

## Cloning and Expression Analysis of a Squalene Epoxidase Gene from *Ginkgo biloba*

Li ZHU, Liang-Qiong MA, Feng XU, Xiao-Juan YAN, Wei-Wei ZHANG\*

Yangtze University, College of Horticulture and Gardening, 434025, Jingzhou, China; [Zhuli322@163.com](mailto:Zhuli322@163.com); [maliangqiong@163.com](mailto:maliangqiong@163.com); [xufeng198@126.com](mailto:xufeng198@126.com); [15335763736@163.com](mailto:15335763736@163.com); [wuzhangchn@163.com](mailto:wuzhangchn@163.com) (\*corresponding author)

### Abstract

Squalene epoxidase is a key enzyme involved in triterpene saponins biosynthetic pathways in plants. In this study, the *SE* gene was isolated from *Ginkgo biloba* by RT-PCR, which designated as *GbSE*, GenBank accession number: KY751713. The length of *GbSE* gene is 1646 bp, contains an open reading frame of 1617 bp encoding 538 amino acids. The theoretical molecular weight and pI of GbSE protein was 58.3 kDa and 8.35, respectively. Sequence multiple alignment found that GbSE protein had high homology with SE proteins from other plants, including several highly conservative motifs and amino acids. Phylogenetic tree analysis showed that *Ginkgo biloba* SE belonged to the gymnospermous group, and was closely related to the SE protein of *Picea sitchensis*. A study of gene expression analysis indicated that the *GbSE* gene was highly expressed in stems, low expressed in fruits. The cDNA sequence of *GbSE* gene was cloned from *G. biloba* and the expression level of this gene was analyzed. This work can provide some theoretical basis for the research on the molecular mechanism of triterpenoid saponin biosynthesis.

**Keywords:** cloning, expression profile, *Ginkgo biloba*, squalene epoxidase, triterpenoid saponin

### Introduction

*Ginkgo biloba*, also named as Gongsun tree, Duck feet, belong to deciduous tree, is the only species of Ginkgo family, and is one of the oldest species in China (Zhou and Zheng, 2003). Due to *G. biloba* has survived from the fourth century of glacial coverage period so was known as “a living fossil” too (Xu *et al.*, 2008). Extracts of *G. biloba* of active compounds is complex, not only contains flavonoids, including biflavone, flavonoid glycoside and flavonoid aglycones, which mainly play a role in eliminating free radical and anti-lipid peroxidation in plants, but also contains terpene lactones, including diterpenes and sesquiterpenes, they have a strong specific inhibitory activity on the receptor (van Beek, 2002; Cheng *et al.*, 2009). Consequently, in the recent years, *G. biloba* especially as a medicinal plant has become one of popular functional plant favored by scholars, but mainly in the study of flavonoids little about triterpenoid saponin (Cao *et al.*, 2012).

Triterpenoid saponin, when *in vitro* of plant, beneficial to human body as immune adjuvant and lower cholesterol, and *in vivo* of plant also defend plants against attack from external environment and enhance the ability of resistance to pests and diseases (Osbourne and Haralampidis, 2002;

Yendo *et al.*, 2010). Triterpenoid saponins skeleton are synthesized by the isoprenoid pathway (Lambert *et al.*, 2011). Oxidosqualene (also named as squalene-2,3-epoxide), the precursor of the triterpenoid saponins in plants, is synthesized by mevalonate (MVA) pathway in the cytoplasm and 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway in plastids. Both pathways can synthetic isopentenyl pyrophosphate (IPP), and then by the farnesyl diphosphate synthase (FPS) converted to farnesyl diphosphate (FPP), subsequently, catalyzed into squalene by squalene synthase (SS) (Kuzuyama, 2002; Vranová *et al.*, 2013). Finally, under the effect of squalene epoxidase (SE) squalene is converted into oxidosqualene, which is the first oxygenation step in triterpenoid saponins biosynthesis (Fig. 1) (Haralampidis *et al.*, 2002). Thus, squalene epoxidase is one of the rate-limiting and a key enzyme in the triterpenoid saponins biosynthetic pathway.

Squalene epoxidase, also known as squalene monooxygenases, which located in the endoplasmic reticulum of eukaryotic cell, with the flavin adenine dinucleotide (FAD) binding site. The FAD is involved in the triterpenoid saponins biosynthesis process of key enzymes important cofactor about redox reaction (Favre and Ryder, 1996). Thus far, squalene epoxidase has been isolated from different species, such as rat (Sakakibara *et al.*, 1995), *Candida albicans* (Favre and Ryder, 1997), *Panax ginseng* (Han *et al.*, 2010), *Arabidopsis thaliana* (Rasbery *et al.*

al., 2007) and *Panax notoginseng*. Since SE is supposed to be a rate limiting enzyme, so its cloning and expression is found to play an important role in the biosynthesis of saponins. However, little is known about the SE involved in triterpenoid saponins biosynthesis in *G. biloba*. In the present study, we report for the first time the isolation, cloning, sequence analysis, tissue-specific expression of squalene epoxidase from *G. biloba*. We attempted to gain greater understanding of the regulatory role of squalene epoxidase in *G. biloba* triterpenoid biosynthesis. The results of this study may provide some theoretical reference and lay some molecular basis for the study of metabolic regulation of triterpenoid saponins and sterols biosynthesis.

## Materials and Methods

### Plant material

*G. biloba* plants were grown in the Yangtze University, Jingzhou, China. The plants were used as source plant material for the present study. Six different tissues (roots, stems, leaves, male flowers, female flowers and fruits) that were used for RNA extraction and expression analysis were collected, fresh tissue samples were then immediately frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until further analysis.

### Total RNA extraction

Different tissues of ginkgo were immediately pulverized in liquid nitrogen, total RNA was extracted from roots, stems, leaves, male flowers, female flowers and fruits following the instruction of TaKaRa MiniBEST Universal RNA Extraction Kit (TaKaRa, Dalian, China). The quality, concentration, and purity total RNA was detected at OD<sub>260</sub>/280 absorbance ratio by using a spectrophotometer and 1% (w/v) agarose gel electrophoresis.

### The cDNA synthesis and cloning of *GbSE*

First-strand cDNA was synthesized by using PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa, Dalian, China). According to the transcriptome sequencing data of *G. biloba*, the specific primer pair SE-FP (5'-AGGTTAAGCCAGTGGCGACCTC-3') and SE-RP (5'-CTAATTATCTACTGAGGAAGCTGAATT-3') were designed and synthesized. The PCR reaction mixture was ddH<sub>2</sub>O 40.5  $\mu\text{L}$ , 10 $\times$ Buffer (Mg<sup>2+</sup>) 5  $\mu\text{L}$ , dNTP (10 mmol·L<sup>-1</sup>) 1  $\mu\text{L}$ , SE-FP and SE-RP 1  $\mu\text{L}$ , cDNA template 1  $\mu\text{L}$ , Taq DNA polymerase (5 U· $\mu\text{L}^{-1}$ ) 0.5  $\mu\text{L}$ . The PCR reaction condition was followed as 94  $^{\circ}\text{C}$  for 5 min ; 30 cycles at 94  $^{\circ}\text{C}$  for 30 s, 60  $^{\circ}\text{C}$  for 30 s, 72  $^{\circ}\text{C}$  for 1 min; with an extension at 72  $^{\circ}\text{C}$  for 10 min. The PCR amplification products were analyzed by 1% gel electrophoresis, and then the target gene was recovered by TaKaRa MiniBEST Agarose Gel DNA Extraction Kit (TaKaRa, Dalian, China). After further purification, it was cloned into pMD19-T vector (TaKaRa, Dalian, China), and transformed into *Escherichia coli* DH5 $\alpha$  strain. The transformed bacteria were evenly applied to the plate for blue-white selection; the positive clones were validated by colony PCR and then sequenced in the Shanghai Sangon Biotechnology Company.

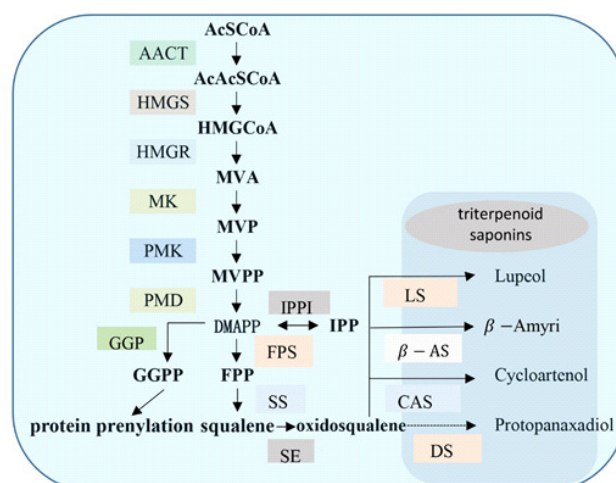


Fig. 1. The triterpenoid saponins biosynthesis of MVA pathway in *G. biloba*. AACT, Acetyl-CoA C-acetyltransferase; HMGS, 3-Hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-Hydroxy-3-methylglutaryl-CoA reductase; MK, MVA kinase; PMK, Phospho-MVA kinase; PMD, Diphospho-MVA decarboxylase; GGP, Geranyl diphosphate synthase; IPPI, Isopentenyl diphosphate; FPS, Farnesyl diphosphate synthase; SS, Squalene synthase; SE, Squalene epoxidase; LS, Lanosterol synthase; CAS, Cycloartenol synthase; DS, Dammarenediol synthase;  $\beta$ -AS,  $\beta$ -amyrin synthase

### Bioinformatics analyses

The isolated nucleotide sequence and deduced amino acid sequence were compared through database search using online bioinformatics tools (NCBI, <http://www.ncbi.nlm.nih.gov/BLAST/>). Using Vector NTI Suite V 11.5 and DNAMAN 8 to analyze cDNA sequence of *GbSE* gene. The online tool ExPASy (<http://web.expasy.org/protparam/>) was used to estimate the physicochemical properties of amino acid sequence of the *GbSE*. Program Align X (Vector NTI Suite V 11.5) was used to perform multiple alignment analysis of SE proteins from various plants, the Clustal X 2.0 and MEGA 6.0 were used to construct the phylogenetic tree of SE proteins with the neighbor-joining (NJ) method.

### Quantitative real-time PCR (qRT-PCR) analysis

Total RNA was isolated from six different samples by using TaKaRa MiniBEST Universal RNA Extraction Kit. First-strand cDNA was synthesized using a PrimeScript™ RT reagent Kit with gDNA Eraser (TaKaRa, Dalian, China). Real-time quantitative PCR were performed using the Mini Opticon™ Real-Time PCR system (Bio-Rad) with SYBR® Premix Ex Taq™ II Kit (TaKaRa, Dalian, China). Relative mRNA levels were normalized to a house-keeping gene (*GAPDH*) from ginkgo. The primer sequences of reference gene (*GAPDH*-F: 5'-CTGGCGTAGAGTATGTGGTTGAAT-3', *GAPDH*-R: 5'-CACGCCAACCAACGAACATG-3') and *GbSE* (SE-RTF: 5'-CGACCGTTGAGTTTGGTTCTG-3', SE-RTR: (5'-TCTGTCTGACACCTTCTGCCTTA-3') for the qRT-PCR were synthesized by Shanghai Sangon Biotechnology Company. qRT-PCR reactions were

performed in a final reaction volume of 25 µl according to the manufacturer's protocol. Each reaction sample was prepared in three technical replicates with a negative control using water as template. The relative expression level of *GbSE* gene was calculated by using the  $2^{-\Delta\Delta CT}$  method (Livak and Schmittgen, 2001).

**Results**

*Cloning of GbSE cDNA sequence*

By PCR amplification, a cDNA sequence of SE was obtained from *G. biloba*, which was designated as *GbSE* (GenBank accession no. KY751713). The cDNA sequence was 1646 bp, contains an open reading frame of 1617 bp, and encodes 538 amino acids (Fig. 2).

*Characterization of the deduced GbSE protein*

Putative GbSE protein contains 538 amino acids. The result of computer pI/Mw analysis showed the molecular weight and isoelectric point of GbSE protein were 58.3 kD

and 8.35, respectively. The secondary structure of *GbSE* via SOPMA tool analysis showed that the percentages of alpha helix, extended strand, random coil and beta turn were 34.57%, 23.79%, 32.16% and 9.48%, respectively. Homology analysis was completed with online tool BLASTP (NCBI) and Align X (Vector NTI 11.5). The results showed that GbSE protein belongs to NADB Rossmann superfamily, which contains the conservative domain structure FAD binding sites. Conservative areas are in red box, including G, L, RMR, RHPLTGGGMTV, and PDRIVGELLQPGG (Fig. 3). G and L are conservative sites that affect the function and associated with the active center of the enzyme. RxR conserved area plays an important role in the complexation of diphosphate after separation from the substrate (Starks and Noel, 1997). Domain motifs of the RHPLTGGGMTV and PDRIxGExxQPGG are the predicted FAD binding site (Sakakibara et al., 1995; Favre and Ryder, 1997). The putative GbSE protein sequence showed high identical to SE proteins from other plants (Fig. 3). The similarity between GbSE and SE proteins from *Picea sitchensis* (ABK24903), *Amborella trichopoda*

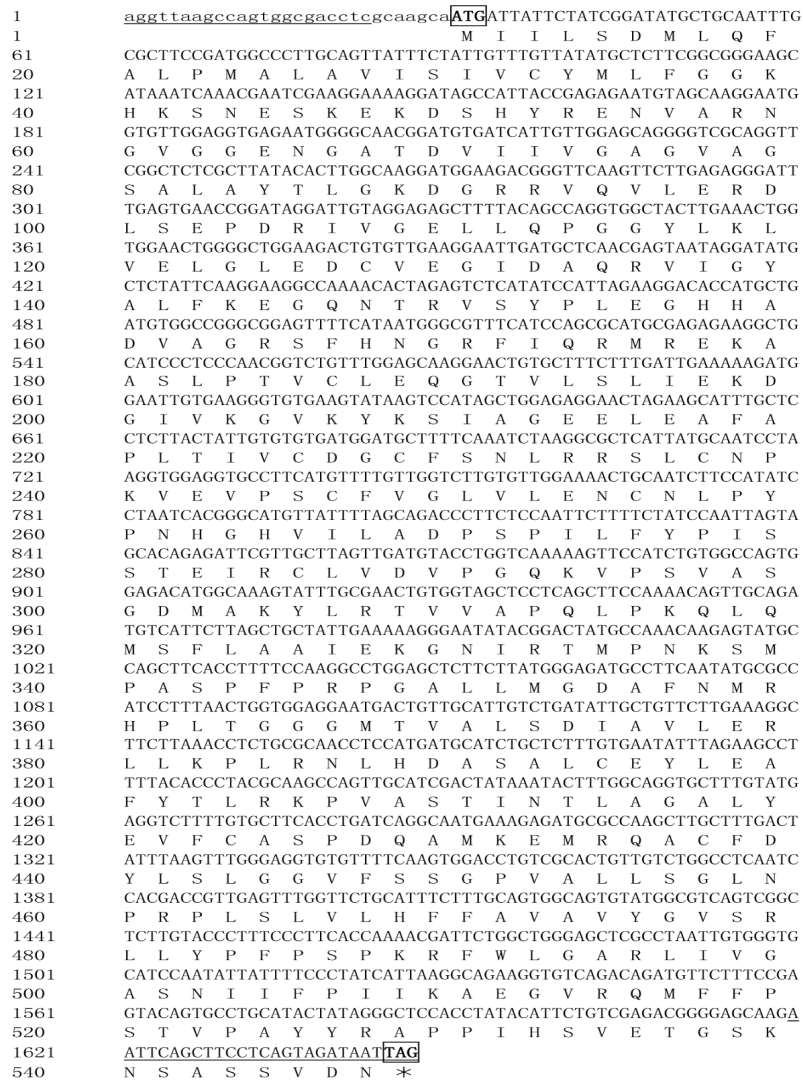


Fig. 2. The cDNA and deduced amino acid sequence of *GbSE*. The initial codon and the stop codon are highlighted in box, letters with underlined indicate the primer sequence

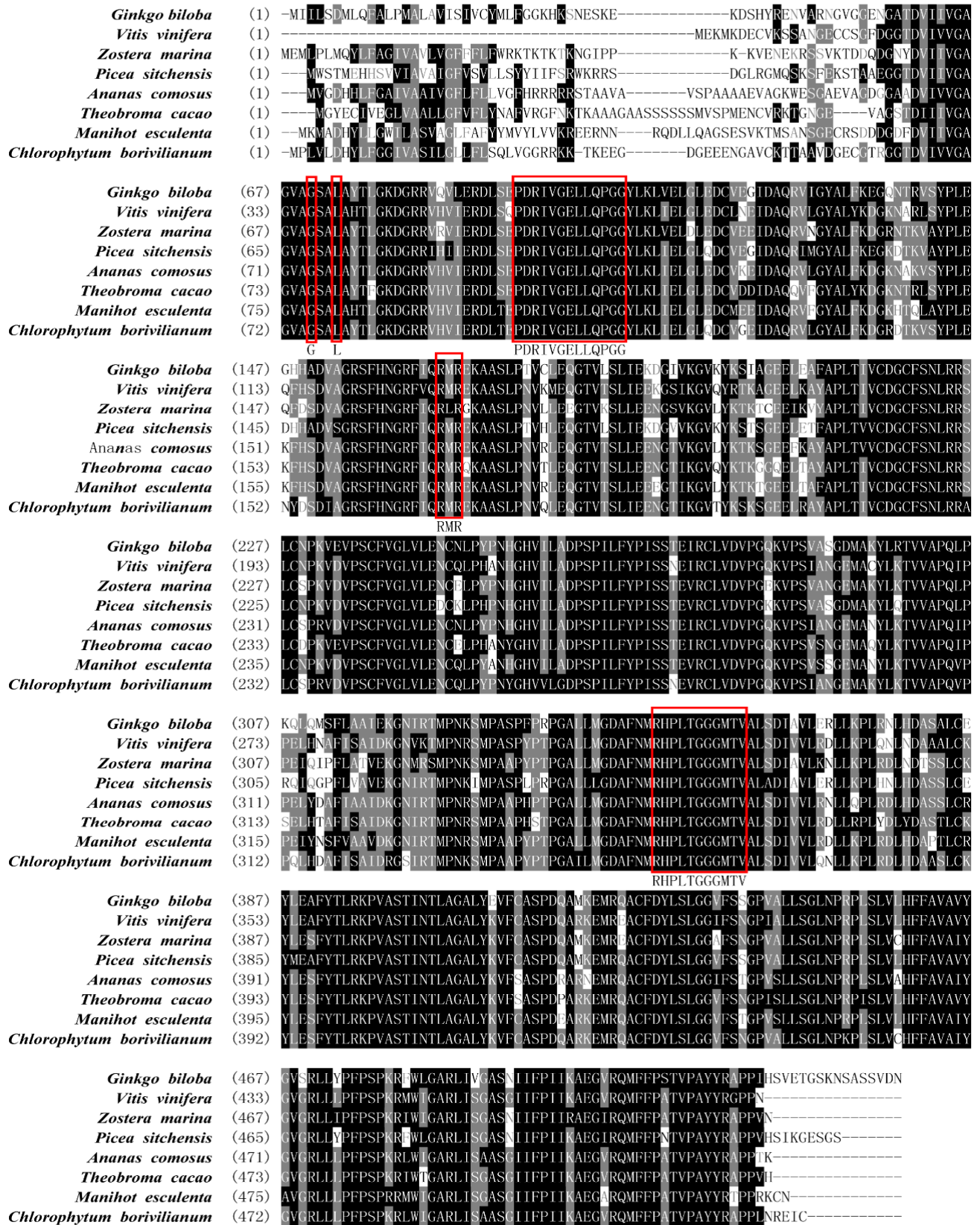


Fig. 3. Multiple alignment of amino acid sequences of SE. The completely identical amino acids are indicated with white foreground and black background. The conserved aminoacids are indicated with white foreground and grey background. Non-similar amino acids are indicated with black foreground and white background

(XP\_006857043), *Zostera marina* (KMZ75443), *Ananas comosus* (OAY63657), *Chlorophytum borivilium* (AFN61200), *Vitis vinifera* (CBI25076), *Zea mays* (ONL95395), *Manihot esculenta* (OAY51053), *Theobroma cacao* (EOX97273) were 89%, 81%, 77%, 83%, 78%, 80%, 81%, 83%, 81%, respectively. The homologous sequence of SE among different species indicated the SE proteins might keep a conservative during the molecular evolution.

#### Phylogenetic analysis of GbSE protein

The phylogenetic tree was constructed by ClustalX 2.0 and MEGA 6.0 to further explore the evolutionary relationship between GbSE from other proteins involved in triterpenoid saponins biosynthesis in plants. Phylogenetic analysis indicated that all of SE proteins descended from a common progenitor. In addition, SE proteins from different plants were clearly clustered into three branches, gymnospermous, dicots and monocots of angiospermous (Fig. 4). As shown in the phylogenetic tree, SE of *G. biloba* and *Picea sitchensis* (ABK24903) were clustered into the

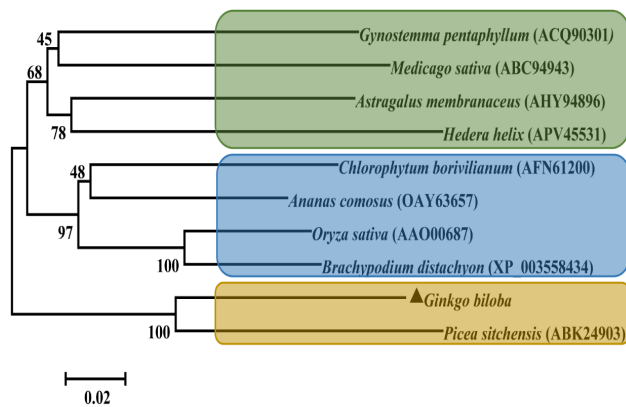


Fig. 4. Phylogenetic tree of SE from different plant species. The numbers at each node represented the bootstrap values (with 1000 replicates)

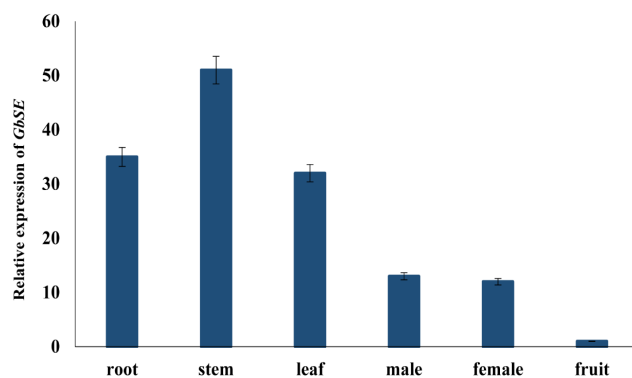


Fig. 5. Expression pattern of GbSE gene among different tissues of *G. biloba*

gymnospermous branch, suggesting they have the closer evolutionary distance. SE proteins from *Ananas comosus* (OAY63657), *Oryza sativa* (AAO00687), *Brachypodium distachyon* (XP\_003558434), *Chlorophytum borivilium* (AFN61200) were clustered into monocots branch of angiospermous, while SE proteins from *Gynostemma*

*pentaphyllum* (ACQ90301), *Hedera helix* (APV45531), *Astragalus membranaceus* (AHY94896), *Medicago sativa* (ABC94943) were clustered into dicots branch of angiospermous. These results suggest that SE genes from the same plant family will be priority clustered together, followed by clustering with the adjacent plant families, which indicated that the SE gene is conserved in the process of evolution, and this is consistent with the natural evolution of plants.

#### Expression profile of the GbSE in different tissues

Relative expression levels of GbSE in different tissues of *G. biloba* was analyzed by Real-time PCR with specific primers. Total RNA was extracted from roots, stems, leaves, male flowers, female flowers and fruits. The results revealed that GbSE was expressed in various tissues, but it was highly expressed in stems. Compared with GbSE mRNA levels in stems, GbSE mRNA levels showed significantly decrease in roots and leaves. The expression levels of GbSE in flowers were low, and no significant difference between male flowers and female flowers. The lowest expression level was observed in fruits (Fig. 5). The expression levels of GbSE in different tissues may be related to its function in *G. biloba*.

#### Discussion

Squalene epoxidase play important roles in triterpenoid saponins biosynthesis. In this study, the SE gene was isolated from *G. biloba*, sequence analysis showed that it contains an open reading frame of 1617 bp, encodes 538 amino acids, the theoretical molecular weight and pI of GbSE protein were 58.3 kD and 8.35, respectively. Multiple alignment results showed that GbSE protein has high homologous with SE proteins from *Picea sitchensis*, *Amborella trichopoda*, *Zostera marina*, *Ananas comosus*, *Chlorophytum borivilium*, *Vitis vinifera*, *Zea mays*, *Manihot esculenta*, *Theobroma cacao*. Phylogenetic tree analysis showed that SE proteins of *Picea sitchensis* and *G. biloba* from gymnospermous were clustered into the same branch, which indicated that GbSE was belonging to SE gene family. GbSE contains the FAD conservative domain structure, which bind to key enzymes as important cofactor of redox reaction in the biosynthesis process of triterpenoid saponins. In addition, RxR conserved is also found in other enzymes such as sesquiterpene synthase in *Chamaemelum*, plays an important role in the complexation of diphosphate after separation from the substrate (Cao et al., 2012).

Quantitative Real-time PCR analysis showed that GbSE display distinct tissue specific expression, it was higher expressed in the stems, lower expressed in the fruits. This phenomenon is probably related to the function of SE in *G. biloba*, Squalene epoxidase maybe also involved in triterpenoid saponins biosynthesis in *G. biloba*. Moreover, triterpenoid saponins can act as a chemical defense substance resistant to pests and diseases in plants (Mahato et al., 1988). In *Ziziphus jujuba*, triterpenoid saponin not only has the ability to scavenge free radicals and antioxidant, but also can be moisturizers and antibacterial for plants (Sun et al., 2012). Hence, the expression levels of GbSE in roots was high may also be related to the response of external

biological stress. In the future, we will study the relationship between *GbSE* expression, triterpenoid saponins content, and transcription factors. This research enables to improve understanding of key enzymes involved in triterpenoid saponins biosynthesis pathway.

### Acknowledgements

This study was supported by the National Natural Science Foundation of China (31370680).

### References

- Cao F, Cheng H, Cheng S, Li L, Xu F, Yu W, Yuan H (2012). Expression of selected *Ginkgo biloba* heat shock protein genes after cold treatment could be induced by other abiotic stress. *International Journal of Molecular Sciences* 13(5):5768-5788.
- Cheng SY, Xu F, Wang Y (2009). Advances in the study of flavonoids in *Ginkgo biloba* leaves. *Journal of Medicinal Plant Research* 3(13):1248-1252.
- Favre B, Ryder NS (1996). Characterization of squalene epoxidase activity from the dermatophyte *Trichophyton rubrum* and its inhibition by terbinafine and other antimycotic agents. *Antimicrobial Agents and Chemotherapy* 40(2):443-447.
- Favre B, Ryder NS (1997). Cloning and expression of squalene epoxidase from the pathogenic yeast *Candida albicans*. *Gene* 189(1):119-126.
- Gawel NJ, Jarret RL (1991). A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. *Plant Molecular Biology Reporter* 9(3):262-266.
- Han JY, Jungyo I, Yongsoo K, Yongeui C (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. *Advances in Biochemical Engineering* 75(75):31-49.
- Kuzuyama T (2002). Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units. *Bioscience Biotechnology and Biochemistry* 66(8):1619-1627.
- Lambert E, Faizal A, Geelen D (2011). Modulation of triterpene saponin production: *in vitro* cultures, elicitation, and metabolic engineering. *Applied Biochemistry and Biotechnology* 164(2):220-237.
- Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta C_T}$  Method. *Methods* 25(4):402-408.
- Mahato SB, Sarkar SK, Poddar G (1988). Triterpenoid saponins. *Phytochemistry* 27(10):3037-3067.
- Osbourn AE, Haralampidis K (2002). Chapter Five Triterpenoid saponin biosynthesis in plants. *Recent Advances in Phytochemistry* 36(02):81-93.
- Rasbery JM, Shan H, Leclair RJ, Norman M, Matsuda SP, Bartel B (2007). *Arabidopsis thaliana* squalene epoxidase 1 is essential for root and seed development. *Journal of Biological Chemistry* 282(23):17002-17013.
- Sakakibara J, Watanabe R, Kanai Y, Ono T (1995). Molecular cloning and expression of rat squalene epoxidase. *Journal of Biological Chemistry* 270(1):17-20.
- Starks CM, Noel JP (1997). Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science* 277(5333):1815-1820.
- Sun YF, Liang ZS, Liu Z, Huang J, Yan H (2012). Study on antimicrobial and antioxidant activities of triterpenoid saponins in sour jujube fruits. *Science and Technology of Food Industry* 33(6):139-142.
- van Beek TA (2002). Chemical analysis of *Ginkgo biloba* leaves and extracts. *Journal of Chromatography A* 967(1):21-55.
- Vranová E, Coman D, Gruissem W (2013). Network Analysis of the MVA and MEP Pathways for Isoprenoid Synthesis. *Plant Biology* 64(64):463-476.
- Xu F, Cheng H, Cai R, Li LL, Chang J, Zhu J, Zhang XF, Chen LJ, Wang Y, Cheng SH, Cheng SY (2008). Molecular cloning and function analysis of an anthocyanidin synthase gene from *Ginkgo biloba*, and its expression in abiotic stress responses. *Molecules and Cells* 26(6):536-547.
- Yendo ACA, Costa FD, Gosmann G, Fett Neto AG (2010). Production of Plant Bioactive Triterpenoid Saponins: Elicitation Strategies and Target Genes to Improve Yields. *Molecular Biotechnology* 46(1):94-104.
- Zhou Z, Zheng S (2003). Palaeobiology: The missing link in *Ginkgo* evolution. *Nature* 423(6942):821-822.