Journal of Plant Biology, September 2005, 48(3) : 263-269

Isolation and Characterization of Squalene Synthase cDNA from *Centella asiatica* (L.) Urban

Ok-Tae Kim[1], Nak-Sul Seong[1], Min-Young Kim[2], and Baik Hwang[2]*

[1]National Institute of Crop Science, RDA, Suwon 441-857, Korea
[2]Department of Biology and Institute of Plant Resources, Chonnam National University, Gwangju 500-757, Korea

We have cloned and characterized a gene for squalene synthase (SQS) from *Centella asiatica* (L.) Urban, a species that produces a large quantity of triterpene saponins such as asiaticoside and madecassoside. Its full-length cDNA clone was isolated by RACE PCR. The sequence of pSQS contains an open reading frame of 1248 nucleotides, which code for 416 amino acids with a molecular mass of 47.3 kDa. Southern analysis revealed that one copy might exist in the *C. asiatica* genome. We also determined that 0.1 mM methyl jasmonate was sufficient to up-regulate those levels of *CaSQS* mRNA.

Keywords: asiaticoside, *Centella asiatica*, squalene synthase, triterpene saponins

Centella asiatica (L.) Urban, a medicinal plant in the Umbelliferae family, grows in some tropical regions of Southeast Asia, Australia, and Southern and Central Africa (Brinkhaus et al., 2000). Its leaves contain large amounts of triterpene saponins, e.g., asiaticoside, which show activity against fungi and pathogens. The regulation of triterpenoid biosynthesis has been studied in plants, because these compounds have commercial value as pharmaceuticals, sweeteners, taste modifiers, and cosmetics (Hostettmann and Marston, 1995). These triterpene saponins can be produced by various methods of plant tissue culture (Ahn et al., 1993; Lian et al., 2002; Kim et al., 2004a).

Two molecules of farnesyl diphosphate are converted by squalene synthase (SQS) into squalene, as a common precursor of sterol and triterpenoid biosynthesis (Fig. 1; Abe et al., 1993). Epoxidation of squalene at the second and third carbon positions results in the formation of 2,3-oxidosqualene. Oxidosqualene cyclases, which are situated at the branching step for biosynthesis of phytosterol and triterpene saponins, catalyze the cyclization of 2,3-oxidosqualene. SQS cDNA has been isolated from several species, including *Arabidopsis thaliana* (Nakashima et al., 1995), *Nicotiana tabacum* (Devarenne et al., 1998), *Capsicum annuum* (Lee et al., 2002), and *Lotus japonicus* (Akamine et al., 2003). In higher plants, SQS plays an important role in the regulation of various triterpenoids. Therefore, molecular cloning of SQS provides a useful tool for studying how saponin biosynthesis is controlled. For example, the application of exogenous methyl jasmonate (MJ) results in the enhancement of saponin content in plant tissues; up-regulation of the SQS gene is observed during its elicitation (Suzuki et al., 2002; Hayashi et al., 2003). However, although a number of genes related to triterpene saponins have been cloned and characterized, no direct evidence has been found for the regulatory function of *SQS* in saponin biosynthesis. Here, we report the isolation and characterization of a cDNA that encodes squalene synthase from *C. asiatica* as a first step in understanding this pathway.

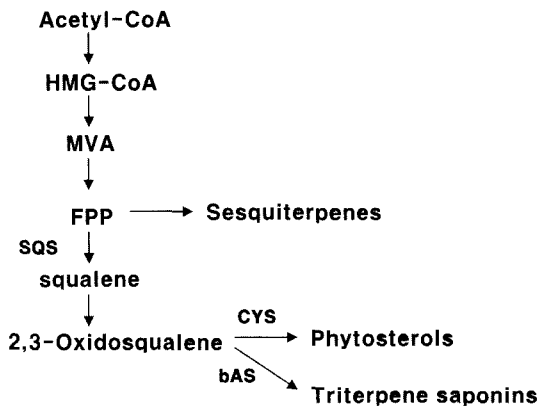

Figure 1. Triterpenoid pathways. HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A; MVA, mevalonic acid; FPP, farnesyl diphosphate; SQS, squalene synthase; CYS, cycloartenol synthase; bAS, β-amyrin synthase.

*Corresponding author; fax +82-62-530-3409
e-mail bhwang@chonnam.ac.kr

MATERIALS AND METHODS

Whole-Plant Cultures

Seeds of *C. asiatica* from Jeju Island, Korea, were sterilized for 10 min in a 3% sodium hypochlorite solution containing 0.1% Tween 20. They were then rinsed twice with sterile distilled water. Whole-plant cultures were maintained in the light at 25°C in 250-ml Erlenmeyer flasks containing 50 mL of liquid B5 medium (Gamborg et al., 1968) supplemented with 3% sucrose. Sub-culturing occurred at six-week intervals (Kim et al., 2004a). For our experiments with MJ elicitation, whole plants were cultured and elicited as described by Kim et al. (2004b). Only leaves of whole plants were harvested after 3 weeks of nodal inoculation for total RNA extractions.

Total RNA Extraction and cDNA Synthesis

Total RNA was extracted from *Centella* leaves, using Trizol reagent and according to Invitrogen procedures. Afterward, mRNA was purified by a Poly(A)$^+$ RNA purification kit (Qiagen, USA), and was reverse-transcribed with dNTP at 37°C using AMV reverse transcriptase (Invitrogen, USA) and an oligo(dT) primer (RACE 32, 5′-GACTCGAGTCGACATCGATTTTTTTTTT-TTTT-3′), as described by Frohman et al. (1988). The resulting cDNA mixture was diluted with 200 μL H_2O and directly used as a template for PCR.

Cloning of Squalene Synthase

Degenerate primers were designed from the highly conserved regions reported by Robinson et al. (1993). Their nucleotide sequences were: coreS = 5′-TAYTGYCAYTAYGTNGC-3′ (YCHYVA) and coreA = 5′-ACYTGNGGDATNGCRCARAA-3′ (FCAIPQV). PCR was carried out for 30 cycles on the GeneAmp® PCR System 2700 (Applied Biosystems, USA) with a program of 94°C, 1 min; 42°C, 2 min; 72°C, 3 min; and a final extension at 72°C, 10 min. A second PCR was performed with the same primers, using 5 μL of the first PCR product as template and under the same conditions as the first PCR. This produced a 371-bp fragment. The PCR products were then cloned into the pGEM-T Easy vector (Promega, USA) and transformed into *Escherichia coli* strain DH5α competent cells. Plasmid DNA was prepared from 15 individual transformants and sequenced with the Automatic Genetic Analyzer 3100 (Applied Biosystems).

Rapid Amplification of cDNA Ends (RACE) PCR

Both 5′ rapid amplification of cDNA ends (RACE) and 3′RACE PCR were performed with gene-specific primers designed against partial sequences obtained through RT-PCR. The primer sequences included: for SQS, 5′RACE 5′-TGAAGCCAAATCTTCAGCCCC-3′ and nested primer 5′-GAAGAGCTTTGACAACCCTAA-3′; and 3′RACE 5′-AGGCAGTGCAGTGCCTCAACG-3 and nested primer 5′-TCTGATTTGCGAGATCCTGCT-3′. PCR was carried out for 30 cycles with a program of 94°C, 1 min; 58°C, 1 min; 72°C, 3 min; and a final extension at 72°C, 10 min. The first products were amplified with the RACE32 primer; filtered products served as template for a second PCR that used a nested primer and an anchor primer (5′-GACTC-GAGTCGACATCGA-3′). After obtaining a full-length cDNA clone, we sequenced *CaSQS* for submission to EMBL, GenBank, and the DDBJ sequence databases.

Construction of CaSQS Expression Vectors and Expression in *E. coli*

To make the recombinant enzymes soluble and to avoid problems regarding their expression, we performed PCR to remove the C-terminal 28 amino acids that constituted a hydrophobic transmembrane segment, using a forward primer (5′-AGTGGATC-**CATG**GGAAGTTTAGGGGCG-3′) (*Bam*HI restriction site underlined and the translation start codon in boldface) as well as a reverse primer (5′-TAA-AAGCTTTCCCGAGTCATTCTCAAG-3′) that harbored a *Hin*dIII site (underlined). The resulting 1161-bp PCR product was digested with *Bam*HI and *Hin*dIII, then gel-purified and ligated into *Bam*HI and *Hin*dIII within the pET32a expression vector, according to the manufacturer's protocol (Novagen, USA). Compared with the wild-type CaSQS, the recombinant fusion protein had 159 amino acids for Trx/His/S-taq at the N terminus. Moreover, a stop codon was eliminated by PCR amplification to fuse the proteins of His-taq (LEHHHHHH) at the C terminus. The constructed pET32a-*CaSQS* was transformed into *E. coli* strain BL21 (DE3) pLysS. CaSQS was expressed by adding isopropylthio-α-D-galactoside (IPTG), at a final concentration of 1 mM, to the exponentially growing recombinant cells (OD_{600} = ~0.5) harboring the pET32a-CaSQS plasmid. This expression was detected by SDS-PAGE by periodically harvesting 100 μL aliquots of the bacterial culture and centrifuging them for 2 min at 5000g. The pellets were re-suspended in 100 μL of 50 mM Tris-HCl (pH 6.8), 10 mM DTT, 2%

SDS, 0.01% bromophenol blue, and 10% glycerol. Afterward, 10 µL aliquots were separated on a 10% SDS-polyacrylamide gel and stained with Coomassie blue for protein detection.

Southern Blot Analysis

Genomic DNA was isolated from *Centella* leaves according to the CTAB procedure (Doyle and Doyle, 1987). Our Southern blotting involved digesting 10 µg of leaf genomic DNA with *Eco*RI, *Hin*dIII, *Pst*I, or *Xba*I, then electrophoresing the digested products on a 0.8% agarose gel (25 V, overnight). This was followed by hybridization with probe cDNA that had been radio-labelled with α-[^{32}P]-dCTP by a random primer kit (Roche, USA). The blot was first washed twice with 5% SDS (containing 1 mM EDTA and 40 mM Na_2HPO_4) for 15 min each at 65°C, then twice (15 min each) with 1% SDS (1 mM EDTA and 40 mM Na_2HPO_4) at 65°C.

Northern Blot Analysis

Small-scale total RNA was isolated from leaf tissue, using Trizol reagent according to Invitrogen procedures. Aliquots of the RNA preparations (30 µg per lane) were fractionated by electrophoresis on 1.5% agarose gels containing formaldehyde, and the fragments were then blotted onto a nylon membrane (Zeta-Probe, Bio-Rad, USA). After being washed for 15 min at room temperature (RT) with 2× SSC/0.1% SDS and for 15 min at RT with 0.5× SSC/0.1% SDS, the blots were washed for 15 min at 43°C with 0.1× SSC/0.1% SDS, and then autoradiographed with an intensifying screen at -80°C for 5 d.

RESULTS AND DISCUSSION

cDNA Cloning of *C. asiatica* SQS

We used homology-based PCR to isolate SQS cDNA from the leaves of C. *asiatica*. Degenerate oligonucleotide PCR primers were designed from the consensus sequence reported by Robinson et al. (1993). The forward primer was based on a consensus sequence in domain (YCHYVA); the reverse primer, from domain (FCAIPQV). Repeat PCRs successfully amplified the 371-bp products of the core fragment for a putative squalene synthase cDNA. To obtain the full-length sequence, 3'- and 5'-RACE PCRs were conducted with oligonucleotide primers specific for the core sequences of SQS. The full-length cDNA, achieved with nested primers, comprises 1573 nucleotides with a 1248-b ORF, flanked by a 103-b 5' untranslated region and a 222-b 3' untranslated region that included a poly(A) tail of 16 b (GenBank accession No. AY787628). This clone encodes a peptide of 416 amino acids with a predicted molecular mass of 47.3 kDa, which is similar to reported values from other plant species (Hayashi et al., 1999; Lee et al., 2004). CaSQS exhibits 86, 77, 47, and 39% amino acid sequence identities with *Panax ginseng* (AB115496), *A. thaliana* (D29017), *Mus musculus* (D29016), and *Saccharomyces cerevisiae* (M63979), respectively.

The conserved segments of known squalene synthases (A, B, and C in Fig. 2) are believed to represent crucial regions for the active site (McKenzie et al., 1992; Robinson et al., 1993). The amino acid residues in those segments are also conserved in the *Centella* polypeptide. Akamine et al. (2003) have shown that a mutation of the Phe residue at position 285 in the C segment dramatically reduces SQS activity. These experiments demonstrated that the phenylalanine residue in the conserved C segment plays an important role in the formation of squalene. Amino acids in the C-terminal region (D segment) exhibited the lowest sequence identity among all SQS proteins. Because this region is very hydrophobic in all SQS enzymes, including CaSQS, it may function as an anchor in the endoplasmic reticulum membrane. Consistent with this suggestion, Lee et al. (2002) have reported that the carboxyl-terminal deletion of residues 389-411 in the SQS enzyme results in the accumulation of a functionally soluble SQS protein.

Using Clustal W to compare the deduced amino acid sequences for CaSQS with those from other species, we created a phylogenetic tree of squalene synthases (Fig. 3). The mammalian enzymes formed a subgroup, being closely related to each other. In contrast, a separate subgroup was established for the plant enzymes, with CaSQS being closely associated with *P. ginseng* SQS.

Southern Analysis

To determine the number of genes for *CaSQS* in the C. *asiatica* genome, we performed a Southern blot analysis. When probed under high-stringency conditions, two bands were observed from genomic DNA digested by *Eco*RI; a single band was detected after digestion with *Hin*dIII, *Pst*I, or *Xba*I (Fig. 4). This suggests that C. *asiatica* contains only one gene for

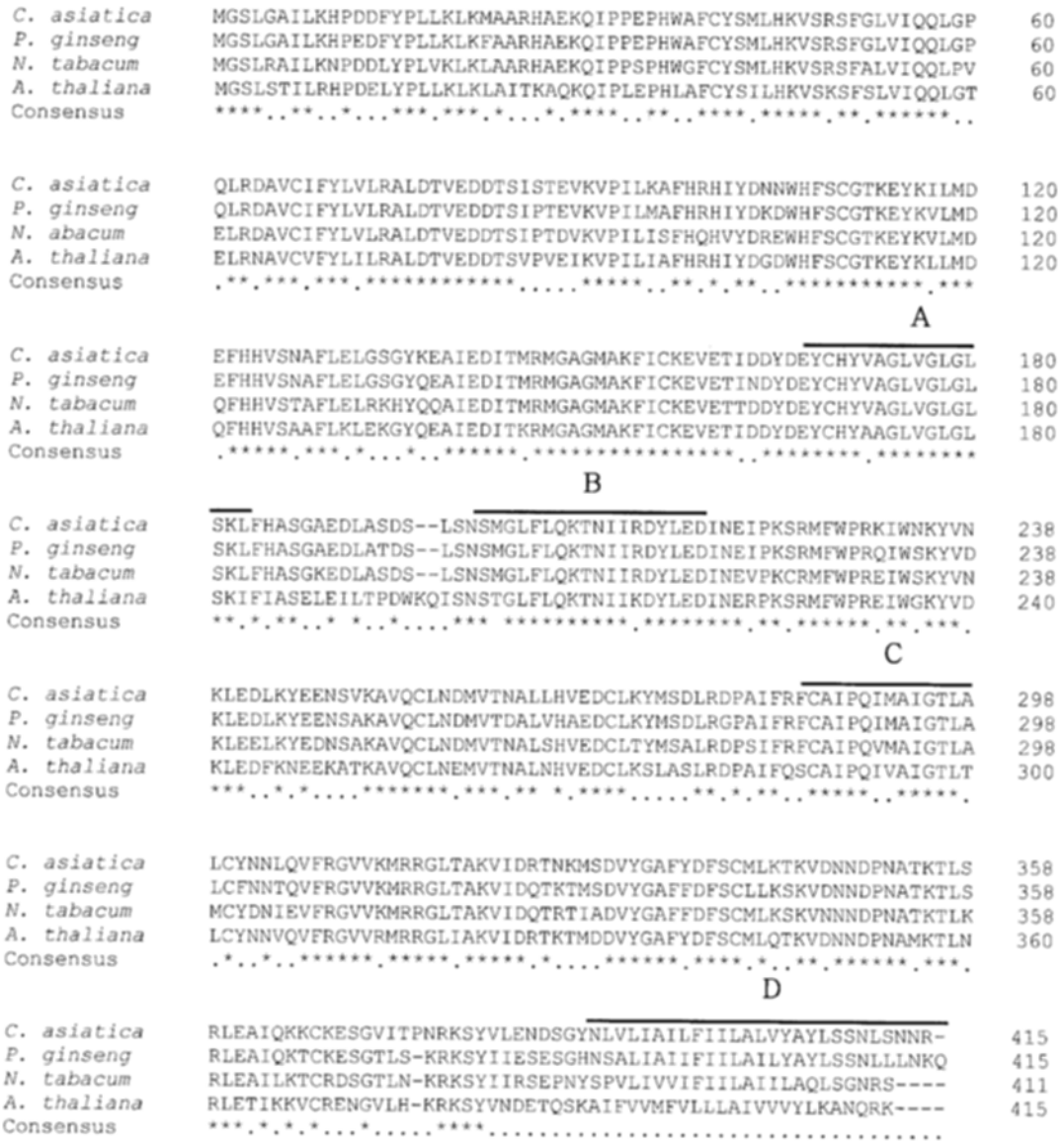

Figure 2. Alignment of the amino-acid sequence of *C. asiatica* SQS with those of proteins encoded by *P. ginseng* (AB115496), *N. tabacum* (U60057), and *A. thaliana* (AF004396). Conserved segments A, B, and C of known squalene synthases are overlined; D segment is very hydrophobic. Asterisks indicate amino acid residues that are identical in all four proteins.

CaSQS. Likewise, the squalene synthase genes in yeast and humans also are reported to be single copies (Robinson et al., 1993). However, at least two SQS genes per genome have been reported in plant species such as *A. thaliana* (Kribii et al., 1997) and *N. tabacum* (Devarenne et al., 1998).

CaSQS Expression in *E. coli*

The full-length cDNA and the truncated cDNA of *CaSQS* consist of a deletion that corresponds to the last 28 amino acids of the carboxyl terminus to make the recombinant enzyme soluble. Both components were inserted into vector pET32a and transformed into *E. coli* strain BL21 (DE3) pLysS. Expression of the truncated CaSQS, which was induced by the addition of 1 mM IPTG, was detected at the predicted molecular mass (61.2 kDa) after 3 h of IPTG induction (Fig. 5). In contrast, that of the empty vector was not found. And a recombinant vector containing the overall ORFs was not induced (data not shown). Similar results have been reported with the squalene synthases of tobacco (Devarenne et al., 1998) and yeast (LoGrasso et al., 1993).

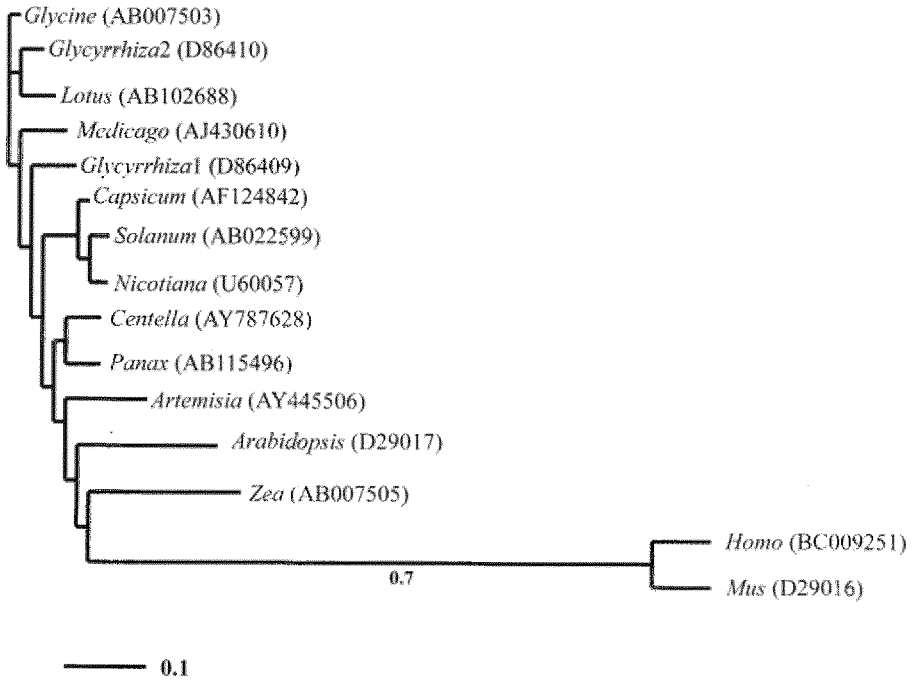

Figure 3. Phylogenetic relationships in SQS from various species. Database accession numbers are indicated in parentheses. Distances between each clone and group are calculated with CLUSTAL W program.

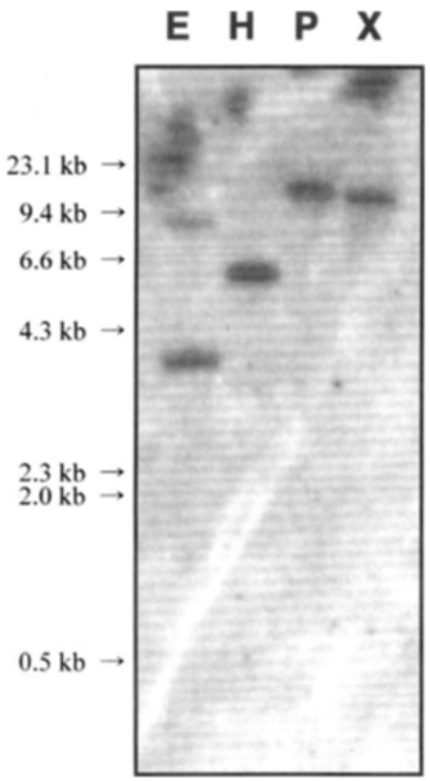

Figure 4. Southern blot analysis. Genomic DNA of *C. asiatica* was digested with restriction enzymes *Eco*RI (E), *Hin*dIII (H), *Pst*I (P), or *Xba*I (X), then Southern-blotted (10 μg DNA/track). Molecular size marker is indicated on the left.

Effect of MJ on Levels of *CaSQS* mRNA

Both triterpene saponin content and the levels of mRNA *SQS* are enhanced when methyl jasmonate is added to the culture media (Suzuki et al., 2002; Hayashi et al., 2003). Therefore, to examine its effect on the mRNA level of *CaSQS*, we treated the medium with various amounts of MJ. Here, the expression of *CaSQS* at 7 d after elicitation was strongly induced at a concentration of 0.1 mM MJ (Fig. 6). This finding is consistent with our previous results that the asiaticoside content (as triterpene saponin) in the leaves of *C. asiatica* is markedly increased by the addition of MJ (Kim et al., 2004b). Our northern analysis results also suggested that *CaSQS* might be associated with the regulation of triterpene saponin biosynthesis, although we found no direct evidence for a regulatory function by the SQS gene in this pathway.

To over-express *SQS* in transgenic plants, Lee et al. (2004) performed transformation with *Agrobacterium tumefaciens*, which harbored a binary vector that included the entire coding region for a squalene synthase (*PgSS1*) in the T-DNA. They previously integrated the *PgSS1* gene into the *P. ginseng* genome, such that the saponin content in transgenic plants was about 1.6 to 3.0 times higher than that measured in the control. This result was the first to report a key

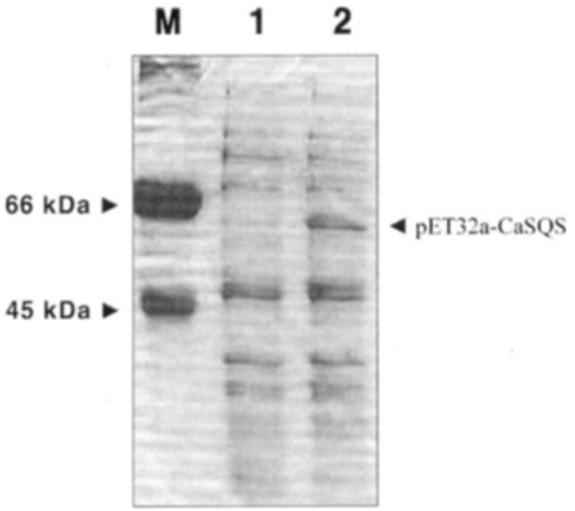

Figure 5. SDS-PAGE analysis of CaSQS expressed in *E. coli*. Lane M, molecular size markers; lane 1, protein was extracted from cells transformed with empty pET32a vector; lane 2, protein was extracted from cells transformed with pET32a-CaSQS. Position of CaSQS fused with Trx/His/S-taq at N terminus and His-taq at C terminus is about 61 kDa.

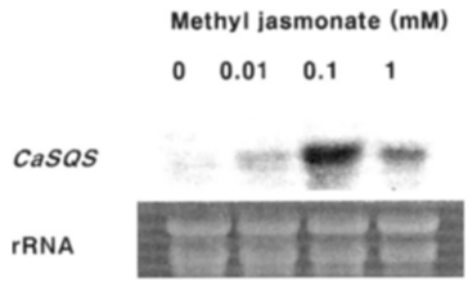

Figure 6. Effect of MJ concentration on mRNA levels of *CaSQS* in leaves of cultured whole plants. MJ was added to media at 5 weeks of culture. After 7 d of elicitation, total RNA was isolated from leaves (shown), resolved by agarose gel electrophoresis, blotted, and probed with *CaSQS*. Relative transcript abundance was calculated by dividing intensity of each transcript with that of corresponding rRNA transcript level.

regulatory enzyme, not only for phytosterol but also for triterpene saponin biosynthesis. Moreover, they showed that *SQS* over-expression resulted in the overproduction of triterpene saponins in *P. ginseng*. Therefore, based on all these results, we now propose that all *SQS* genes, including *CaSQS*, are candidates for further research to better understanding of this biosynthetic pathway, thereby enabling a molecular approach for improving triterpene saponin production in plants.

ACKNOWLEDGEMENT

This work was supported by a Korea Research Foundation Grant (KRF-2004-037-C00032).

Received December 13, 2004; accepted January 14, 2005.

LITERATURE CITED

Abe I, Rohmer M, Prestwich GD (1993) Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. Chem Rev 93: 2189-2206

Ahn JC, Paek YW, Sung CK, Kang GH, Hwang B (1993) Production of saponin by hairy root culture of *Bupleurum falcatum* L. I. Comparison of saponin content and pattern in callus, adventitious root, hairy root and cultivated root. Kor J Bot 36: 43-47

Akamine S, Nakamori K, Chechetka SA, Banba M, Umehara Y, Kouchi H, Izui K, Hata S (2003) cDNA cloning, mRNA expression, and mutational analysis of the squalene synthase gene of *Lotus japonicus*. Biochim Biophys Acta 1626: 97-101

Brinkhaus B, Lindner M, Schuppan D, Hahn EG (2000) Chemical, pharmacological and clinical profile of the East Asian medical plant *Centella asiatica*. Phytomedicine 7: 427-448

Devarenne TP, Shin DH, Back KW, Yin S, Chappell J (1998) Molecular characterization of tobacco squalene synthase and regulation in response to fungal elicitor. Arch Biochem Biophys 349: 205-215

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19: 11-15

Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85: 8998-9002

Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension culture of soybean root cells. Exp Cell Res 50: 195-202

Hayashi H, Hirota A, Hiraoka N, Ikeshiro Y (1999) Molecular cloning and characterization of two cDNAs for *Glycyrrhiza glabra* squalene synthase. Biol Pharm Bull 22: 947-950

Hayashi H, Huang PY, Inoue K (2003) Up-regulation of soyasaponin biosynthesis by methyl jasmonate in cultured cells of *Glycyrrhiza glabra*. Plant Cell Physiol 44: 404-411

Hostettmann K, Marston A (1995) Saponins. Cambridge University Press, Cambridge

Kim OT, Kim MY, Huh SM, Ahn JC, Seong NS, Hwang B (2004a) Effect of growth regulators on asiaticoside production in whole plant cultures of *Centella asiatica* (L.) Urban. J Plant Biol 47: 361-365

Kim OT, Kim MY, Hong MH, Ahn JC, Hwang B (2004b)

Stimulation of asiaticoside production from *Centella asiatica* whole plant cultures by elicitors. Plant Cell Rep 23: 339-344

Kribii R, Arro M, Del Arco A, Conzalez 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

Lee JH, Yoon YH, Kim HY, Shin DH, Kim DU, Lee IJ, Kim KU (2002) Cloning and expression of squalene synthase cDNA from hot pepper (*Capsicum annuum* L.). Mol Cells 13: 436-443

Lee MH, Jeong JH, Seo JW, Shin CG, Kim YS, In JG, Yang DC, Yi JS, Choi YE (2004) Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* overexpressing squalene synthase gene. Plant Cell Physiol 45: 976-984

Lian ML, Chakrabarty D, Peak KY (2002) Effect of plant growth regulators and medium composition on cell growth and saponin production during cell-suspension cultures of Mountain Ginseng (*Panax ginseng* C.A. Mayer). J Plant Biol 45: 201-206

LoGrasso PV, Soltis DA, Boettcher BR (1993) Overexpression, purification, and kinetic characterization of a carboxyl-terminal-truncated yeast squalene synthetase. Arch Biochem Biophys 307: 193-199

McKenzie TL, Jiang G, Straubhaar JR, Conrad DG, Shechter I (1992) Molecular cloning, expression, and characterization of the cDNA for the rat hepatic squalene synthase. J Biol Chem 267: 21368-21374

Nakashima T, Inoue T, Oka A, Nishino T, Osumi T, Hata S (1995) Cloning, expression, and characterization of cDNAs encoding *Arabidopsis thaliana* squalene synthase. Proc Natl Acad Sci USA 92: 2328-2332

Robinson GW, Tsay YH, Kienzle BK, Smith-Montroy CA, Bishop RW (1993) Conservation between human and fungal squalene synthetases: Similarities in structure, function, and regulation. Mol Cell Biol 13: 2706-2717

Suzuki H, Achnine L, Xu R, Matsuda SPT, Dixon RA (2002) A genomic approach to the early stages of triterpene saponin biosynthesis in *Medicago truncatula*. Plant J 32:1033-1048