula platyphylla*) were derived from the Northeast Forestry University, Harbin, China. The seedlings were used as source plant material for the present study. Different tissues that were used in this study were collected from 4-week-old tissue culture birch seedlings. For abscisic acid (ABA), ethephon, methyl jasmonate (MeJA), and salicylic acid (SA) treatments, the plants were mist sprayed all over the visible surfaces with ABA (10 μM), and ethephon (5 mM), MeJA (100 μM), and SA (50 mg/L) solution were similarly applied to both sides of the leaf until the liquid dripped from the leaves. Birch sapling leaves were harvested at 6, 12, 24, and 48 h after different hormone treatments. Fresh tissue samples were then immediately frozen in liquid nitrogen and kept at −80 °C for further analysis.

Cloning and sequencing of squalene synthase gene and squalene epoxidase gene

Total RNA was extracted from the cultivated seedlings of *Betula platyphylla* using the CTAB-based method. Absorbance values of the RNA samples were determined using a NanoDrop ND-1000 spectrophotometer. RNA quality was assessed by the optical density (OD) ratios that were measured at 260 nm/280 nm and by separation through 1 % agarose gels followed by visual evaluation of band integrity. Total RNA was treated with DNase I (Takara, Dalian, China) to eliminate genomic DNA present in the samples. For complementary DNA (cDNA) synthesis, 1 μg of DNase I-

treated total RNA was reverse-transcribed using Revert-aid premium reverse transcription kit with an oligo (dT) primer in a total reaction volume of 20 μL for 15 min at 42 °C, as described in the instruction manual (Takara, Dalian, China). The resultant reaction mixture was used as a template for PCR amplification.

A pair of degenerate oligonucleotide primers, SS-F/SS-R and SE-F/SE-R, was designed according to the conserved sequences of other plant *SS* and *SE* genes and was used for the amplification of the core cDNA fragments of *BpSS* and *BpSE*. Using the abovementioned *Betula platyphylla* cDNA as a template, degenerate PCR was carried out under the following conditions: 35 cycles of 94 °C for 45 s, 45 °C for 30 s, and 72 °C for 45 s followed by a final extension of 72 °C for 10 min. A 450-bp *BpSS* DNA core fragment and a 530-bp *BpSE* core fragment were recovered, then cloned into the pMD18-T vector (Takara, Dalian, China), and subjected to nucleotide sequencing. BLAST analysis showed that the 450- and 530-bp sequenced fragments had high homology to other plant *SS* and *SE* genes. Both fragments were subsequently used for designing gene-specific primers for cloning the 5′ and 3′ ends of the *BpSS* and *BpSE* by RACE.

The 5′ and 3′ ends of the *BpSS* and *BpSE* cDNA were obtained by RACE, using the SMART™ RACE cDNA amplification kit (Clontech, USA). The 5′ RACE-Ready first-strand cDNA and 3′ RACE-Ready first-strand cDNA were synthesized following the protocol provided by the manufacturer. The 5′ end of the *BpSS* and *BpSE* cDNA was amplified using the 5′ end of the gene-specific primers (GSPs) and the universal primers that were provided with the kit. For the 5′ RACE PCR amplification, BpSS-5, and the UPM (i.e., Universal Primer-A Mix), BpSE-5 and UPM were used separately as primers and the 5′ RACE-Ready cDNA as the template. The 5′ RACE PCR amplification procedures were carried out under the following conditions: 3 min at 94 °C, 35 cycles of 45 s at 94 °C, 30 s at 60 °C, and 2 min at 72 °C, and a final step of 10 min at 72 °C. Similarly, the 3′ end of *BpSS* and *BpSE* cDNA were amplified separately using BpSS-3 gene-specific primers and UPM for PCR. Both of the 3′ RACE PCR amplification procedures were carried out under the following conditions: 3 min at 94 °C, 35 cycles of 45 s at 94 °C, 30 s at 65 °C, and 2 min at 72 °C, and a final step of 10 min at 72 °C. The PCR amplification products of both 5′ and 3′ RACE procedures were purified and cloned into the pMD18-T vector followed by DNA sequencing.

By aligning the 3′ RACE and 5′ RACE product sequences, the full-length cDNA products of *BpSS* and *BpSE* were obtained. The *BpSS* and *BpSE* open reading frame (ORF) was amplified using a single-stranded cDNA template and gene-specific primers BpSSFull-F/BpSSFull-R and BpSEFull-F/BpSEFull-R separately, which covered the start and the stop codon regions, respectively, which were then cloned into the pMD18-T vector.

Sequence analysis

The full-length nucleotide sequence was translated using the Translate tool (found at http://www.expasy.ch/tools/dna.html), and the properties of the deduced amino acid sequence were estimated using ProtParam and TMHMM software programs. Structurally and functionally important regions were identified in the deduced protein sequence at the National Center for Biotechnology Information (NCBI) BLAST Conserved Domains. In addition, the secondary structure was determined by using the SOPMA program.

Phylogenetic analysis

Protein sequences were retrieved from GenBank using the Blastp algorithm at the NCBI using *BpSS* and *BpSE* sequences as query entries, and then several *SS* and *SE* homologues from different organisms were subsequently selected. Sequences were aligned using the ClustalW software program found at the European Bioinformatics Institute (http://www.ebi.ac.uk) using default parameters. Neighbor-joining trees were constructed with MEGA6.0 software.

Tissue-specific and phytohormone treatment expression analysis by quantitative real-time analysis

Total RNA from the leaves, root, and stalk tissues and those plants treated with abscisic acid (ABA), ethephon, methyl jasmonate (MeJA), and salicylic acid (SA) were extracted as described above. DNA contamination was removed from the RNA samples using DNase I. RNA concentrations were measured at 260 nm by a spectrophotometer. The quality and purity of the preparations were determined at an OD 260/OD 280-nm absorption ratio (i.e., 1.8–2.0) and the integrity of the preparation was determined by electrophoresis through a 1.2 % agarose gel containing formaldehyde. About 1 μg total RNA was reverse-transcribed into first-strand cDNA using the reverse transcription enzyme (Takara, Dalian, China) with the oligo-(dT) primers and done according to the manufacturer's protocols. Real-time qPCR reactions were performed in triplicate using SYBR Premix Ex Taq polymerase (Takara, Dalian, China) in 96-well optical plates using an ABI StepOne Real-time qPCR system. The PCR reaction (20 μL) included 2 μL of the cDNA template, 200 nM each of the primers, and 10 μL of the SYBR Premix Ex Taq polymerase. The cycling parameters were 94 °C for 30 s, followed by 40 cycles of 94 °C for 50 s, 58 °C for 30 s, and 72 °C for 40 s. The specificity of each primer pair was validated by a dissociation curve (i.e., a single peak was observed for each primer pair). Raw data of relative quantification values were calculated using the $2^{-\Delta\Delta CT}$ algorithm approach (Livak and Schmittgen 2001). The target gene expression level was plotted as $\log_{10}$ RQ (i.e., $RQ=2^{-\Delta\Delta CT}$) using the housekeeping β-tubulin (Tu) gene as the internal controls (Yin et al. 2012).

Isolation of BpSS and BpSE promoters

Genome Walking products were constructed following the user's manual guidance that was provided by the manufacturer (Genome Walking™ Kit, Takara, Dalian). *Betula platyphylla* genomic DNA was isolated from fresh leaves using the CTAB method. The isolated DNA was also confirmed for purity considerations on a 0.8 % agarose gel. The primary PCR assay was performed using AP1 (as provided in the kit) and SSSP1, with SESP1 as the primers, and construction of the DNA template via the following protocol: 5 cycles at 94 °C for 25 s and 70 °C for 3 min, with 30 cycles at 94 °C for 25 s, and 65 °C for 3 min, and a final reaction condition at 67 °C for 7 min. The product was diluted 10-fold and used in the nested PCR assay, which was performed using AP1 and SSSP2, with SESP2 as the nested primers under the following conditions: 5 cycles at 94 °C for 25 s and 70 °C for 3 min, with 30 cycles at 94 °C for 25 s and 70 °C for 3 min, and 30 cycles at 94 °C for 25 s and 67 °C for 3 min, followed by a final reaction of 67 °C for 7 min. The PCR products were purified and cloned into the pMD18-T vector and subsequently DNA sequenced. Putative *cis*-acting elements upstream of the start codon of both the *BpSS and BpSE* cDNA were identified by interrogating the Plant Care and PLACE publically available databases.

Subcellular localization analysis of transiently expressed BpSS protein

The full-length BpSS coding region was amplified using primers BpSS-SalI (5′ CGGTCGACCGATGGGGTGCGTG TGTTTCGTAGT 3′, SalI site underlined) and BpSS-PstI (5′ CGCTGCAGCGACCAGCCAGCCATGAACTCAAGT 3′, PstI site underlined). The coding sequence of the gene was inserted in SalI and PstI sites of the pCAMBIA1303 vector (the vector contains 35S promoter and GFP gene sequence, Biovector, INC, USA). The fusion construct pCAMBIA1303-BpSS (p35S BpSS::GFP) and control pCAMBIA1303 (p35S-GFP) were mobilized in *Agrobacterium tumefaciens* strain GV3101 by triparental hybridization method. Agrobacterium-mediated transient expression in onion epidermal cells (Nagegowda 2010). These onion epidermal cells were previously incubated on MS agar plates at 28 °C under light illumination for 24 h. After transformation, the tissues were placed on the same agar plates and incubated at 25 °C for 48 h and a photoperiod of 16/8 h (day/night). The transient expression image of the BpSS/GFP fusion protein under blue excitation (475 nm) was observed by Axioplan 2 imaging MOT Software acquisition and analysis using a fluorescence microscope (Axioskop2plusFL, Zess, Germany).

Expression BpSS and BpSE in yeast INVScl1 strain

The ORF of *BpSS* and *BpSE* PCR products was digested with appropriate restriction enzyme sites and ligated into the corresponding sites of pYES2 (Invitrogen) to construct the plasmid pYES2-*BpSS* and pYES2-BpSE. Each plasmid and empty vector control plasmid were transferred to INVScl1 strain (Clontech, USA), using PEG/LiAc method. The transformants were inoculated in 25-ml synthetic complete medium without uracil (SC-U) and incubated at 30 °C for 2 days. Then, media were changed to SC-U with the same supplement and 2 % galactose in place of glucose which is used to induce the GLA1 promoter. Cells of the INVScl1 strain bearing pYES2-BpSS, pYES2-BpSE, and empty vector control plasmid respectively were cultured at 30 °C for 12 h with shaking.

HPLC analysis of BpSS products accumulated in INVScl1 transformants

The transformants were cultured for 12 h, and expression of *BpSS* was induced at the same condition as described above. Then, cells were collected. The cellular squalene was analyzed according to Lu et al. (2004)with some modifications. Dried cells (50 mg) were saponified by 5 ml 10 % (*w*/*v*) KOH-75 % (*v*/*v*) ethanol solution at 50 °C for 15 min. The mixture was extracted with 4 ml of hexane for three times. The hexane layer was collected and evaporated to dryness. The residue was dissolved in 2.5 ml of acetonitrile for subsequent high-performance liquid chromatography (HPLC) analysis. Waters 1525 HPLC equipped with a Waters 2489 UV/Visible detector were used. The separation of squalene was achieved by using a mobile phase of 100 % acetonitrile at a constant flow rate of 1.0 ml/min. The squalene was monitored at wavelength of 210 nm. The peak of squalene was identified based on the retention time and UV spectra against squalene standard. The quantification was done according to the external calibration graphs obtained from peak areas vs. different concentrations of squalene standards (Sigma, USA). All samples and standards were filtered through 0.45-μm membrane filters before injection.

Detection of squalene epoxidase activity in INVScl1 transformants

The cells were harvested by centrifugation at 1000×*g* for 5 min, resuspended in 0.5 ml of 0.1 M potassium phosphate buffer (pH 7.4 containing 0.45 M sucrose, 1 mM EDTA, and 1 mM dithiothreitol), and broken by vortex with acid-washed glass beads. Cell homogenates were centrifuged at 13,300×*g* for 15 min, and the supernatant served as the enzyme preparation. The squalene epoxidase activity of cells of the INVScl1 strain bearing pYES2-BpSE and empty vector control plasmid was detected using the Squalene epoxidase activity

quantitative colcrimetric assay kit (GENMED, USA). The substrates and buffer which are required in the reaction were supplied by the kit, and each reaction adds 100 μg yeast protein and then was incubated at 37 °C for 30 min. The reaction was treated with sulfuric acid and formaldehyde to show color, the results were monitored by spectrophotometer at 400 nm, and squalene epoxidase activity was calculated by standard curve.

Results

Cloning and sequence analysis of BpSS and BpSE

A full-length cDNA of *BpSS and BpSE* was obtained from the tissue culture seedlings of *Betula platyphylla* by degenerate PCR and RACE procedures. Based upon the conserved amino acid domains of squalene synthase and squalene epoxidase, forward degenerate primers and reverse degenerate primers were designed. Approximately 450- and 530-bp core fragments of *BpSS and BpSE* were obtained from the initial RT-PCR reaction. To obtain 5′ and 3′ ends of the core amplified fragment, the RACE technique was used. Gene-specific primers used for carrying out RACE were designed using partial length sequences that were obtained from the amplified core fragment (Table 1). After assembling the 5′ and 3′ RACE sequences, a *BpSS* cDNA sequence of 1588 bp with a poly-A tail was obtained that contained an ORF of 1241 bp encoding a protein of 413 amino acids (Fig. 1). A 52-bp 5′ untranslated region (UTR) was upstream of the start codon, and the coding region was followed by a 265 bp 3′ UTR. Similarly, a *BpSE* cDNA sequence of 2040 bp with a poly-A tail was obtained that contained an ORF of 1581 bp encoding a protein of 526 amino acids (Fig. 2). A 213-bp 5′ UTR was upstream of the start codon and coding region that was followed by a 234-bp 3′ UTR.

Analysis of the deduced protein sequence

The calculated molecular mass and predicted isoelectric point of the *BpSS*-deduced polypeptide sequence were 47.57 kDA and 7.90, respectively. The amino acid sequence of BpSS shared 88 % homology with that of *Glycine max*, *Camellia oleifera*, and *Glycyrrhiza uralensis*, and 85 % with *Diospyros kaki*, *Theobroma cacao*, and *Eleutherococcus senticosus*. Multiple alignment of *BpSS* deduced amino acid sequences with other *SS* proteins suggested the presence of six (I–VI) highly conserved SS signature domains of 14–23 amino acids that are important for catalytic/functional activity of squalene synthase (Fig. 3). Conserved domains (NCBI) predict that the BpSS-deduced polypeptide sequence contains a Trans_IPPS_HH conserved domain, a squal_synth sequence of approximately 35–367 aa, and a SQS_PSY sequence of approximately 40–329 aa. Neighbor-joining phylogenetic trees were constructed with the MEGA 6.0 software program from different organisms that were retrieved from the NCBI GenBank database (Fig. 4). The degree of relatedness correlated well with the amino acid sequence similarity among these proteins. The *BpSS* was included in a separate clade consisting of members of the *Theobroma cacao* family to

Table 1 List of primers used in the study

No.	Primer name	Direction	Sequence (5′→3′)
1.	SS-F	Forward	GC(A/T)CTTGA(C/T)ACTGTTGAGG
2.	SS-R	Reverse	GGTATCTCATT(T/A)AT(G/A)TCCTC
3.	SE-F	Forward	GG(G/T)CT(A/T)GA(G/A)GATTGT(G/T)TG
4.	SE-R	Reverse	C(A/C)CCATT(T/A)GAAATAGAAGG
5.	BpSS-5	Reverse	TGCGAGTGGCATTTTTCGTGTGG
6.	BpSS-3	Forward	GCCATTCCTGCACCCATTCTTTTGG
7.	BpSE-5	Reverse	TCGGATTGCAAAGAGTGCGGCGCA
8.	BpSE-3	Forward	CGGATGTGGCTGGGAGGAGCTTCCA
9.	UPM	Forward/reverse	Long: 5′-CTAAT ACGAC TCACT ATAGG GCAAGCAGTG GTATC AACGC AGAGT-3′ Short: 5′-CTAAT ACGAC TCACT ATAGG GC-3′
10.	BpSSFull-F	Forward	ATGGGGTGCGTGTGTTTCGTAGT
11.	BpSSFull-R	Reverse	ACCAGCCAGCCATGAACTCAAGT
12.	BpSEFull-F	Forward	AAGGAAAAACAAATCACAGCCA
13.	BpSEFull-R	Reverse	CAAGTATTCACCATCCCCTCAAC
14.	SSSP1	Reverse	GTTGAGCATGGTGTAGCAGAAGGC
15.	SSSP2	Reverse	CCAAACTCCCCATTGCTCCTCTCA
16.	SESP1	Reverse	GCTCTTCATGCCCTCGCTTCTCG
17.	SESP2	Reverse	AGGCGAATATCCACCCAAACAAG

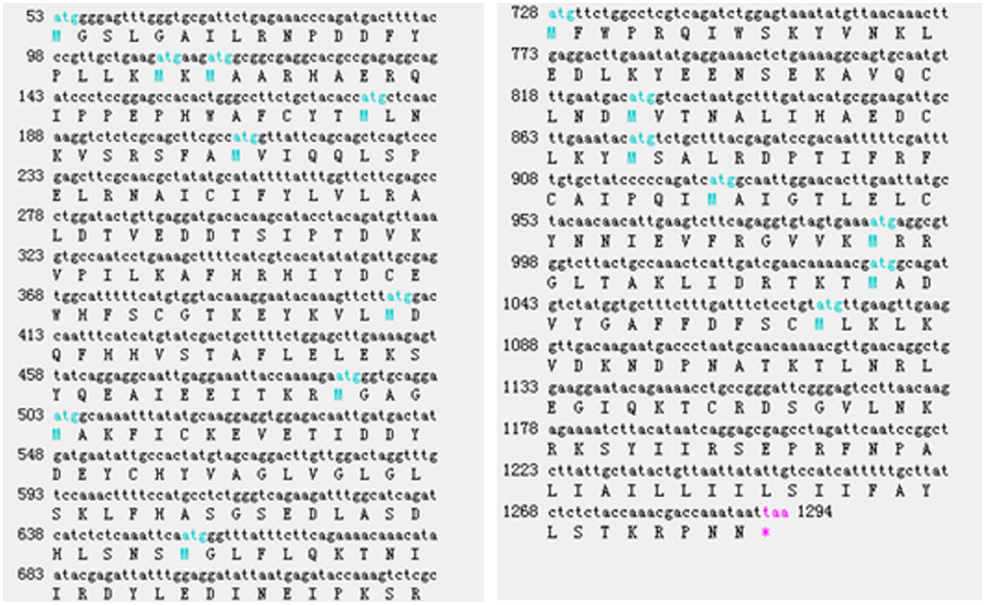

Fig. 1 Nucleotide and deduced amino acid sequence of *BpSS* from *Betula platyphylla.* The ATG start codon at position 53 and the TAA stop codon at position 1294

which *BpSS* belongs. Amino acid alignment results, in conjunction with phylogenetic analysis, reflected that *BpSS* cDNA encoded the full-length functional *SS* polypeptide from *Betula platyphylla*. Also, secondary structural analysis of the *BpSS* protein by the SOPMA program revealed that *BpSS* consist of α-helixes (66.83 %), β-turns 93.63 % that were joined by extended strands (5.57 %), and random coils (23.97 %). Moreover, SS is a conserved class of catalyzing synthase, and two transmembrane helices in the deduced BpSS protein are positioned at residues 171–187 and at 389–406 aa.

The calculated molecular mass and predicted isoelectric point of the *BpSE*-deduced polypeptide sequence were 57.11 kDA and 8.83, respectively. The amino acid

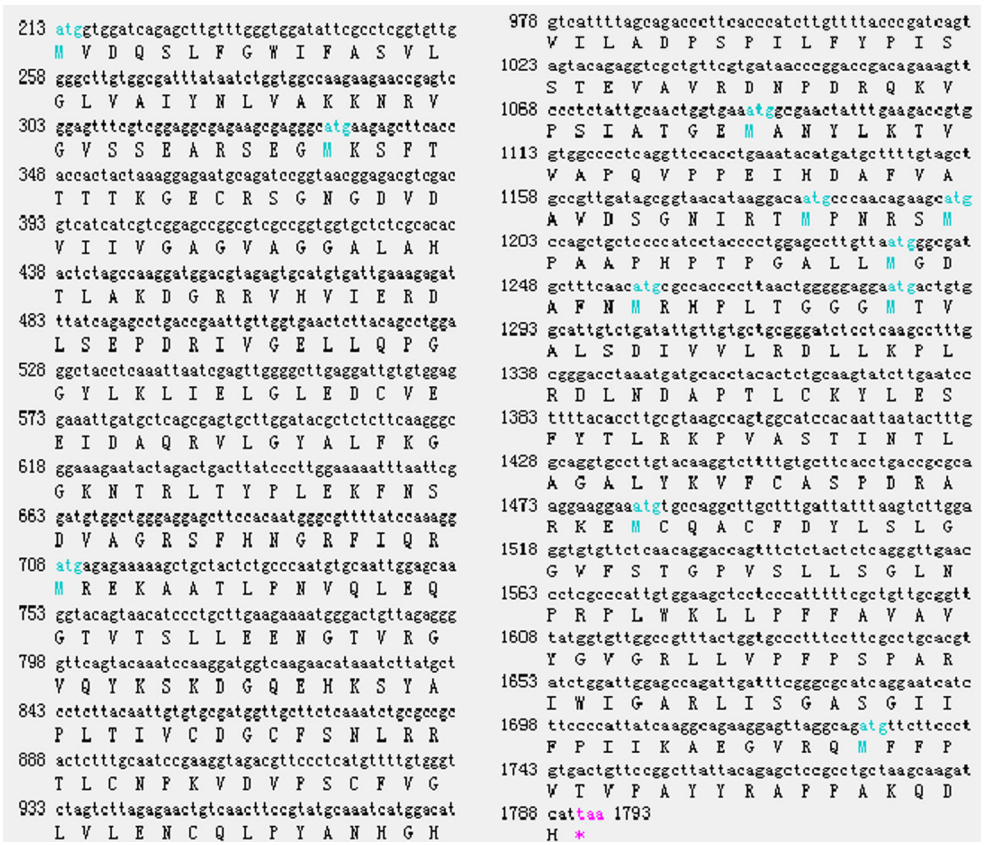

Fig. 2 Nucleotide and deduced amino acid sequence of *BpSE* from *Betula platyphylla.* The ATG start codon at position 213 and the TAA stop position at 1793

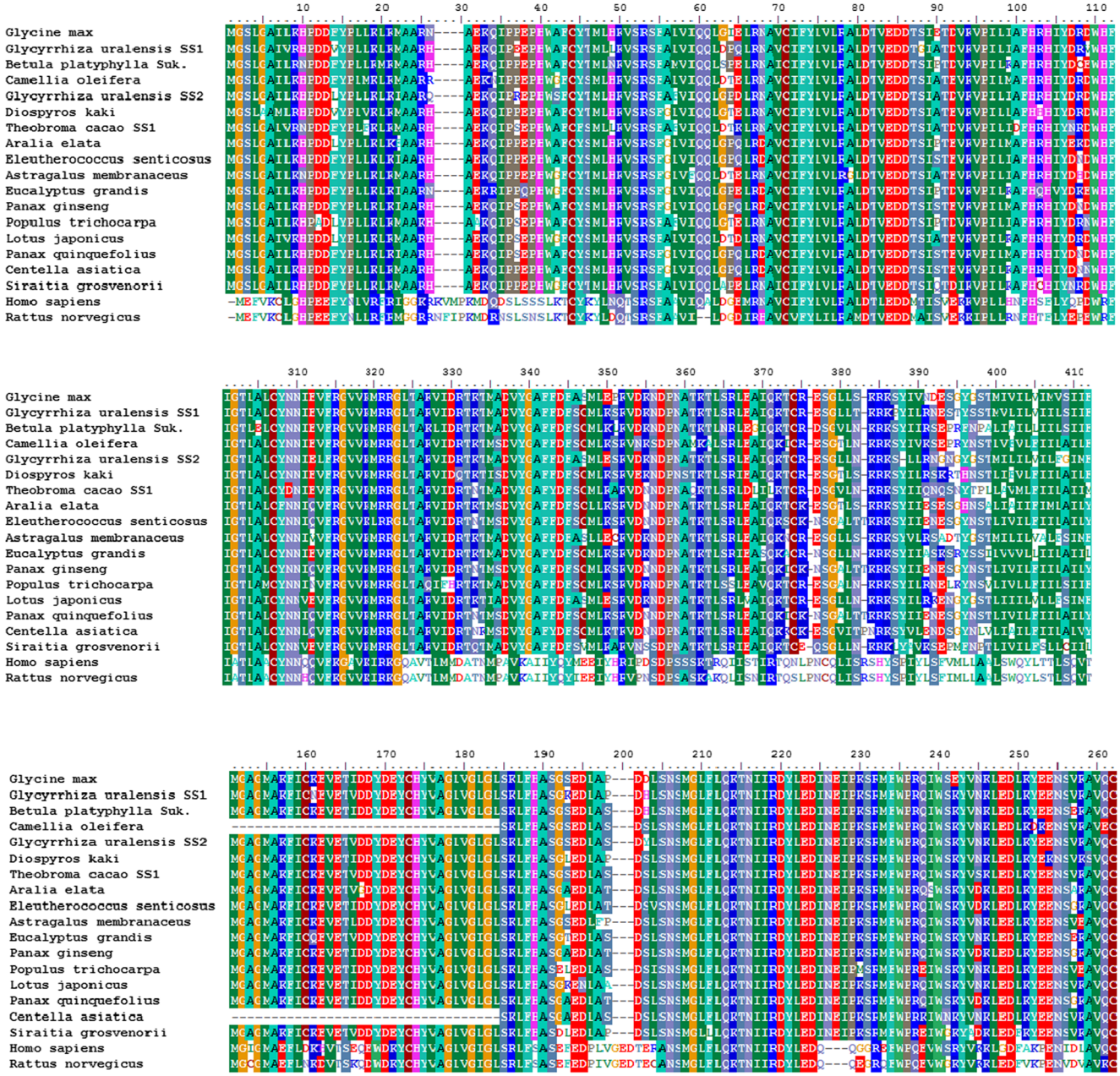

Fig. 3 Multiple alignment of the deduced amino acid sequence of *BpSS* with known *SS*s of the other organisms using the ClustalW multiple alignment tool. The aligned sequences are derived from *Glycine max* (AB007503), *Glycyrrhiza uralensis* SS1 (ADG36702.1), *Camellia oleifera* (AGB05603.1), *Glycyrrhiza uralensis* SS2 (ADG36719.1), *Dioapyros kaki* (ACN69082.1), *Theobroma cacao* (XM_007041985.1), *Aralia elata* (ADC32654.1), *Eleutherococcus senticusus* (AEA41712.1), *Astragalus membranaceus* (ADW27427.1), *Eucalyptus grandis* (LOC104453372), *Panax ginseng* (ACV88718.1), *Populus trichocarpa* (XP_002313765.1), *Lotus japonicas* (BAC56854.1), *Panax quinquefolius* (CAJ58418.1), *Centella asitica* (AAV58897.1), *Siraitia grosvenorii* (AEM42980.1), *Homo sapiens* (NP_00453.3), and *Rattus norvegicus* (AAA42179.1)

sequence of BpSE shared 85 % homology with that of *Theobroma cacao*, *Jatropha curcas*, 80 % homology with *Gynostemma pentaphyllum*, and *Panax vietnamensis var.* Multiple alignment of the *BpSE*-deduced amino acid sequence with other *SE* proteins is shown in Fig. 5. Conserved domains (NCBI) that predicted the BpSE-deduced polypeptide sequence contained a SDR superfamily of approximately 209–428 aa. The neighbor-joining phylogenetic tree was constructed with MEGA 6.0 software from different organisms that were retrieved from the NCBI GenBank database (Fig. 6). The degree of relatedness correlated well with the amino acid sequence similarity among these proteins. The *BpSE* was included in a separate clade that consisted of members of the *Withania somnifera* family to which *BpSS* belongs. Amino acid alignment results and the phylogenetic analysis informed that *BpSE* cDNA encoded

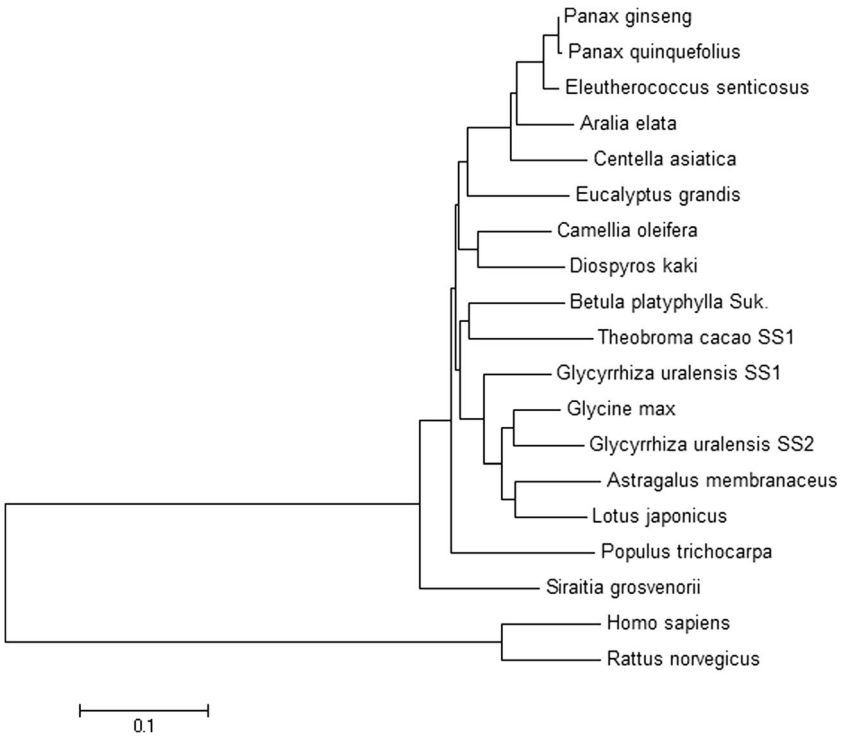

Fig. 4 Neighbor-joining phylogenetic tree constructed from the deduced amino acid sequences of *BpSS* from *Betula platyphylla* and *SS*s from other organisms using the MEGA 6.0 software program. *Numbers above the branches* indicate bootstrap values

the full-length functional *SE* polypeptide from *Betula platyphylla*. Secondary structural analysis of the *BpSE* protein by SOPMA program analysis revealed that *BpSS* consists of α-helixes (35.89 %) and β-turns (6.91 %) that were joined by extended strands (16.51 %) and random coils (40.69 %). The predicted BpSE protein possessed six transmembrane helices at positions 6–24, 59–78, 353–372, 453–472, and 482–499 aa.

Isolation and analysis of BpSS and BpSE promoter regions

Using genome walking methodology, we isolated 965 and 1193 bp fragments upstream of the start codon including the 189 bp 5′ UTR and the 145 bp 5′ UTR that corresponded to the putative promoter region of *BpSS* and *BpSE* (Figs. 7 and 8). They possessed a typically high A and T content that is commonly found in other plant promoters. Computational analysis using PlantCare and PLACE databases revealed several *cis* elements including potential TATA box sequences within the promoter regions of *BpSS* and *BpSE* at positions +889, −722, +693, +686, −573, +381, +262, −236, −121, +69 and +940, +286, +576, −575, respectively. Another consensus eukaryotic *cis* element known as the CAAT-box was also observed at −954, +616, +615, +471, +82, −85 and 751, +496, −863, −833, −862, −532, −385, and −615 in the promoter sequences of *BpSS* and *BpSE*. Furthermore, several stress-related *cis* elements were identified in the promoter regions of *BpSE*. MBS is an MYB recognition site that was identified during this analysis. Members of both the Myb and Myc family of proteins play an important role in plant responses to pathogens, low temperature extremes, and drought conditions. TC-rich repeats present in the promoters of many defense-related genes of other plants were also found in the promoter region of *BpSE* at positions −656 and −1162. In addition, several important phytohormone responsive elements were also predicted within the *BpSS* and *BpSE* promoter regions that included (1) the CGTCA motif that is involved in methyl-jasmonate responsiveness, (2) the TGA element that is involved in auxin responsiveness, (3) the GARE-motif that is involved in gibberellin responsiveness, and (4) the ABRE that is involved in abcisic acid responsiveness. The important *cis*-regulatory elements identified in both *BpSS* and *BpSE* along with their putative roles are listed in Tables 2 and 3.

Differential tissue-specific expression of both BpSS and BpSE under different phytohormone treatments

To investigate the *BpSS* and *BpSE* expression pattern in different tissues of *Betula platyphylla*, cDNA was prepared from the total RNA that was purified from the roots, stems, and leaves that were then used as the template in quantitative RT-PCR analysis. Results indicated that *BpSS* and *BpSE* were expressed in a constitutive manner in the tissues of roots, stems, and leaves, with the highest expression levels found in leaves and the lowest expression levels found in the roots (Fig. 9). Similar patterns of gene expression were found in

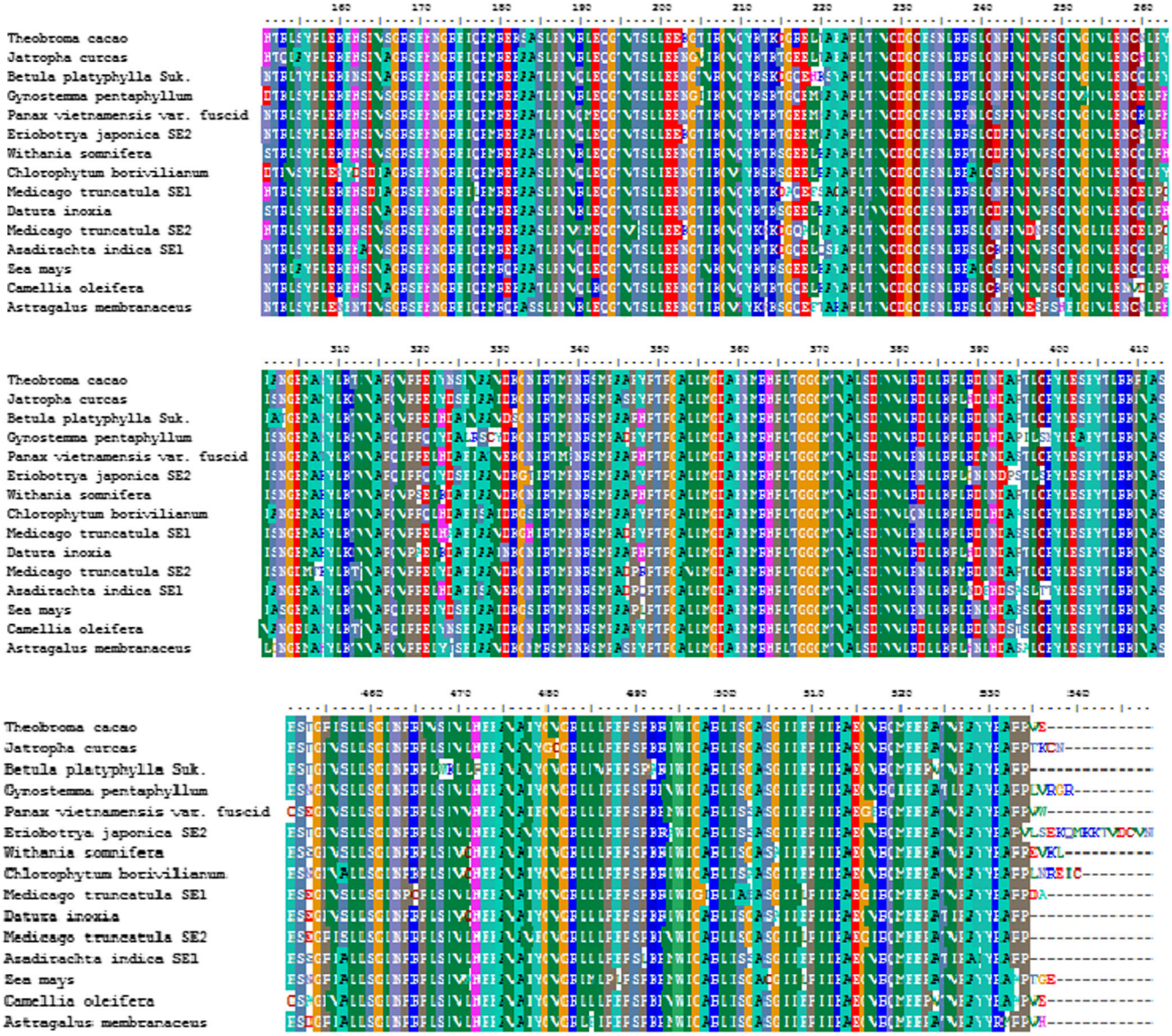

Fig. 5 Multiple alignment of the deduced amino acid sequences of *BpSE* with the known *SE*s of other organisms as determined by the ClustalW multiple alignment tool. The aligned sequences are derived from *Theobroma cacao* (XM_007041985.1), *Jatropha curcas* (AGW81843.1), *Gynostemma pentaphyllum* (ACQ90301.1), *Panax viethamensis var. fuscid* (AIK23028.1), *Eriobotrya japonica* (,AFI33134.2), *Withania somifera* (ADD17678.1), *Chlorophytum borivilianum* (AFN61200.1), *Medicago truncatula* (CAD23249.1), *Datura inoxia* (AAY22200.1), *Asadirachta indica* (AGC82085.1), *Zea mays* (AFW89101.1), *Camellia oleifera* (AGB05603.1), and *Astragalus membranaceus* (ADW27427.1)

Withania somnifera, *P. ginseng*, and *Stevia rebaudiana*. In addition, several studies have reported that the phytohormone elicitor upregulates numerous genes that are involved in secondary metabolism in cultured plant cells. We used ABA (10 μM), ethephon (5 mM), MeJA (100 μM), and SA (50 mg/L) to treat the samples. The expression of *BpSS* was significantly increased by 12 h after ABA treatment as compared to control, whereas *BpSE* was not induced by ABA treatment. The greatest change in expression of *BpSS* transcripts was observed in response to ethephon treatment, which resulted in an 80-fold increase during the 6-h period of post-elicitation. Additionally, *BpSE* transcripts were upregulated following ethephon treatment for 12–24 h. The expression of *BpSS* and *BpSE* was both induced by MeJA at 24 h. However, the SA treatment did not affect the *BpSS* and *BpSE* transcripts notably (Fig. 10).

BpSS localization in the cytoplasm

To examine the subcellular localization of *BpSS* in living cells, we performed transient expression assays using onion epidermal cells with constructs expressing GFP alone and

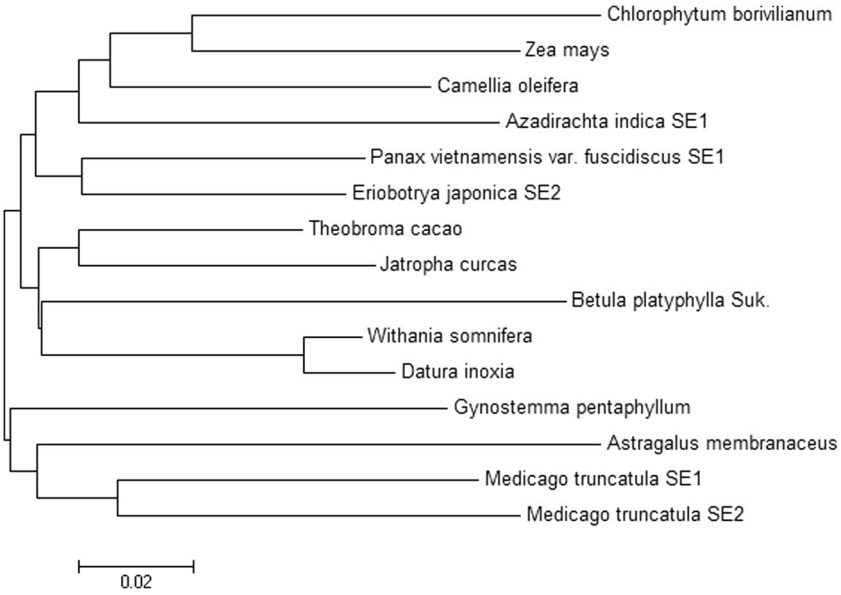

Fig. 6 Neighbor-joining phylogenetic tree constructed from the deduced amino acid sequences of *BpSE* derived from *Betula platyphylla* and *SE*s from other organisms using MEGA 6.0 software. *Numbers above the branches* indicate bootstrap values

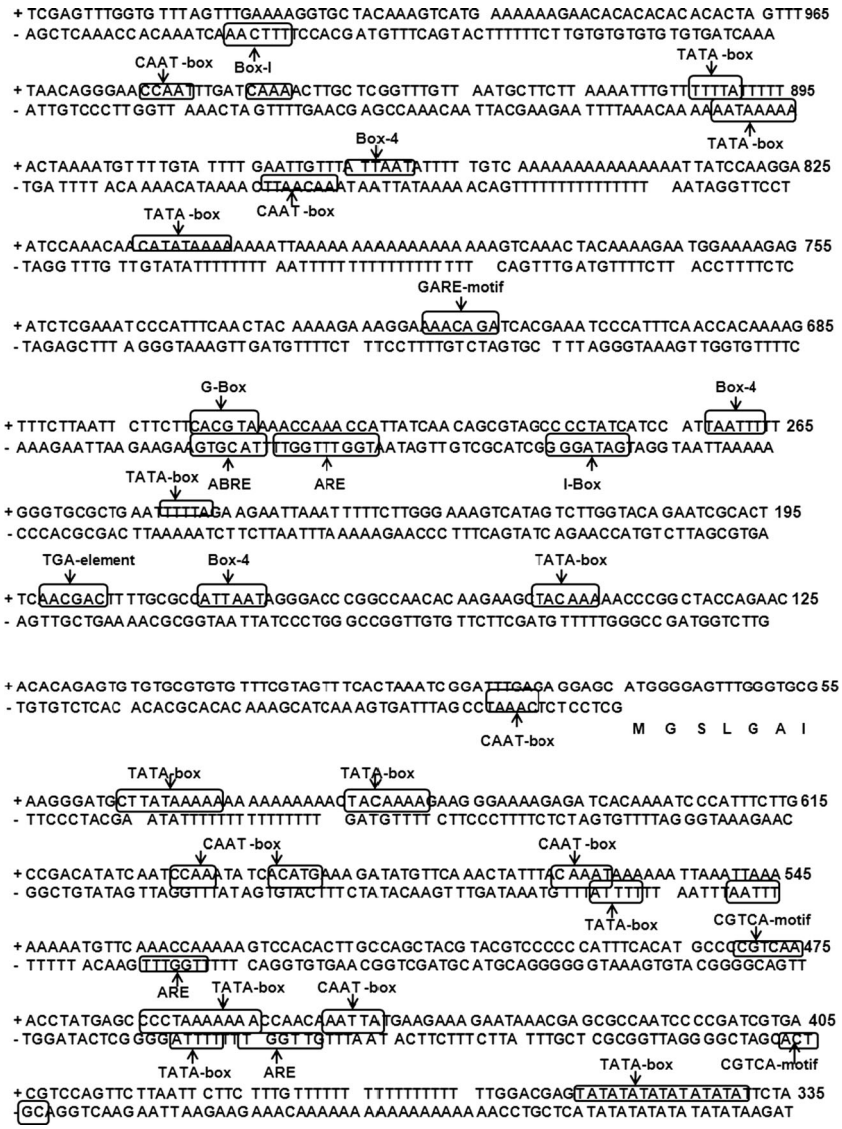

Fig. 7 The DNA sequence of the *Betula platyphylla BpSS* promoter as amplified by the genome walking method. The putative TATA box, CAAT box, and other important *cis*-regulatory elements are marked out

Fig. 8 The DNA sequence of the *Betula platyphylla BpSE* promoter as amplified by the genome walking method. The putative TATA box, CAAT box, and other important *cis*-regulatory elements are marked out

the BpSS/GFP fusion. Fluorescence microscopy showed that the BpSS/GFP fusion protein was targeted into the cytoplasm. By contrast, the control GFP alone was distributed in the entire cell including the cytoplasm and nucleus (Fig. 11). These results indicated that BpSS is localized in the cytoplasm.

Table 2 Putative *cis*-acting regulatory elements identified in the promoter region of *BpSS* using PLANTCARE databases (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)

Cis element	Position	Sequence	Function of site
TATA-box	+889, −722, +693, +686, −573, +381, +262, −236, −121, +69, and 44 positions	TAATA, TATA, ATTATA	Core promoter element around −30 transcription start
CAAT-box	−954, +616, +615, +471, +82, −85, and 13 positions	CAAT	Common *cis*-acting element in promoter and enhancer regions
ABRE	−718	CACGTG	ABA-responsive element
CGTCA-motif	−629, +556	CGTCA	*Cis*-acting regulatory element involved in the MeJA-responsiveness
GARE-motif	+317	AAACAGA	Gibberellin-responsive element
TGA element	+844	AACGAC	Auxin-responsive element
G-box	+718, −718	CACGTG	Light *cis*-acting regulatory element
GATA-motif	−751	GATA	Part of a light responsive element
Box-4	+859, +762, +171	ATTAAT	Conserved DNA sequences involved in the regulation of light
Box-I	−19	TTTCAAA	Light- responsible element
ATC-motif	+614	GCCAATCC	Part of a light responsive element
I-box	−751	GATAA	Light regulatory conserved sequence
Skn-1-motif	+38, +851	GTCAT	Endosperm expression of *cis*-acting element

HPLC of BpSS products accumulated in INVScl1 transformants

Based on the obtained sequences, PCR products of *BpSS* and *BpSE* were digested with corresponding restriction enzymes and ligated to the downstream of *GAL1* promoter of pYES2 to construct the plasmid pYES2-BpSS and pYES2-BpSE. The resulting plasmids were transferred to the INVScl1 strain. The transformants were cultured, and *GAL1* promoter was induced by replacing the carbon source from glucose to galactose. At cultivation time of 12 h, the squalene content of galactose-induced pYES2-BpSS expression yeast cells (0.466 mg/g) was about 13.2 times of the empty vector control yeast cells (0.0338 mg/g) and 3.23 times of the uninduced pYES2-BpSS expression yeast cells (0.144 mg/g) (Figs. 12 and 13). These results indicated that *BpSS* gene indeed involved in the synthesis of squalene and triterpenoids.

Table 3 Putative *cis*-acting regulatory elements identified in the promoter region of *BpSE* using PLANTCARE databases (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)

Cis element	Position	Sequence	Function of site
TATA-box	+940 , +286, +576, −575, and so on	TAATA, TATA, ATTATA	Core promoter element around −30 transcription start
CAAT-box	+751, +496, −863, −833, −862, −532, −385, −615, and 18 positions	CAAT	Common *cis*-acting element in promoter and enhancer regions
5UTR Py-rich repeats	$_{+1167}$	ATTCTCTTC	Efficient transcription element
ARE	+331, +973	TGGTTT	Anaerobic induction responsive element
CGTCA-motif	−311, −739, −568	CGTCA	*Cis*-acting regulatory element involved in the MeJA-responsiveness
GARE-motif	+1017	AAACAGA	Gibberellin-responsive element
AuxRR-core	−801	GGTCCAT	Auxin-responsive element
G-box	+480, −483	CACGTG	Light *cis*-acting regulatory element
GATA-motif	+1029	GATA	Part of a light responsive element
MBS	+44, +946	TAACTG	Drought induced related and MYB binding site
TC-rich repeats	−656, −1162	ATTCTCTAAC	Defense and stress responsive element
GAG-motif	+734	GGAGATG	Part of a light responsive element
GATA-motif	+1029	AAGGATAAGG	Part of a light responsive element
GC-motif	−642	GCCCCGG	Auxin induced enhancer element
CATT-motif	+167, −608	GCATTC	Part of a light responsive element
Sp1	−679, −793, −1045	CC(G/A)CCC	Light responsive element
Skn-1-motif	−310, −696, −738	GTCAT	Endosperm expression of *cis*-acting element

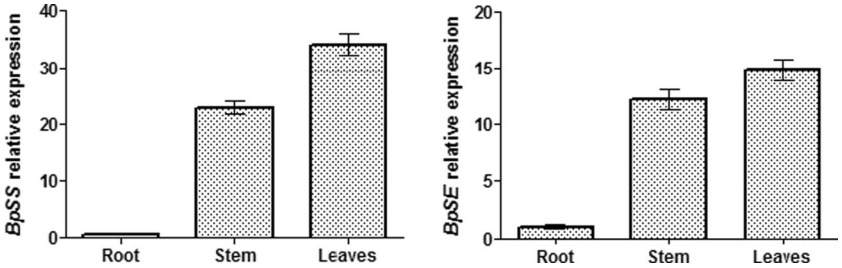

Fig. 9 Tissue-specific expression of *BpSS* and *BpSE* of *Betula platyphyll* seedlings from 4 weeks old. The root, stem, and leaves of 4-week-old birch were collected, and expression of *BpSS* and *BpSE* genes was analyzed using real-time quantitative PCR. Gene expression data were calculated relative to the reference genes Tu after implementation of the $2^{-\Delta\Delta Ct}$ method. A total of three biological and three technical replicates were performed, data are mean±SE of $n=9$

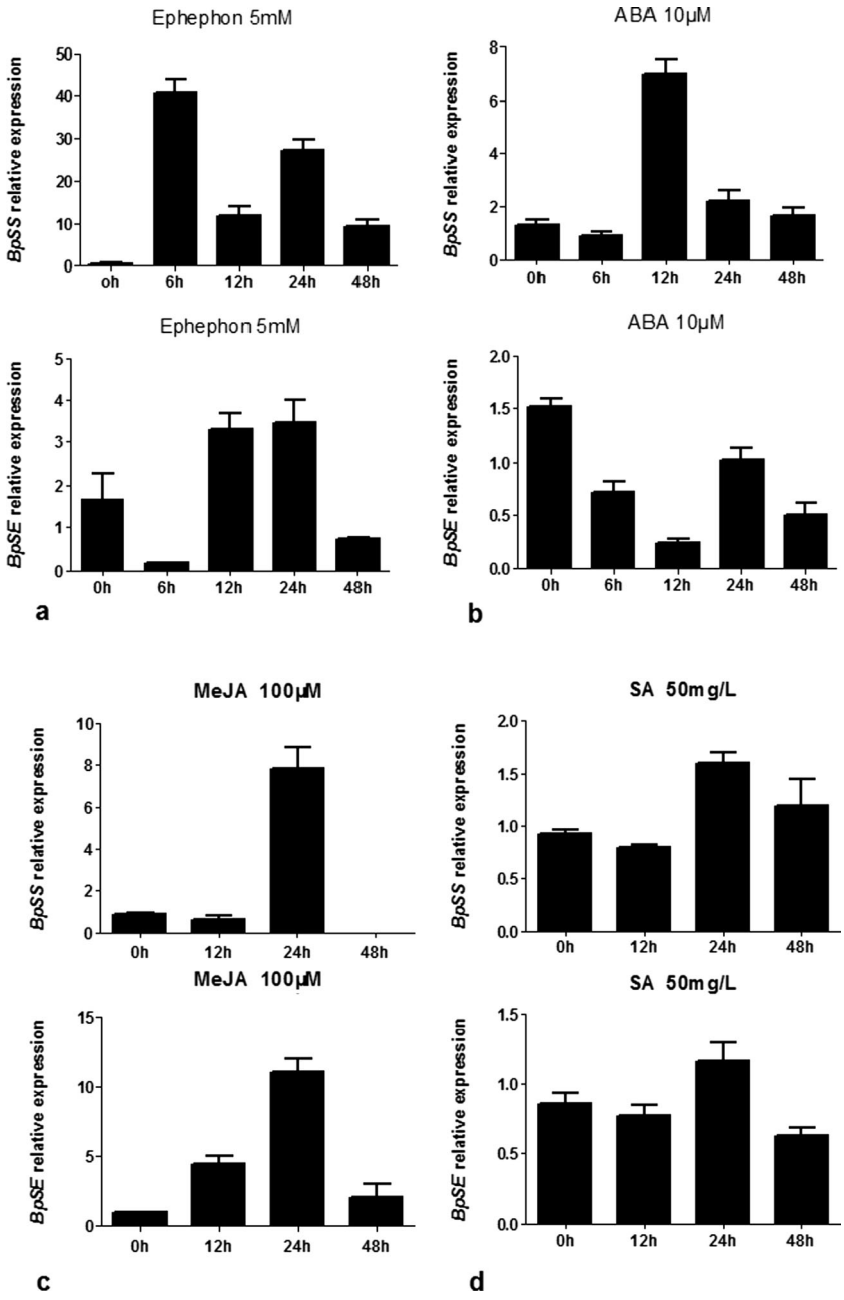

Fig. 10 **a**, **b** Relative quantification of *BpSS* and *BpSE* genes of birch saplings under different hormone treatment. **a** Relative quantification of *BpSS* and *BpSE* genes under 5 mM/L Ephophon treatment. **b** Relative quantification of *BpSS and BpSE* genes under10 μM/L ABA. **c** Relative quantification of *BpSS and BpSE* genes under 100 μM/L MeJA. **d** Relative quantification of *BpSS* and *BpSE* genes under 50 mg/L SA. Birch saplings of 4-week use be different hormone treatment is described in "Materials and Methods." Birch saplings leaf were harvested on 6, 12, 24, and 48 h after different hormone treatment, and expression of *BpSS* and *BpSE* genes was analyzed using real-time quantitative PCR. Gene expression data were calculated relative to the reference genes *Tu* after implementation of the $2^{-\Delta\Delta Ct}$ method. A total of three biological and three technical replicates were performed; data are mean±SE of $n=9$

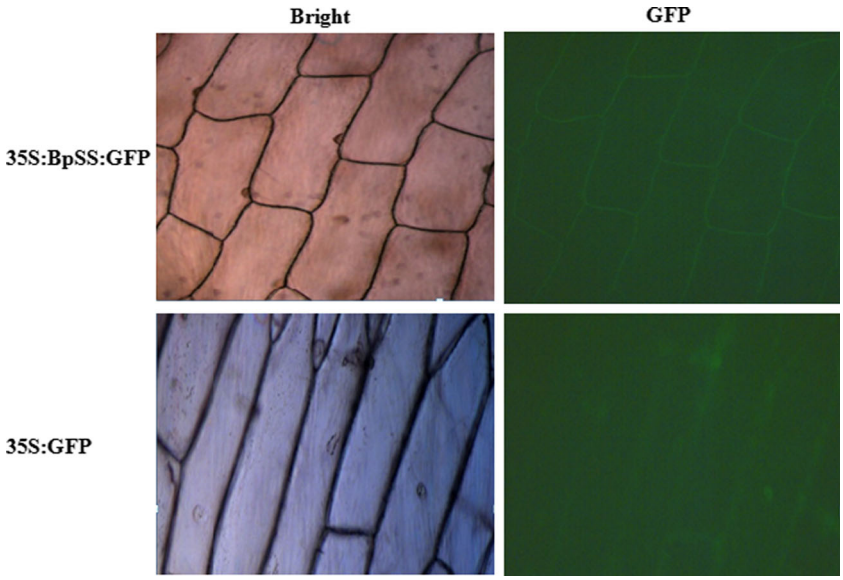

Fig. 11 Subcellular localization of the BpSS/GFP fusion protein onion epidermal cells. The cells with constructs expressing fluorescence protein (GFP) alone and the BpSS/GFP fusion protein were analyzed under bright and fluorescence field after 48 h

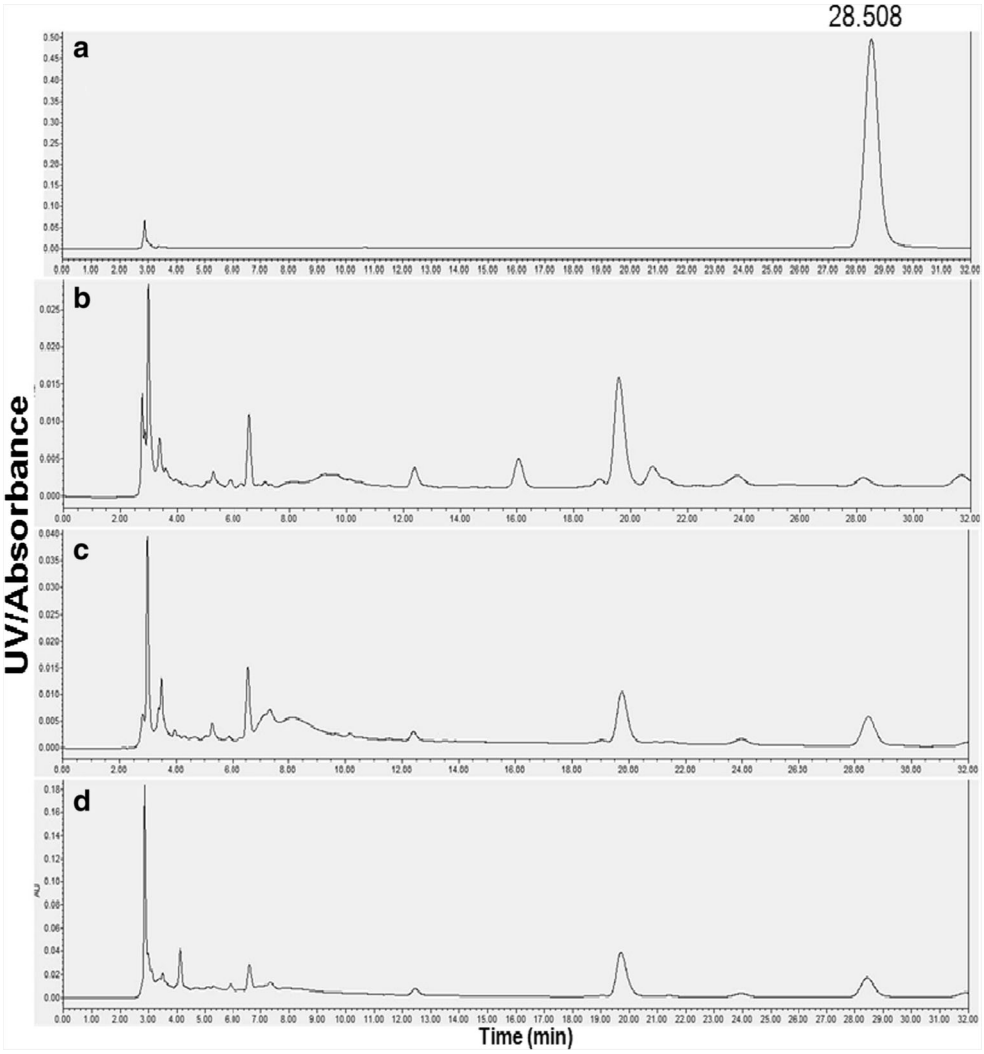

Fig. 12 The HPLC profile of squalene from standard (**a**), the empty vector yeast cells (**b**), the uninduced pYES2-BpSS yeast cells (**c**), and induced pYES2-BpSS yeast cells (**d**). Squalene peak time was at 28.508 min by UV absorption at 210 nm

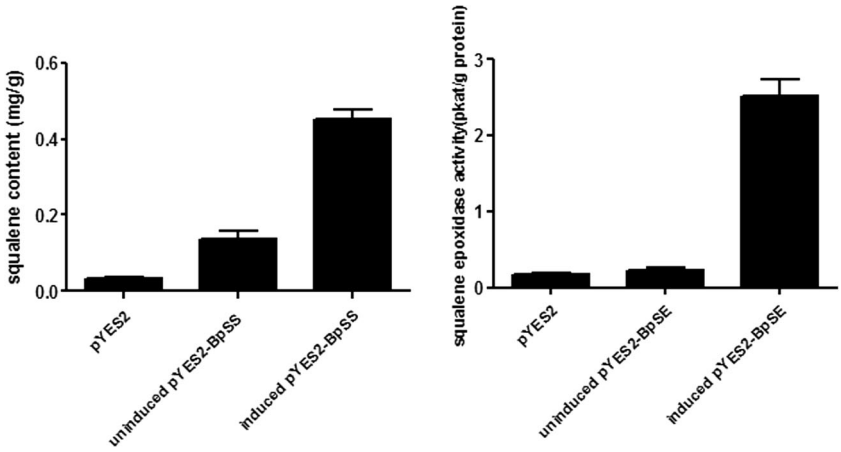

Fig. 13 The content of squalene and squalene epoxidase activity of different yeast cells. The cellular squalene was analyzed according to Lu et al. (2004) with some modifications. The squalene epoxidase activity (pkat/g protein, represented as mean±SD) of each empty vector control yeast cells, uninduced pYES2-BpSE yeast cells, and induced pYES2-BpSE yeast cells. A total of three biological and three technical replicates were performed

Detection of squalene epoxidase activity in INVScl1 transformants

Under the same condition, we examined the absorbance at 400 nm of the reactants and then calculated squalene epoxidase activity. There were no obvious differences between squalene epoxidase activity of uninduced pYES2-BpSE yeast cells and vector empty control yeast cells. In contrast, the galactose-induced pYES2-BpSE yeast cells show more squalene epoxidase activity and had about 11.84 times of control (Fig. 13), and its squalene epoxidase activity was 2.131 pkat/g protein. The squalene epoxidase activity of uninduced pYES2-BpSE yeast cells and vector empty control yeast cells was 0.2314 and 0.1832 pkat/g protein, respectively.

Discussion

Squalene synthase and squalene epoxidase play important roles in terpenes biosynthesis. An enhancement of phytosterol and saponin accumulation by SS and SE gene overexpression has been reported in some species, including *P. ginseng* (Lee et al. 2004) and *Eleutherococcus senticosus* (Seo et al. 2005). In this study, by using degenerate primers and the RACE PCR strategy, full-length cDNA that encoded squalene synthase and squalene epoxidase were isolated from *Betula platyphylla.* In addition, their promoter regions were cloned and characterized, and the full-length and promoter regulatory elements and functions were analyzed using bioinformatics. NCBI BLAST alignment results showed similarity of the *BpSS* cDNA full-length sequence and *Glycine max*, *Theobroma cacao*, *Eleutherococcus senticosus*, and other plants at greater than 82 % homology. The *BpSS*-encoding amino acids contained a typical SQS_PSY conserved domain, and with the Trans_IPPS_HH conserved domains, that might all belong to isoprenoid biosynthesis enzymes, of the class 1 family (Robinson et al. 1993; Huang et al. 2007). In the *BpSE* cDNA full sequences with known *Theobroma cacao* squalene epoxidase SE gene expression and *Jatropha curcas* of the squalene epoxidase gene, it was shown that the SE similarity was 82 % and 80 %, respectively. Conserved domain prediction found that *BpSE* belonged to the SDR superfamily, which contains the squalene epoxidase-specific domain that is comprised of conserved regions, the coenzyme Q binding domain (COQ6) and an FAD-dependent chloride reductase domain (UbiH), as well as other species that are specific to squalene epoxidase that also have the same conserved domains (He et al. 2008; Gomelsky and Klug 2002; Favre and Ryder 1996; Han et al. 2010)

In this study, we isolated *BpSS* and *BpSE* gene promoters in the *Betula platyphylla* genome, with fragment lengths of 965 and 1193 bp. The promoter software program plantCARE displayed results as displayed in Figs. 7 and 8. The *BpSS* and *BpSE* promoter contains a variety of other plants in the presence of promoters such as the common promoter elements including the TATA and CAAT box. The *BpSE* promoter contains the stress-related *cis* element and the MYB binding site, which collectively will empower the *BpSE* gene with the ability of adapting to the environment. In addition to the promoters of both genes that contained *cis*-acting elements that were associated with the induction of plant hormones, the ABRE abscisic acid response element, CGTCA-motif methyl jasmonate-responsive element, the GARE-motif GA-responsive element, the TGA-motif auxin response elements, and the AuxRR-core core auxin response elements collectively indicated that both *BpSS* and *BpSE* could be induced by the plant hormones. By analyzing the expression patterns of *BpSS* and *BpSE*, the results are the same as our prediction.

Jasmonate (JA) and its methyl ester (MeJA) are generally considered to modulate many physiological events in higher plants, such as defense responses, flowering, and senescence (Wasternack and Parthier 1997). Treatments of these signaling molecules are one of the major biotechnological approaches to modify the biosynthesis of secondary metabolites. It was found that following treatment for 2 h in the *Betula platyphylla* study, gene expression of both *BpSS* and *BpSE*

was significantly upregulated, which indicates that it plays a role as a signaling molecule in the biosynthesis of triterpenes in *Betula platyphylla*. MeJA could also augment the accumulation of *SS* mRNA in *Bupleurum tenue* (Kim et al. 2011b). When MeJA was administered to cultured roots of *Bupleurum falcatum*, the level of BfSS1 mRNA greatly increased in a dose-dependent manner (Young et al. 2011). Moreover, others have found that *Withania somnifera* of the gene promoters of *SS and SE* displayed multiple plant hormones that were associated *cis*-acting elements, and after MeJA treatment for 12 h, the expression of both genes was increased (Bhat et al. 2012; Razdan et al. 2003). A few studies have reported the effects of SA on the expression of genes involved in triterpene biosynthesis. In the present experiment, the accumulation of BpSE and BpSS mRNA was not affected by SA treatment. Aoyagi et al. (2001) found that SA treatment at various concentrations did not promote saikosaponin production in *Bupleurum falcatum*. It was reported that exogenous application of SA or its derivatives blocks not only JA biosynthesis but also the action of JA after wounding in tomato (Pena-Cortes et al. 1993; Doares et al. 1995), suggesting a possible mechanism for the inverse relationship between SA and JA in triterpene metabolism. Other plants, such as *P. ginseng* and *Panax notoginseng*, informed us that other hormones could enhance the expression of both genes (Niu et al. 2014). Interestingly, the accumulation of BpSS mRNA increased after treatment with ABA and ethephon, which have the same result with BfSS1 treatment in *Bupleurum falcatum.* These results can indicate that these stress hormones can promote the expression of SS and SE, and this process might be responsible for the biosynthesis of secondary metabolites. Although ABA is not a regular signal to induce secondary metabolism, it has also been proved that can regulated secondary metabolism of some kind of plants. Ethylene has been reported that can increase the production of anthocyanin, flavonoid, and stilbenoid by upregulating their biosynthetic genes in grape (Vitisvinifera) cell cultures (El-Kereamy et al. 2003).

An important molecular event initiates secondary metabolite biosynthesis wherein the transcription factor binds the gene promoter and then activates triterpene biosynthesis-related gene expression. Studies have found that a variety of terpene biosynthetic transcription factors can specifically bind with the G-Box start codon for terpene synthesis-related gene expression, such as Siberil (Pré et al. 2000). This uses the STR gene promoter in the G-Box as an inducer from the periwinkle-isolated cDNA library, which was identified as the bZIP family CrGBF (*Catharanthus roseus* G-box binding factor) transcription factor: CrGBF1 and CrGBF2. Both the *BpSS* and *BpSE* promoter sequences exist in the G-Box element *Betula platyphylla*, which may have sequence homology to bZIP. Moreover, the bHLH transcription factor can bind G-Box to start *BpSS* and *BpSE* mRNA transcription and then regulate triterpenoid metabolism. The MYB transcription factor plays a significant role in plant metabolism and development, and its main function is to regulate secondary metabolism and control cell morphogenesis (Dubos et al. 2008). Many studies have shown that MYB transcription factors are involved in plant secondary metabolism (Shen et al. 2012). In the *BpSE* promoter sequences, MBS is the MYB-binding site that can specifically bind an MYB transcription factor and then start *BpSE* mRNA transcription. We speculate that in the *BpSS* and *BpSE* promoter regions of G-Box, some other *cis*-acting elements may also exist, which represents the sole transcription factor that can bind multiple triterpenoid biosynthetic genes in the regulation of triterpene metabolism.

SS protein has been demonstrated in some other species that is membrane bound and localized in endoplasmic reticulum. But, it has not confirmed in *Betula platyphylla* yet. Most membrane proteins have hydrophobic regions which span the hydrophobic core of the membrane bilayer. C-terminal hydrophobic sequences of SS protein are important for anchoring the enzyme to the endoplasmic erticulum membrane (Robinson et al. 1993; Kribii et al. 1997). Membrane-spinning amino acids derived from *BpSS* were predicted using TMHMM, a program that predicts transmembrane helices based on a hidden Markov model. BpSS not only had C-terminal hydrophobic sequences but also had additional membrane-spanning helix near the region of domain C (the 389~406 amino acid chain is outside-transmembrane helix). We performed transient expression assays using onion epidermal cells with constructs expressing GFP alone and the BpSS/GFP fusion, which also indicated the BpSS protein located in the cytoplasm. But, if the BpSS protein is membrane bound and localized in endoplasmic reticulum or not will require further experimental analysis. SE protein is revealed by a perusal of literature that is a microsomal protein and also present in lipid droplets, but only ER associated protein has been found to be active (Leber et al. 1998). We will study the cellular localization of BpSE protein in the future.

HPLC analysis of *BpSS* products accumulated and detection of squalene epoxidase activity in INVScl1 transformants indicated that the *BpSS* and *BpSE* were indeed involved in the synthesis of triterpenoids. The data confirm the function of two genes.

In this study, we cloned and detected the expression patterns of the genes encoding the enzymes *BpSS* and *BpSE* that are involved in birch triterpene biosynthesis. In addition, we isolated *BpSS* and *BpSE* promoters and analyzed *cis*-acting elements. In the future, we plan to study the relationship between *BpSS* and *BpSE* expression, secondary metabolite synthesis, and transcription factors. Therefore, the regulation of the gene expression of key enzyme and their promoters that are involved in triterpene biosynthesis not only enables investigators to improve understanding of the pathways involved but might also have important research and development value.

Acknowledgments This work was supported by the Fundamental Research Funds for the Central Universities (257-2015DY01), Natural Science Foundation of China (31200428, 31570589), and National basic science training fund (J1210053).

Compliance with ethical standards

Conflict of interest We have no conflicts of interest to declare.

References

Alakurtti S, Mäkelä T, Koskimies S, Yli-Kauhaluoma J (2006) Pharmacological properties of the ubiquitous natural product betulin. Eur J Pharm Sci 29:1–13

Aoyagi H, Kobayashi Y, Yamada K, Yokoyama M, Kusakari K, Tanaka H (2001) Efficient production of saikosaponins in *Bupleurum falcatum* root fragments combined with signal transducers. Appl Microbiol Biotechnol 57:482–488

Bhat WW, Lattoo SK, Razdan S et al (2012) Molecular cloning, bacterial expression and promoter analysis of squalene synthase from *Withania somnifera* (L.) Dunal[J]. Gene 499(1):25–36

Busquets A, Keim V, Closa M et al (2008) Arabidopsis thaliana contains a single gene encoding squalene synthase. Plant Mol Biol 67(1-2):25–36

Dang Z, Qian K, Ho P et al (2012) Synthesis of betulinic acid derivatives as entry inhibitors against HIV- 1 and bevirimat- resistant HIV-1 variants. Bioorg Med Chem Lett 22(16):5190–5194

Doares SH, Narvaez-Vasquez J, Conconi A, Ryan CA (1995) Salicylic acid inhibits synthesis of proteinase inhibitors in tomato leaves induced by systemin and jasmonic acid. Plant Physiol 108:1741–1746

Dubos C, Le Gourrierec J, Baudry A et al (2008) MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. Plant J 55(6):940–953

El-Kereamy A, Chervin C, Roustan JP, Cheyhnier V, Souquet JM, Moutounet M, Raynal J, Ford C, Latche´ A, Pech JC, Bouzayen M (2003) Exogenous ethylene stimulates the long term expression of genes related to anthocyanin biosynthesis in grape berries. Physiol Plant 119:175–282

Favre B, Ryder NS (1996) Characterization of squalene epoxidase activity from the dermatophyte *Trichophyton rubrum* and its inhibitor by terbinafine and other antimycotic agents. Antimicrob Agents Chemother 40(2):443

Gomelsky M, Klug G (2002) A novel FAD-binding domain involved in sensory transduction in microorganisms. Trends Biochem Sci 27(10):497

Han JY, In JG, Kwon YS et al (2010) Regulation of ginsenoside and phytosterol biosynthesis by RNA interferences of squalene epoxidase gene in *Panax ginseng*. Phytochemistry 71(1):36–46

Haralampidis K, Trojanowska M, Osbourn AE (2002) Biosynthesis of triterpenoid saponins in plants[M]//History and Trends in Bioprocessing and Biotransformation. Springer, Berlin Heidelberg, pp 31–49

Huang ZS, Jiang KJ et al (2007) Molecular cloning and characterization of Yew gene encoding squalene synthase from *Taxus cupidate*. Biochem Mole Biol 40(5):625–635

He F, Zhu Y, He M et al (2008) Molecular cloning and characterization of the gene encoding squalene epoxidase in *Panax notoginseng*. DNA Seq 19(3):270

Kim TD, Han JY, Huh GH et al (2011a) Expression and functional characterization of three squalene synthase genes associated with saponin biosynthesis in *Panax ginseng*. Plant Cell Physiol 52(1):125–137

Kim YS, Cho JH, Park S (2011b) Gene regulation patterns in triterpene biosynthetic pathway driven by overexpression of squalene synthase and methyl jasmonate elicitation in *Bupleurum falcatum*. Planta 233:343–355

Kribii R, Arro M, Del Arco A, Gonza´lez V, Balcells L, Delourme D, Ferrer A, Karst F, Boronat A (1997) 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 249:61–69

Laule O, Furholz A, Chang HS, Zhu T, Wang X, Heifetz PB, Gruissem W, Lange BM (2003) Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in Arabidopsis thaliana. Proc Natl Acad Sci USA 100:6866–6871

Lee MH, Jeong JH, Seo JW et al (2004) Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* overexpressing squalene synthase gene. Plant Cell Physiol 45(8):976–984

Leber R, Landl K, Zinser E, Ahorn H, Spok A, Kohlwein SD, Turnowsky F, Daum G (1998) Dual localization of squalene epoxidase, Erg1p, in yeast reflects a relationship between the endoplasmic reticulum and lipid particles. Mol Biol Cell 9(2):375–386

Livak K.J, Schmittgen T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25(4):402–408

Lu H, Liu J, Yang W et al (2007) Isolation of Squalene Synthase Genes of *Glycyrrhiza Uralensis* and Construction of Plant Expression Vector. Pharm Biotechnol Beijing 14(4):255

Lu HT, Jiang Y, Chen F (2004) Determination of squalene using high-performance liquid chromatography with diode array detection. Chromatographia 59:367–71

Nagegowda DA (2010) Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. FEBS Lett 584:9

Niu Y, Luo H, Sun C et al (2014) Expression profiling of the triterpene saponin biosynthesis genes FPS, SS, SE, and DS in the medicinal plant Panax notoginseng. Gene 533(1):295–303

Patočka J (2003) Biologically active pentacyclic triterpenes and their current medicine signification. J Appl Biomed 1(1):7–12

Pré M, Sibéril Y, Memelink J et al (2000) Isolation by the yeast one-hybrid system of cDNAs encoding transcription factors that bind to the G-box element of the strictosidine synthase gene promoter from *Catharanthus roseus*. Int J Biochromatogr 5:229–244

Pena-Cortes H, Albrecht T, Prat S, Weiler EW, Willmitzer L (1993) Aspirin prevent wound-induced gene expression in tomato leaves by blocking jasmonic acid biosynthesis. Planta 191:123–128

Razdan S, Bhat WW, Rana S et al (2003) Molecular characterization and promoter analysis of squalene epoxidase gene from *Withania somnifera* (L.) Dunal. Mol Biol Rep 40(2):905–916

Robinson GW, Tsay YH, Kienzle BK et al (1993) Conservation between human and fungal squalene synthetases: similarities in structure, function, and regulation. Mol Cell Biol 13(5):2706–2717

Seo JW, Jeong JH, Shin CG et al (2005) Overexpression of squalene synthase in *Eleutherococcus senticosus* increases phytosterol and triterpene accumulation. Phytochemistry 66(8):869–877

Shen H, He X, Poovaiah CR et al (2012) Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. New Phytol 193(1):121–136

Sun I, Chen C, Kashiwada Y et al (2002) Anti-AIDS agentssynthesis, anti-HIV, and anti-fusion activities of IC9564 analogues based on betulinic acid. J Med Chem 45(19):4271–4275

Wasternack C, Parthier B (1997) Jasmonate-signalled plant gene expression. Trends Plant Sci 2:302–307

Yin J, Ren CL, Zhan YG, Li CX, Xiao JL, Qiu W, Li XY, Peng HM (2012) Distribution and expression characteristics of triterpenoids and OSC genes in white birch (*Betula platyphyllasuk.*). Mol Biol Rep 39(3):2321–2328

Young SK, Jung HC, Sangkyu P et al (2011) Gene regulation patterns in triterpene biosynthetic pathway driven by overexpression of squalene synthase and methyl jasmonate elicitation in *Bupleurum falcatum*. Planta 233:343–355

Zhang H, Shibuya M, Yokota S et al (2003) Oxidosqualene cyclases from cell suspension cultures of *Betula platyphylla var.* japonica: molecular evolution of oxidosqualene cyclases in higher plants. Biol Pharm Bull 26(5):642–650

Zuco V, Supino R, Righetti SC et al (2002) Selective cytotoxicity of betulinic acid on tumor cell lines, but not on normal cells. Cancer Lett 175(1):17–25