

Cloning and Expression Analysis of Mevalonate Kinase and Phosphomevalonate Kinase Genes Associated with MVA Pathway in Santalum Album

Meiyun Niu

South China Institute of Botany

Yuping Xiong

South China Institute of Botany

Haifeng Yan

Guangxi Academy of Agricultural Sciences

Xinhua Zhang

South China Institute of Botany

Yuan Li

South China Institute of Botany

Jaime A. Teixeira da Silva

P.O. Box 7, Miki-cho Post Office, Miki-cho, Ikenobe 3011-2, Kagawa-ken, 761-0799, Japan

Guohua Ma (✉ magh@scib.ac.cn)

South China Institute of Botany

Research Article

Keywords: Santalum album L., MVA pathway, Sesquiterpenoids, Mevalonate kinase, Phosphomevalonate kinase

Posted Date: December 29th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-132912/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on August 19th, 2021. See the published version at <https://doi.org/10.1038/s41598-021-96511-4>.

Abstract

Sandalwood is highly valued for its fragrant heartwood and its extracted oil. The major oil component santalols are terpenoids, which are biosynthesis through the MVA pathway. MK and PMK are the major enzymes on the MVA pathway. Little is known about the genes encoding MK and PMK in *Santalum album* on its expression regulation mechanism. The analysis of MK and PMK genes and their functions are important for the further study of the biosynthesis of santalol. These results will help to further study the role of MK and PMK genes in *S. album* santalol biosynthesis. The total RNA of sandalwood leaves was extracted, then the First-strand cDNA synthesis was obtained through the PrimeScript first-strand cDNA synthesis kit. Then sequence comparison and bioinformatics analyses of the genes homology of SaMK and SaPMK with MKs and PMKs, We also investigated subcellular localization of SaMK and SaPMK proteins. Its functional complementation of SaMK and SaPMK in yeast were also investigated. Atlast, MeJA was used to induce tissue-specific analysis and expression profiles of SaMK and SaPMK. The results showed that the full-length cDNA sequences of SaMK and SaPMK were 1409 bp and 1679 bp containing a 1381 bp open reading frame (ORF) encoding a polypeptide of 460 amino acids and a 1527 bp ORF encoding a polypeptide of 508 amino acids, respectively. Sequence comparison and bioinformatics analyses indicated that SaMK and SaPMK showed high homology with MKs and PMKs, respectively from other plant species. Further functional complementation of SaMK in an MK-deficient mutant yeast strain YMR208W and SaPMK in a PMK-deficient mutant yeast strain YMR220W confirmed that cloned SaMK and SaPMK cDNA encode a functional MK and PMK, respectively and mediated MVA biosynthesis in yeast. Tissue expression pattern analysis revealed that SaMK and SaPMK were constitutively expressed in all the tested tissues. SaMK was highly expressed in young leaves but least expressed in sapwood while SaPMK was highly expressed in roots and mature leaves, and least expressed in young leaves.

Introduction

Santalum album L., known as Indian sandalwood, belongs to the Santalaceae, and is a slow growing, evergreen, root semi-parasitic tree widely distributed in tropical and temperate regions such as India, Sri Lanka, Malay Archipelago and southern China (Bhargava et al. 2018; Akhtar and Shahzad 2019). Sandalwood is highly valued for its fragrant heartwood and its extracted oil that is used in incense, perfumes, cosmetics, pharmaceutical and ornamental carvings (Rao et al. 2007). Sandalwood oil shows a variety of biological activities, such as antiviral, anticarcinogenesis, antipyretic, antiseptic, antiscabietic, antitumor and diuretic effects (Kim et al. 2006; Misra and Dey 2013a; Bommareddy et al. 2019). Sandalwood usually yields 3-7% essential oil depending on the region and hemisphere (Demole et al. 1976). The value of a sandalwood tree depends on three important characteristics: the volume of heartwood, and the concentration and quality of its heartwood oil (Subasinghe et al. 2014). Specifications for sandalwood oil reported in the Food Chemicals Codex (FCC 2003) indicate that the international (ISO) standard for sandalwood oil stipulates a minimum free alcohols (calculated as santalol) content of 90%. The principal components of *S. album* essential oil distilled from heartwood are sesquiterpenoids, including four main sesquiterpene, namely α -santalol, β -santalol, epi- β -santalol and α -exo-bergamotol

(Howes et al. 2004; Jones et al. 2007; Baldovini et al. 2011). Previous studies showed that the α -santalol and β -santalol are responsible for the pleasant fragrance of sandalwood and that α -santalol is responsible for most of the essential oil's bioactivity (Burdock and Carabin 2008). However, global sandalwood resources are diminishing and are far from being applicable for commercial-scale production due to habitat destruction, over-exploitation and complex requirements for cultivation (Demole et al. 1976; Sandeep and Manohara 2019), limitations that can to some extent be overcome through the use of biotechnology (Teixeira da Silva et al. 2016). Therefore, it is necessary to improve the production of santalol by biotechnological methods to meet rapidly increasing commercial demands. Different strategies have been studied to increase the content of santalol (Zhang et al. 2017), including the treatment of *S. album* trees with exogenous substances (Radomiljac 1998), chemical synthesis (Lewis et al. 1967; Corey and Yamamoto 1970; Baumann and Hoffmann 1979; Brocke et al. 2008), heterologous expression (Ajikumar et al. 2010; Diaz-Chavea et al. 2013; Srivastava et al. 2015) and *in vitro* culture and transgenic strategies (Misra and Dey 2013b; Cheng et al. 2019). However, until now, the yields of santalol have been low and new strategies such as the use of key genes for increasing santalol production by genetic engineering are imperative. Because of the structural complexity of terpenoids, chemical synthesis of terpenoids is very difficult. Metabolic engineering of the MVA pathway provides an alternative approach to traditional synthesis of terpenoids (Chang and Keasling 2006).

In higher plants, there are two distinct routes to biosynthesize isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) which are the central five-carbon precursors of all isoprenoids: the mevalonate (MVA) pathway in the cytosol and the 2-methylerythritol-4-phosphate (MEP) pathway in plastids (Chappell 1995; Lichtenthaler 1999; Kuzuyama 2002). The MVA pathway predominantly synthesizes sesquiterpenoids, triterpenoids such as sterol, ubiquinones and other polyterpenoids (Dubey et al. 2003; Gutensohn et al. 2013). Separately, the MEP pathway mainly involves the biosynthesis of monoterpenoids, diterpenoids and other terpenoids such as hormones, plant pigments and plastoquinone (Laule et al. 2003). As santalol are sesquiterpenoids, we focused on the genes in the MVA metabolic pathway. Mevalonate kinase (MK, ATP: mevalonate-5-phosphotransferase; EC2.7.1.36) is the fourth enzyme that catalyzes the conversion of MVA into mevalonate-5-phosphate, and then phosphomevalonate kinase (PMK, E.C. 2.7.4.2) catalyzes the conversion of mevalonate-5-phosphate into mevalonate-5-diphosphate. MK belongs to the GHMP superfamily and catalyzes the first phosphorylation reaction in the MVA pathway (Liao et al. 2006). Studies have shown that geranyl diphosphate (GPP), farnesyl diphosphate (FPP) and geranylgeranyl diphosphate (GGPP) can inhibit MK activity (Riou et al. 1994). Since FPP, GPP and GGPP are important precursors of terpenoid synthesis, MK can play an important role in regulating the biosynthesis of terpenoids (Dorsey and Porter 1968). **Phylogenetic analysis** of different MK genes of **microorganism**, yeast, plant and mammal show that MK genes have high homology (Oulmouden and Karst 1991). This might indicate that the conservation properties of MK is under evolutionary pressure suggesting that MK may play an important role in general metabolism (Oulmouden and Karst 1991). PMK also belongs to the GHMP superfamily (Liao et al. 2006). Proteomics data has shown that MK and PMK are expressed at relatively low levels and may be targets for increasing overall isoprenoid production (Chu et al. 2003; Garcia 2014).

Recently, some *MK* genes have been isolated from plant species, such as *Hevea brasiliensis* (Williamson and Kekwick 1965), *Agave americana* (García-Peregrín 1972), *Arabidopsis thaliana* (Riou et al. 1994), *Catharanthus roseus* (Schulte et al. 2000), *Zea mays* (Alexandrov et al. 2009) and *Ginkgo biloba* (Chen et al. 2017). Some *PMK* genes also have been isolated from plant species, such as *A. thaliana* (Xiao et al. 2015), *H. brasiliensis* (Sando et al. 2008) and *Salvia miltiorrhiza* (Ma et al. 2012). However, previous reports on sandalwood terpenoids mainly focused on downstream enzymes such as sesquiterpene synthase which can yield sesquiterpene hydrocarbons (Jones et al. 2008; Jones et al. 2011; Rani et al. 2013; Zhang et al. 2019) and cytochrome P450 oxygenase which convert sesquiterpene hydrocarbons to corresponding sesquiterpene alcohols (Diaz-Chavea et al. 2013; Celedon et al. 2016). Little is known about the genes encoding *MK* and *PMK* in *S. album* or its expression regulation mechanism. Studies shows that overexpressed *MK* and *PMK* genes in *E. coli* DH10B can significantly improved *MK* and *PMK* protein levels and over three-fold improved final amorpha-4,11-diene titer (Batth et al. 2011). Therefore, the analysis of *MK* and *PMK* genes and their functions are important for the further study of the biosynthesis of santalol.

In the present study, two novel *MK* and *PMK* cDNAs, named as *SaMK* and *SaPMK*, respectively, were cloned and characterized from *S. album* by RACE technology for the first time. Their structure and function were analyzed by bioinformatics analysis and yeast complementation assays. In addition, the expression profiles of *SaMK* and *SaPMK* in various tissues, including roots, shoots, young leaves, mature leaves, sapwood and heartwood, were examined. The expression patterns of *SaMK* and *SaPMK* following the induction by methyl jasmonate (MeJA) were also investigated.

Results

Cloning and characterization of the full-length cDNA of *SaMK* and *SaPMK*

The full-length cDNA sequences of *SaMK* and *SaPMK* were obtained through RT-PCR and 5'/3' RACE. Full-length *SaMK* is 1409 bp and contains a 1170 bp ORF that encodes 389 deduced amino acid residues. Full-length *SaPMK* is 1679 bp and contains a 1527 bp ORF that encodes 508 deduced amino acid residues. The results of BLASTN analysis on NCBI revealed that the *SaMK* and *SaPMK* sequences were highly homologous to the *MK* and *PMK* genes from other plants (Table 2). The *SaMK* nucleotide sequence exhibited 76%, 75%, 75%, 74% and 73% similarity with *H. brasiliensis*, *Morus alba*, *Platycodon grandiflorus*, *Panax notoginseng* and *C. roseus*, respectively. The *SaPMK* nucleotide sequence exhibited 76% similarity with *Tripterygium wilfordii*, *M. alba* and *H. brasiliensis* and 74% similarity with *P. ginseng* and *Siraitia grosvenorii*. Therefore, these genes were designated as *SaMK* (GenBank accession No. MH018696) and *SaPMK* (GenBank accession No. MH018697).

Bioinformatics analysis of the deduced *SaMK* and *SaPMK* proteins

The ExPASy online tool was used to calculate the relative molecular weight, theoretical isoelectric point, instability index, aliphatic index and grand average of hydropathicity of the deduced SaMK and SaPMK proteins. The results are shown in Table 2. The relative MW of SaMK is 41.3 kDa which is similar with the MK protein of *A. thaliana* (Riou et al. 1994). The predicted MW of *A. thaliana* MK protein is 40.7 kDa. The MWs of PMKs from different species vary widely. In *S. cerevisiae*, the MW of PMK is 47.3 kDa (Tsay and Robinson 1991). In *C. roseus*, the estimated MW is 128 kDa (Schulte et al. 2000) and in *Matricaria chamomilla*, it is 53.58 kDa (Xu et al. 2018). The relative MW of SaPMK protein is 54.6 kDa which is similar with PMK protein of *M. chamomilla*. The theoretical isoelectric points of SaMK and SaPMK are 5.23 and 5.92, respectively. The instability index of SaMK protein is 33 and the instability index of SaPMK protein is 46 mean that they are both stable proteins (Table 3). The total average hydropathicity of SaMK protein is 0.113 means that SaMK protein is a hydrophobic protein and the total average hydropathicity of SaPMK protein is -0.073 means that SaPMK protein is a hydrophilic protein (Table 3). Transmembrane domain and signal peptides were predicted by TMHMM Server and SignalP, respectively. Results showed that SaMK and SaPMK have no transmembrane domain or signal peptide (Fig. 1) which indicated that SaMK and SaPMK are non-secretory proteins.

Predicted protein structure and functional domain

The amino acid sequences deduced from SaMK and SaPMK genes were analyzed by the NCBI Conserved Domains database (Fig. 2). The SaMK protein contains the N-terminal conserved region of GHMP kinase (GHMP) from aa 134-212 and C-terminal conservative region from aa 296-365 (Fig. 2a). SaMK has significant mevalonate kinase activity from aa 1-389 and SaPMK has significant phosphomevalonate kinase activity from aa 1-482 (Fig. 2b). The active site of mevalonate kinase exists between amino acids 139 and 150 (LPLGSGLGSSAA) in SaMK and is an ATP binding domain sequence of GHMP kinase (Fig 3a). SaPMK proteins (Fig. 3b) also contains the N-terminal conserved region of GHMP kinase from aa 182-252 and have three conserved motifs, namely GKVLLAGGY (10-18), GLGSSA (187-193) and GGGVPGAGG (448-456), which are the same as in previous studies in *Enterococcus faecalis* and *Streptococcus pneumoniae* (Romanowski et al. 2002; Doun et al. 2005). These findings confirm that SaMK and SaPMK have similar catalytic functions to the corresponding MK and PMK from other plants.

Molecular evolution of the deduced SaMK and SaPMK proteins

To investigate the evolutionary relationships among deduced SaMK and SaPMK proteins with other MKs and PMKs from angiosperms, gymnosperms, fungi, and bacteria, phylogenetic trees were constructed using the NJ method with MEGA 7. As shown in Fig. 4a, MKs from different species seemed to evolve into four different groups, with bacteria as an ancient group. SaMK belonged to the angiosperms group and was clustered into one group with *Siraitia grosvenorii* and *H. brasiliensis*. As shown in Fig. 4b. PMKs from different species also evolve into four different groups and with bacteria as an ancient group. SaPMK

belonged to the angiosperms group and was clustered into one group with *H. brasiliensis* and *Tripterygium wilfordii*. These results suggest that SaMK and SaPMK shared a common evolutionary origin with other plant MK and PMK proteins based on their conserved structures and sequence characteristics, such as amino acid homologies and conserved motifs.

Subcellular localization of SaMK and SaPMK proteins.

Predicted sub-cellular localization of SaMK and SaPMK proteins by PSORT showed that SaMK and SaPMK proteins may be localized in the cytoplasmic. To further verify the sub-cellular localization of SaMK and SaPMK protein, sub-cellular localization of SaMK-YFP and SaPMK-YFP (yellow fluorescent protein) were studied using a modified polyethylene glycol method to transform SaMK-YFP and SaPMK-YFP constructs to *A. thaliana* protoplasts. We found both SaMK and SaPMK proteins were located in cytosol (Fig 5). Our results suggested that SaMK and SaPMK cloned in this study may be involved in mevalonate pathway in *S. album*.

Functional complementation of SaMK and SaPMK in *Saccharomyces cerevisiae*

In yeasts, the MVA pathway is a biosynthetic pathway that is essential to survival. Disrupting MVA pathway genes in yeast strains can be fatal (Servouse and Karst 1986). Disruption of the chromosomal ERG8 coding region by integration of URA3 or HIS3 marker fragments was lethal in haploid cells, indicating that this gene is essential (Tsay and Robinson 1991). To verify the function of SaMK and SaPMK, the recombined expression vectors, pYES2-SaMK and pYES2-SaPMK, were successfully constructed. YMR208W, which harbored pYES2-SaMK, and YMR220W, which harbored pYES2-SaPMK, grew well on YPG medium. However, neither YMR208W which harbored pYES2-SaMK nor YMR220W which harbored pYES2-SaPMK or pYES2-SaHMGS could grow on YPD medium (Fig. 6). These results indicate that SaMK and SaPMK have MK and PMK activity, respectively.

Tissue-specific expression of SaMK and SaPMK

To determine the tissue-specific expression pattern of *SaMK* and *SaPMK* genes in *S. album*, total RNA was extracted from roots, heartwood, sapwood, young leaves, mature leaves and shoots, and qRT-PCR was performed. The results of qRT-PCR are shown in Fig. 7. SaMK and SaPMK were constitutively expressed in all detected tissues of *S. album*. As shown in Fig. 7a, the lowest transcript level of *SaMK* was observed in sapwood, and the highest expression level in young leaves followed by mature leaves and shoots, approximately 7.77-, 6.59- and 2.72-fold higher than sapwood. The expression level of *SaPMK* (Fig. 7b) in young leaves was lowest but highest in roots followed by mature leaves and sapwood, approximately 5.84-, 5.38- and 3.93-fold higher than young leaves. These studies revealed that *MK* and *PMK* genes may have distinct spatial and temporal expression patterns in different plant species.

Expression of *SaMK* and *SaPMK* in response to MeJA

MeJA is plant-specific signaling molecule that is involved in the regulation of various biological processes (Pauwels et al. 2008). In the present study, we measured the expression level of *SaMK* and *SaPMK* in *S. album* roots, shoots and leaves after treatment with 100 μ M MeJA (Fig. 8). The expression of both genes was significantly induced by MeJA. The trend in the change in transcript level of *SaMK* in *S. album* shoots and leaves and the expression level of *SaPMK* in *S. album* roots, shoots and leaves after MeJA treatment was consistent, all increasing gradually and peaking at 12 h and then gradually decreasing compared with control seedlings. However, there was no obvious change in the expression level of *SaMK* in roots.

Discussion

Terpenoids, including monoterpenes, sesquiterpenes and diterpenes, plays an important role in plant physiology and ecology (Nagegowda 2010). In recent years, many studies have documented the molecular regulation of sesquiterpene biosynthesis (Degenhardt et al. 2009). Santalol, a sesquiterpenoid, is the most dominant aromatic and active ingredient in sandalwood essential oil (Demole et al. 1976). Santalol is mainly synthesized via the MVA pathway. MK is one a rate-limiting enzyme in the MVA pathway that catalyzes the phosphorylation of mevalonic acid to produce mevalonate-5-phosphate. PMK is a potential regulatory enzyme of the isoprenoid biosynthetic pathway (Lluch et al. 2000). The functions of MK and PMK proteins in the MVA pathway have been studied in many plant species (Riou et al. 1994; Schulte et al. 2000). Mevalonate kinase expression level is related to the precursor of terpenoids biosynthesis, IPP and DMAPP, which can indirectly regulate the biosynthesis of terpenoids. Studies show that overexpression of PMK gene can increase the content of terpenoids (Woo et al. 2013). Yuan (Yuan et al. 2013) founded that there was a significant correlation between *plpmk* gene expression level and accumulation of paeoniflorin and other substances in *Paeonia lactiflora*. However, there are currently no reports on the genes upstream of the MVA pathway in the biosynthesis of sesquiterpenoids in *S. album*. Thus, we attempted to examine the molecular biology of the MVA pathway via santalol biosynthesis in *S. album* by cloning, characterization, and functional analysis of full-length cDNAs of *SaMK* and *SaPMK* genes.

In this study, a 1409 bp full-length cDNA of the *SaMK* gene and a 1679 bp full-length cDNA of the *SaPMK* gene were isolated from *S. album*. The deduced *SaMK* protein contained 389 amino acids and weighed 41.30 kDa, which is consistent with previous studies. The deduced *SaPMK* protein contained 508 amino acids and weighed 54.60 kDa. Multiple alignments showed that the deduced *SaMK* and *SaPMK* protein sequences were very similar to other plant MKs and PMKs, respectively. A phylogenetic tree revealed that the *SaMK* protein had highest homology with MK from *Hevea brasiliensis* and *Siraitia grosvenorii* while the *SaPMK* protein had highest homology with PMK from *H. brasiliensis* and *Tripterygium wilfordii*. This trend indicates that *SaMK* and *SaPMK* shared common evolutionary origins with other MK and PMK proteins based on amino acid sequences and functional domains. Studies show that MK protein of *Catharanthus roseus* localized in cytosolic while PMK protein of *Catharanthus roseus* localized in peroxisomes (Simkin et al. 2011). However, *SaMK* and *SaPMK* proteins both localized in cytosolic. Further

researches are needed to make sure where the final steps of the plant MVA pathway are localised. SaMK and SaPMK complementation assays in yeast revealed that the expression of *SaMK* and *SaPMK* provided basic nutrients for the survival of yeast, thereby confirming the catalytic function of SaMK and SaPMK (Tsay and Robinson 1991; Chen et al. 2017).

Sandalwood is considered to be one of the most valuable trees in the world (Fox 2000). Its value lies mainly in heartwood and the essential oil extracted from heartwood (Rao et al. 2007). Santalol is responsible for the pleasant fragrance of sandalwood (Burdock and Carabin 2008) and most of the pharmacological activity of the oil (Kim et al. 2006). Thus, it is important to investigate whether or not *SaMK* and *SaPMK* transcripts may be related to the accumulation of santalol in different *S. album* tissues. qRT-PCR showed that *SaMK* and *SaPMK* genes were constitutively expressed in all the detected tissues but were expressed differently in various tissues, thus having different expression patterns. The transcript level of *SaMK* in young leaves was significantly higher than in other tissues, and its expression level was lowest in sapwood. A similar expression pattern of the *SaMK* gene was found in *Ginkgo biloba*, in which *GbMK* genes were highly expressed in roots and leaves (Chen et al. 2017). In *H. brasiliensis*, the *HbMK* gene was highly expressed in latex, followed by xylem and mature leaves (Sando et al. 2008). The level of *SaPMK* transcript in roots was significantly higher than in other tissues, and its expression level was lowest in young leaves. In *H. brasiliensis*, the *HbPMK* gene was highly expressed in xylem, followed by latex (Sando et al. 2008). In *Panax ginseng*, the *PgPMK* gene was highly expressed in fine roots, followed by lateral roots (Kim et al. 2014).

MeJA is an important cellular regulator involved in diverse developmental processes, such as seed germination, root growth, fertility, fruit ripening, and senescence (Cheong and Choi 2003). There is a relationship between MeJA and terpene metabolism, including campestene-triol (Robert-Seilaniantz, A. et al. 2011; Farag et al. 2017). MeJA promoted the production of monoterpenoids and sesquiterpenoids in *Ocimum basilicum* (Jiang et al. 2016). A cembranoid was upregulated in *Sarcophyton glaucum* treated with 200 μ M MeJA (Cheong and Choi 2003). MeJA enhanced the production of the diterpenoid tanshinone in *Salvia miltiorrhiza* and *Ginkgo biloba* (Gao et al. 2009; Liao et al. 2015). Zhang (Zhang et al. 2019) found that MeJA can induce the expression of *SaTPS1* and *SaTPS2* in leaves of *Santalum album*. Hence, studying the expression profiles of *SaMK* and *SaPMK* genes following treatment with MeJA is important, as it may provide more insight into the regulation of *SaMK* and *SaPMK* in santalol biosynthesis. In the present study, we measured the expression levels of *SaMK* and *SaPMK* in *S. album* roots, shoots and leaves after treatment with 100 mM MeJA. The *SaMK* expression level in shoots and leaves, and the *SaPMK* expression level in roots, shoots and leaves were significantly induced by MeJA. However, there was no obvious change in the expression level of *SaMK* in roots. The trend in the change in transcript level of *SaMK* in *S. album* shoots and leaves after MeJA treatment was consistent, all increasing gradually and peaking at 12 h. The trends in the change in transcript level of *SaPMK* in *S. album* roots, shoots and leaves after MeJA treatment were consistent, all gradually increasing and peaking at 12 h after MeJA treatment, then gradually decreasing. The *MK* gene transcript was induced by 1 mM MeJA in *Ginkgo biloba* (Chen et al. 2017) and the *PMK* gene transcript was induced by 100 μ M MeJA in *Panax ginseng* (Kim et al. 2014). Similar to previous reports, our results showed that *SaMK* and

SaPMK expression levels in *S. album* increased obviously after MeJA treatment. The expression of *SaMK* and *SaPMK* genes was upregulated by MeJA, indicating that these inducible genes might be involved in signal molecule-related responses to environmental stimuli.

The characterization and expression profiles of *SaMK* and *SaPMK* genes may contribute to an understanding of the biosynthesis of sesquiterpenes in *S. album* at the molecular level and the regulatory mechanisms involved in the MVA pathway.

Materials And Methods

Five-year-old sandalwood trees that had been growing in South China Botanical Garden, Guangzhou, China, were used. The young and mature leaves, heartwood, sapwood, roots and shoots were collected and wrapped in tin foil paper, frozen immediately in liquid nitrogen, and stored at -80°C for later use. Two-month-old young seedlings (6-8 leaves) of *S. album* were sprayed with 100 µM MeJA until the leaf surfaces were wet. 2% alcohol served as the control for each treatment. Samples were collected at 0 h, 2 h, 6 h, 12 h, 24 h, 48 h and 72 h after treatment and stored at -70°C for further analyses. Each treatment was repeated three times.

Cloning of the full-length putative cDNA of *SaMK* and *SaPMK* by RACE

The total RNA of sandalwood leaves was extracted using Column Plant RNAOUT (Tiandz, Beijing, China) according to the manufacturer's instruction. The concentration and quality of RNA were measured using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, NC, USA) and agarose gel electrophoresis.

First-strand cDNA synthesis was obtained through the PrimeScript first-strand cDNA synthesis kit (Takara Bio Inc., Dalian, China). 5' and 3' RACE was performed according to the SMARTer RACE cDNA Amplification Kit (Clontech Laboratories Inc., CA, USA) manual. Primers were designed on the basis of the initial data of MK and PMK unigenes in the transcriptome (Zhang et al. 2015) (Table 1). The sequence information of 5' and 3' RACE PCR product clones were used to design primers from the start and stop codon to obtain the internal fragments. The amplified PCR products were purified by a gel DNA purification kit (Tiangen, Beijing, China) and ligated into pMD18-T vector (Takara Bio Inc.). The recombinant plasmids were transformed into *Escherichia coli* DH5α competent cells (Takara Bio Inc.) and sequenced at the Beijing Genomics Institution (BGI, Shenzhen, China).

Bioinformatics analysis and molecular evolution analysis of *SaMK* and *SaPMK*

SaMK and *SaPMK* gene sequences were assembled and translated into amino acid sequences using DNAMAN software. The open-reading frame of the *SaMK* and *SaPMK* genes were predicted by ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>). Sequence comparison was performed with NCBI BLAST online tools (<http://www.ncbi.nlm.nih.gov/BLAST/>). The physicochemical properties of the deduced SaMK and SaPMK proteins were calculated by ExPASy (<http://cn.expasy.org>). Protein domains and active sites were predicted by the CDD database in NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Transmembrane domains and signal peptides were predicted by the TMHMM Server (<http://www.cbs.dtu.dk/services/TMHMM/>) and SignalP (<http://www.cbs.dtu.dk/services/SignalP/>), respectively. Multiple sequence alignment was performed with CLUSTALX 2.0 (Conway Institute, UCD Dublin, Dublin, Ireland) and phylogenetic tree of SaMK and SaPMK proteins from *S. album* and other plants were constructed by MEGA 7 through the neighbor-joining (NJ) method with 1000 bootstrap replicates (Saitou and Nei 1987).

Subcellular localization of SaMK and SaPMK proteins.

A vector pSAT6-EYFP containing enhanced yellow fluorescent protein (EYFP) open reading frame was used in this study. The cDNA encoding SaMK and SaPMK were amplified with two pairs of primers YFP-MK-F and YFP-MK-R, YFP-PMK-F and YFP-PMK-R, respectively (Table 1). The PCR products and pSAT6-EYFP vector were digested with corresponding **endonuclease** restriction enzymes. The digested fragment was ligated into pSAT6-EYFP linearized vector to generate pSAT6-EYFP-SaMK and pSAT6-EYFP-SaPMK fusion constructs. The fusion expression vectors and the pSAT6-EYFP vector were transformed into arabidopsis mesophyll protoplasts through PEG-mediated transformation followed the method described previously (Yoo et al. 2007). Using a confocal laser-scanning microscope (Leica TCS SP8 STED 3X, Wetzlar, Germany) to observe YFP fluorescence in transformed protoplasts after overnight incubation at 22°C.

Functional complementation of *SaMK* and *SaPMK* in yeast

The two recombined plasmids, pYES2-SaMK and pYES2-SaPMK, were constructed by the In-Fusion HD Cloning Kit (Takara Bio Inc.) according to the manufacturer's instructions. The pYES2 vectors (Invitrogen, Carlsbad, CA, USA), containing a yeast galactose-dependent promoter that can promote high levels of expression of target genes, were used as carriers for target genes in this study. The recombined plasmids (pYES2-SaMK and pYES2-SaPMK) were extracted and transformed into YMR208W (Δ ERG12) and YMR220W (Δ ERG8) (Dharmacon, Chicago, IL, USA), respectively with the Frozen-EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA, USA). Transformants were spotted on SC (-Ura) medium (6.7% yeast nitrogen base without amino acids, 2% galactose) (Chen et al. 2017). Positive clones were further confirmed by PCR. Subsequently, transformed diploid cells were induced to sporulate and formed haploid cells containing pYES2-SaMK and pYES2-SaPMK. To further observe their growth conditions, the diploid *Saccharomyces cerevisiae* strain YSC1021 and transformed haploid strains YMR208W and YMR220W

were grown separately on YPD (1% yeast extract, 2% bacto peptone, 2% glucose) and YPG (1% yeast extract, 2% bacto peptone, 2% galactose) media, respectively (Albers and Larsson 2009; Tao et al. 2016).

Tissue-specific analysis and expression profiles of *SaMK* and *SaPMK* induced by MeJA

To investigate the expression levels of *SaMK* and *SaPMK* genes in different tissues (roots, sapwood, heartwood, young leaves, mature leaves and shoots) and their expression profiles after MeJA treatment, qRT-PCR was carried out according to the manufacturer's instructions. About 1.0 µg of total RNA was reverse transcribed into first-strand cDNA using the PrimeScript RT Reagent Kit (Takara Bio Inc.) according to the manufacturer's protocols. The reactions were performed on ABI7500 fluorescence quantitative PCR (Applied Biosystems, Thermo Fisher Scientific, MA, USA) using iTaq Universal SYBR Green supermix as the buffer (Applied Biosystems, USA). PCR amplification was performed under the following conditions: 95°C for 30 s, followed by 35 cycles of 95 °C for 15 s and 60 °C for 60 s and melting curve analyses were performed. The housekeeping gene, β-actin, was selected as the internal control (Zhang et al. 2015) for normalization of all the reactions. All experiments were performed in triplicate and the mean value was analyzed. The $2^{-\Delta\Delta CT}$ method was used to analyze the relative expression level of genes (Schmittgen et al. 2008).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was financially supported by the National Key Research Plan of China (Grant Number: 2016YFC050304) and the National Natural Science Foundation of China (grant numbers 31470685 and 31270720) and a Guangdong Key Science and Technology project (2015B020231008).

Authors' contributions

MYN, YPX, HFY, XHZ, YL, JATdS and GHM designed the experiment and provided guidance for the study. MYN and XHZ prepared samples for all analyses. YPX and HFY participated in statistical analyses. MYN, GHM and JATdS co-wrote the manuscript. GHM interpreted the data and experimental results. All authors read and approved the manuscript for publication.

References

1. Akhtar, R. and Shahzad, A. Morphology and ontogeny of directly differentiating shoot buds and somatic embryos in *Santalum album* J. Forestry Res. 30:1179–1189 (2019). <https://doi.org/10.1007/s11676-018-0679-5>
2. Bhargava, P. et al. A modified and improved protocol development for *in vitro* clonal propagation of *Santalum album* from internodal explants. Trop. Plant Res. 5:193–199 (2018). <https://doi.org/10.22271/tpr.2018.v5.i2.024>
3. Rao, M. N. et al. Assessing threats and mapping sandal (*Santalum album*) resources in peninsular India: Identification of genetic hot-spot for *in-situ* conservation. Conserv. Genet. 8: 925–935 (2007). <https://doi.org/10.1007/s10592-006-9247-1>
4. Bommareddy, et al. Medicinal properties of alpha-santalol, a naturally occurring constituent of sandalwood oil: review. Nat. Prod. Res. 33:527–543 (2019). <https://doi.org/10.1080/14786419.2017.1399387>
5. Kim, T. H. et al. New antitumor sesquiterpenoids from *Santalum album* of Indian origin. Tetrahedron 62: 6981–6989 (2006). <https://doi.org/10.1016/j.tet.2006.04.072>
6. Misra, B. B. and Dey, S. Evaluation of in vivo anti-hyperglycemic and antioxidant potentials of α -santalol and sandalwood oil. Phytomedicine (2013). <https://doi.org/10.1016/j.phymed.2012.12.017>
7. Demole, E. et al. A chemical investigation of the volatile constituents of east Indian sandalwood oil (*Santalum album*). Helv. Chim. Acta 59: 737–747 (1976). <https://doi.org/10.1002/hlca.19760590304>
8. Subasinghe, U. et al. Essential oil content and composition of Indian sandalwood (*Santalum album*) in Sri Lanka. J. of Forestry Res. 24:4 (2014). <https://doi.org/10.1007/s11676-013-0331-3>
9. Sandalwood oil, East Indian type. Food Chemicals Codex, fifth ed. National Academy Press, Washington DC, p. 395(2003).
10. Jones, C. G. et al. Non-destructive sampling of Indian sandalwood (*Santalum album*) for oil content and composition. J. Essent. Oil Res. 19: 157–164 (2007).

<https://doi.org/10.1080/10412905.2007.9699250>

11. Baldovini, N. et al. Phytochemistry of the heartwood from fragrant *Santalum* species: a review. *Flavor Frag. J.* 26: 7–26 (2011). <https://doi.org/10.1002/ffj.2025>
12. Howes, M. J. et al. Evaluation of the quality of sandalwood essential oils by gas chromatography-mass spectrometry. *J. Chromatog. A.* 1028: 307–312 (2004).
<https://doi.org/10.1016/j.chroma.2003.11.093>
13. Burdock, G. A. and Carabin, I. G. Safety assessment of sandalwood oil (*Santalum album*). *Food Chem. Toxicol.* 46: 421–432 (2008). <https://doi.org/10.1016/j.fct.2007.09.092>
14. Sandeep, C. and Manohara, T. N. Sandalwood in India: historical and cultural significance of *Santalum album* as a basis for its conservation. *Nebio* 10:235–242 (2019).
15. da Silva Teixeira, J. A. et al. Sandalwood: basic biology, tissue culture, and genetic transformation. *Planta* 243: 847–887 (2016). <https://doi.org/10.1007/s00425-015-2452-8>
16. Zhang, Y. Y. et al. Multiple strategies for increasing yields of essential oil and obtaining sandalwood terpenoids by biotechnological methods in sandalwood. *Trees* 32: 17–28 (2017).
<https://doi.org/10.1007/s00468-017-1558-y>
17. Radomiljac, A. M. *Santalum album* plantations: a complex interaction between parasite and host. PhD thesis, Murdoch University, Perth, Australia, pp. 154–181 (1998).
18. Lewis, R. G. et al. Total synthesis of alpha-santalol. *Tetrahedron Lett.* 8: 401–406 (1967).
[https://doi.org/10.1016/s0040-4039\(00\)90958-4](https://doi.org/10.1016/s0040-4039(00)90958-4)
19. Corey, E. J. and Yamamoto, H. Modification of reaction of permit stereospecific synthesis of certain trisubstituted olefins stereospecific synthesis alpha-santalol. *J. Am. Chem. Soc.* 92: 226–228 (1970).
<https://doi.org/10.1021/ja00704a052>
20. Baumann, M. and Hoffmann, W. Synthese von rac-β-Santalol und rac-β-dehydrosantalol. *Liebigs Ann. Chem.* 126: 743–750 (1979). <https://doi.org/10.1002/jlac.197919790602>
21. Brocke, C. et al. Recent developments in the chemistry of sandalwood odorants. *Chem. Biodivers.* 5: 1000-1010 (2008). <https://doi.org/10.1002/cbdv.200890080>
22. Ajikumar, P. K. et al. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* 330: 70–74 (2010). <https://doi.org/10.1126/science.1191652>
23. Diaz-Chavea, M. L. et al. Biosynthesis of sandalwood oil: *Santalum album* CYP76F cytochromes P450 produce santalols and bergamotol. *PLoS ONE* 8: e75053 (2013).
<https://doi.org/10.1371/journal.pone.0075053>
24. Srivastava, P. L. et al. Functional characterization of novel sesquiterpene synthases from Indian sandalwood, *Santalum album*. *Sci. Rep.* 5: 10095 (2015). <https://doi.org/10.1038/srep10095>
25. Misra, B. B. and Dey, S. Culture of East Indian sandalwood tree somatic embryos in air-lift bioreactors for production of santalols, phenolics and arabinogalactan proteins. *AoB Plants* 5: plt025 (2013).
<https://doi.org/10.1093/aobpla/plt025>

26. Cheng, Q. W. et al. *Santalum album* callus co-cultured with fungus *Colletotrichum gloeosporioides* accumulates santalenes and bisabolene. *Trees* 33: 305–312 (2019).
<https://doi.org/10.1007/s00468-018-1758-0>
27. Chang, M. C. and Keasling, J. D. Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* 2:674–681 (2006). <https://doi.org/10.1038/nchembio836>
28. Chappell, J. Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. *Annu. Rev. Plant Biol.* 46: 521–547 (1995). <https://doi.org/10.1146/annurev.pp.46.060195.002513>
29. Lichtenthaler, H. K. The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Rev. Plant Biol.* 50: 47–65 (1999). <https://doi.org/10.1146/annurev.arplant.50.1.47>
30. Kuzuyama, T. Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units. *Biosci. Biotech. Bioch.* 66: 1619–1627 (2002). <https://doi.org/10.1271/bbb.66.1619>
31. Dubey, V. S. et al. An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *J. Biosciences* 28: 637–646 (2003). <https://doi.org/10.1007/BF02703339>
32. Gutensohn, M. et al. Cytosolic monoterpene biosynthesis is supported by plastid-generated geranyl diphosphate substrate in transgenic tomato fruits. *Plant J.* 75: 351–363 (2013).
<https://doi.org/10.1111/tpj.12212>
33. Laule, O. et al. Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Natl. Acad. Sci. USA* 100: 6866–6871 (2003).
<https://doi.org/10.1073/pnas.1031755100>
34. Liao, Z. H. et al. Isoprenoid biosynthesis in plants: pathways, genes, regulation and metabolic engineering. *J. Biol. Sci.* 6: 371–374 (2006). <https://doi.org/10.3923/jbs.2006.209.219>
35. Riou, C. et al. Isolation and characterization of a cDNA encoding *Arabidopsis thaliana* mevalonate kinase by genetic complementation in yeast. *Gene* 148: 293–297 (1994).
[https://doi.org/10.1016/0378-1119\(94\)90701-3](https://doi.org/10.1016/0378-1119(94)90701-3)
36. Oulmouden, A. and Karst, F. Nucleotide sequence of the *ERG12* gene of *Saccharomyces cerevisiae* encoding mevalonate kinase. *Curr. G* 19: 9–14 (1991). <https://doi.org/10.1007/BF00362081>
37. Dorsey, J. K. and Porter, J. W. The inhibition of mevalonic kinase by geranyl and farnesyl p J. Biol. Chem. 243:4667 (1968). [https://doi.org/10.1016/S0993-9857\(04\)97346-3](https://doi.org/10.1016/S0993-9857(04)97346-3)
38. Chu, S. et al. Expression and purification of Arg196 and Lys272 mutants of mevalonate kinase from *Methanococcus jannaschii*. *Protein Expres. Purif.* 30:210-218 (2003). [https://doi.org/10.1016/S1046-5928\(03\)00101-3](https://doi.org/10.1016/S1046-5928(03)00101-3)
39. Garcia, D. E. Kinetics of phosphomevalonate kinase from *Saccharomyces cerevisiae*. *Plos One*, 9:e87112 (2014). <https://doi.org/10.1371/journal.pone.0087112>
40. Williamson, I. P. and Kekwick, R. G. The formation of 5-phosphomevalonate by mevalonate kinase in *Hevea brasiliensis* *Biochem. J.* 96: 862–871 (1965). <https://doi.org/10.1042/bj0960862>
41. García-Peregrín, E. Mevalonate phosphorylation in *Agave americana*. *Phytochemistry* 11: 2495–2498 (1972). [https://doi.org/10.1016/S0031-9422\(00\)88523-4](https://doi.org/10.1016/S0031-9422(00)88523-4)

42. Schulte, A. E. et al. Purification and characterization of mevalonate kinase from suspension-cultured cells of *Catharanthus roseus* (L.) G. Don. Arch. Biochem. Biophys. 378: 287–298 (2000).
<https://doi.org/10.1006/abbi.2000.1779>
43. Alexandrov, N. N. et al. Insights into corn genes derived from large-scale cDNA sequencing. Plant Mol. Biol. 69: 179 (2009). <https://doi.org/10.1007/s11103-008-9415-4>
44. Chen, Q. et al. Molecular cloning, characterization, and functional analysis of acetyl-CoA C-acetyltransferase and mevalonate kinase genes involved in terpene trilactone biosynthesis from *Ginkgo biloba*. Molecules 22: 74 (2017). <https://doi.org/10.3390/molecules22010074>
45. Xiao, W. et al. Genome-wide identification, classification and expression analysis of GHMP genes family in *Arabidopsis thaliana*. Plant Syst. Evol. 301: 2125–2140 (2015).
<https://doi.org/10.1007/s00606-015-1219-9>
46. Sando, T. et al. Cloning and characterization of mevalonate pathway genes in a natural rubber producing plant, *Hevea brasiliensis*. Biotech. Bioch. 72: 2049–2060 (2008).
<https://doi.org/10.1271/bbb.80165>
47. Ma, Y. et al. Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in *Salvia miltiorrhiza*. J. Exp. Bot. 63: 2809–2823 (2012).
<https://doi.org/10.1093/jxb/err466>
48. Jones, C. G. et al. Isolation of cDNAs and functional characterisation of two multi-product terpene synthase enzymes from sandalwood, *Santalum album* Arch. Biochem. Biophys. 477: 121–130 (2008). <https://doi.org/10.1016/j.abi.2008.05.008>
49. Jones, C. G. et al. Sandalwood fragrance biosynthesis involves sesquiterpene synthases of both the terpene synthase (TPS)-a and TPS-b subfamilies, including santalene synthases. J. Biol. Chem. 286: 17445–17454 (2011). <https://doi.org/10.1074/jbc.m111.231787>
50. Rani, A. et al. Molecular regulation of santalol biosynthesis in *Santalum album* Gene 527: 642–648 (2013). <https://doi.org/10.1016/j.gene.2013.06.080>
51. Diaz-Chavez, M. L. et al. Biosynthesis of sandalwood oil: *Santalum album* CYP76F cytochromes P450 produce santalols and b PLoS ONE 9: e75053 (2013).
<https://doi.org/10.1371/journal.pone.0075053>
52. Celedon, J. M. et al. Heartwood-specific transcriptome and metabolite signatures of tropical sandalwood (*Santalum album*) reveal the final step of (*Z*)-santalol fragrance biosynthesis. Plant J. 86: 289–299 (2016). <https://doi.org/10.1111/tpj.13162>
53. Batth, T. S. et al. Targeted proteomics for metabolic pathway optimization Metabolic Engineering 13:194–203 (2011). <https://doi.org/10.1016/j.ymben.2010.12.005>
54. Zhang, X. H. et al. RNA-Seq analysis identifies key genes associated with haustorial development in the root hemiparasite *Santalum album*. Plant Sci. 6: 661 (2015).
<https://doi.org/10.3389/fpls.2015.00661>
55. Saitou, N. and Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406–425 (1987). <https://doi.org/10.1266/jjg.61.559>

56. Yoo, S. D. et al. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2:1565–1572 (2007). <https://doi.org/10.1038/nprot.2007.199>
57. Albers, E. and Larsson, C. A comparison of stress tolerance in YPD and industrial lignocellulose-based medium among industrial and laboratory yeast strains. *Journal of Industrial Microbiol.* 36:1085–1091 (2009). <https://doi.org/10.1007/s10295-009-0592-1>
58. Tao, T. et al. Molecular cloning, characterization, and functional analysis of a gene encoding 3-hydroxy-3-methylglutaryl-coenzyme A synthase from *Matricaria chamomilla*. *Genes Genom.* 38: 1179–1187 (2016). <https://doi.org/10.1007/s13258-016-0463-x>
59. Schmittgen, T. D. et al. Analyzing real-time PCR data by the comparative C (T) method. *Nature Protocols* 3:1101–1108 (2008). <https://doi.org/10.1038/nprot.2008.73>
60. Tsay, Y. H. and Robinson, G. W. Cloning and characterization of ERG8, an essential gene of *Saccharomyces cerevisiae* that encodes phosphomevalonate kinase. *Mol. Cell. Biol.* 11: 620–631 (1991). <https://doi.org/10.1128/MCB.11.2.620>
61. Xu, Y. et al. Cloning and sequence analysis of phosphomevalonate kinase gene (*McPMK*) from *Matricaria chamomilla*. *J. Curr. Res. Biosci. Plant Biol.* 5: 10–17 (2018). <https://doi.org/10.20546/ijcrbp.2018.502.002>
62. Romanowski, M. J. et al. Crystal structure of the *Streptococcus pneumoniae* phosphomevalonate kinase, a member of the GHMP kinase superfamily. *Proteins.* 47: 568–571 (2002). <https://doi.org/10.1002/prot.10118>
63. Doun, S. S. et al. *Enterococcus faecalis* phosphomevalonate kinase. *Protein Sci.* 14: 1134–1139 (2005). <https://doi.org/10.1110/ps.041210405>
64. Servouse, M. and Karst, F. Regulation of early enzymes of ergosterol biosynthesis in *Saccharomyces cerevisiae*. *J.* 240: 541–547 (1986). <https://doi.org/10.1042/bj2400541>
65. Pauwels, L. et al. Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* P. *Natl. Acad. Sci. USA* 105: 1380–1385 (2008). <https://doi.org/10.1073/pnas.0711203105>
66. Nagegowda, D. A. Plant volatile terpenoid metabolism: Biosynthetic genes, transcriptional regulation and subcellular compartmentation. *FEBS Lett.* 584: 2965–2973 (2010). <https://doi.org/10.1016/j.febslet.2010.05.045>
67. Degenhardt, J. et al. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70: 1621–1637 (2009). <https://doi.org/10.1016/j.phytochem.2009.07.030>
68. Lluch, M. A. et al. Molecular cloning and expression analysis of the mevalonate kinase gene from *Arabidopsis thaliana*. *Plant Mol. Biol.* 42: 365–376 (2000). <https://doi.org/10.1023/A:1006325630792>
69. Woo, H. M. et al. Application of targeted proteomics and biological parts assembly in *coli* to optimize the biosynthesis of an anti-malarial drug precursor, amorpho-4,11-diene. *Chem. Eng. Sci.* 103:21–28 (2013). <https://doi.org/10.1016/j.ces.2013.04.033>

70. Simkin, A. J. et al. Peroxisomal localisation of the final steps of the mevalonic acid pathway *in planta*. *Planta* 234: 903–914 (2011). <https://doi.org/10.1007/s00425-011-1444-6>
71. Fox, J. E. D. Sandalwood: the royal tree. *Biologist* 47: 31 (2000). PMID: 11190217.
72. Kim, Y. K. et al. Molecular cloning and characterization of mevalonic acid (MVA) pathway genes and triterpene accumulation in *Panax ginseng*. *J. Korean Soc. Appl. Bi.* 57: 289–295 (2014). <https://doi.org/10.1007/s13765-014-4008-1>
73. Cheong, J. J. and Choi, Y. D. Methyl jasmonate as a vital substance in plants. *Trends Genet.* 19: 409–413 (2003). [https://doi.org/10.1016/S0168-9525\(03\)00138-0](https://doi.org/10.1016/S0168-9525(03)00138-0)
74. Robert-Seilaniantz, A. et al. Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. *Annu. Rev. Phytopathol.* 49: 317–343 (2011). <https://doi.org/10.1146/annurev-phyto-073009-114447>
75. Farag, M. A. et al. Effect of oxylipins, terpenoid precursors and wounding on soft corals' secondary metabolism as analyzed via UPLC/MS and c Molecules 22:1–15 (2017). <https://doi.org/10.3390/molecules22122195>
76. Jiang, Y. et al. Regulation of floral terpenoid emission and biosynthesis in sweet basil (*Ocimum basilicum*). *J. Plant Growth Regul.* 35: 921–935 (2016). <https://doi.org/10.1007/s00344-016-9591-4>
77. Gao, W. et al. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. *Lett.* 11: 5170–173 (2009). <https://doi.org/10.1021/ol902051v>
78. Liao, Y. L. et al. Promoter analysis and transcriptional profiling of *Ginkgo biloba* 3-hydroxy-3-methylglutaryl coenzyme A reductase (GbHMGR) gene in abiotic stress r Not. Bot. Horti Agrobi. 43:25–34 (2015). <https://doi.org/10.15835/nbha4319416>
79. Zhang, X. H. et al. Identification and functional characterization of three new terpene synthase genes involved in chemical defense and abiotic stresses in *Santalum album*. *BMC Plant Biol.* 19:115 (2019). <https://doi.org/10.1186/s12870-019-1720-3>

Tables

Table 1
Primers used for related experiments in this study

Primer purpose	Primer name	Primer sequence(5'→3')
5' RACE	MK-5(1)/(2)	CGCAGATGAACCCAATCCAGAAC/CCTCGGGAATGTTCTGCTCTTCA
	PMK-5(1)/(2)	TGTCAGAGGGAGCCCACGTGCCT/AAGAGTACTGCACCGCTTGTCT
3' RACE	MK-3(1)/(2)	CGACAACACGGTCAGCACATA/AATGAATCAGGGGTTGCTCCA
	PMK-3(1)/(2)	ATGGCTGTAGTTGCTTCTGCTCC/GTGAACTAACATCTCCTCAGCTC
ORF	MK-O(F)/(R)	ATGGAGGTGAGGGCTCGAGCTC/TGAAGAACCACCGAGACAAATCT
	PMK-O(F)/(R)	ATGGCTGTAGTTGCTTCTGCTCC/TCCAATGTGAACTGAAGAAACAG
qRT-PCR	q-MK-F/R	GCTTCCTCTAGTTTCTGGATTG/CTGCGGTCCAAAGTTACTGTAT
	q-PMK-F/R	GACTGGCGGTTACCTCATTT/GCTTGACATCATCGTGAATTGG
Functional complementation	pYES2-MK-F	CCGGAATTCGAATGGAGGTGAGGGCTCGAGCTC
	pYES2-MK-R	TAAAGCGGCCGCGTGAAGAACCACCGAGACAAATCT
	pYES2-PMK-F	CCGGAATTCGAATGGCTGTAGTTGCTTCTGCTCC
	pYES2-PMK-R	TAAAGCGGCCGCGTCCAATGTGAACTGCTGAAG
Subcellular localization	YFP-MK-F	CCGGAATTCATGGAGGTGAGGGCTCGAGCTC
	YFP-MK-R	TCCCCGGGTGAAGAACCACCGAGACAAATCT
	YFP-PMK-F	CCGGAATTCATGGCTGTAGTTGCTTCTGCTCC
	YFP-PMK-R	CGCGGATCCATCCAATGTGAACTGAAGAAACA

Table 2
Nucleotide sequences of *SaMK* and *SaPMK* and similarity to genes from other plant species

a: *SaMK*

Species	Accession number	Identity (%)
<i>Hevea brasiliensis</i>	JN036543.1	76
<i>Morus alba</i>	KX387386.1	75
<i>Platycodon grandiflorus</i>	KC439364.1	75
<i>Panax ginseng</i>	JQ957844.1	74
<i>Catharanthus roseus</i>	HM462019.1	73

Table 2
b: *PMK*

Species	Accession number	Identity (%)
<i>Tripterygium wilfordii</i>	KR260990.1	76
<i>Morus alba</i>	KX387387.1	76
<i>Hevea brasiliensis</i>	JN036535.1	76
<i>Panax ginseng</i>	KJ804170.1	74
<i>Siraitia grosvenorii</i>	HQ128558.1	74

Table 3
Physicochemical properties of deduced proteins in *Santalum album*

Proteins	Molecular weight (kD)	Theoretical isoelectric point	Number of acidic amino acids	Number of basic amino acids	Instability index	Aliphatic index	Total average hydropathicity
SaMK	41.3012	5.23	42	33	32.42	99.02	0.113
SaPMK	54.5954	5.92	52	46	32.87	89.84	-0.073

Figures

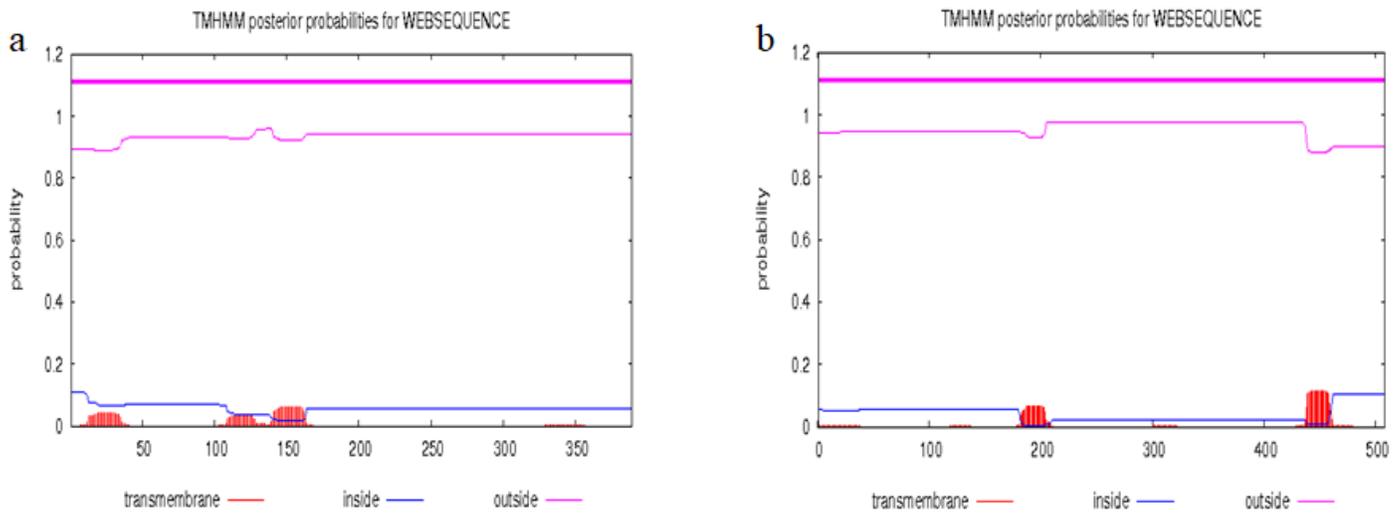


Figure 1

Prediction of transmembrane domains of SaMK (a) and SaPMK (b) proteins in *Santalum album*.

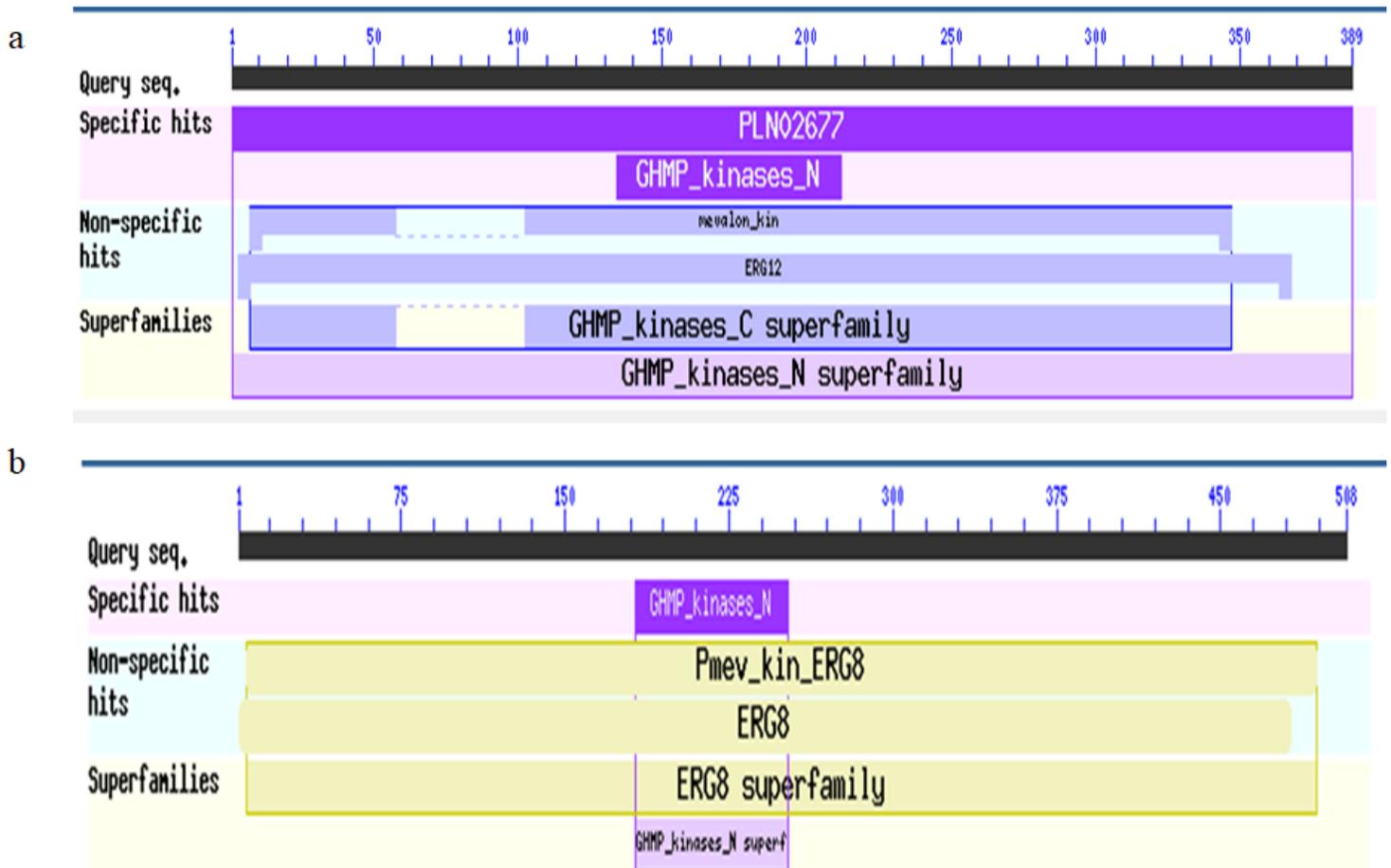
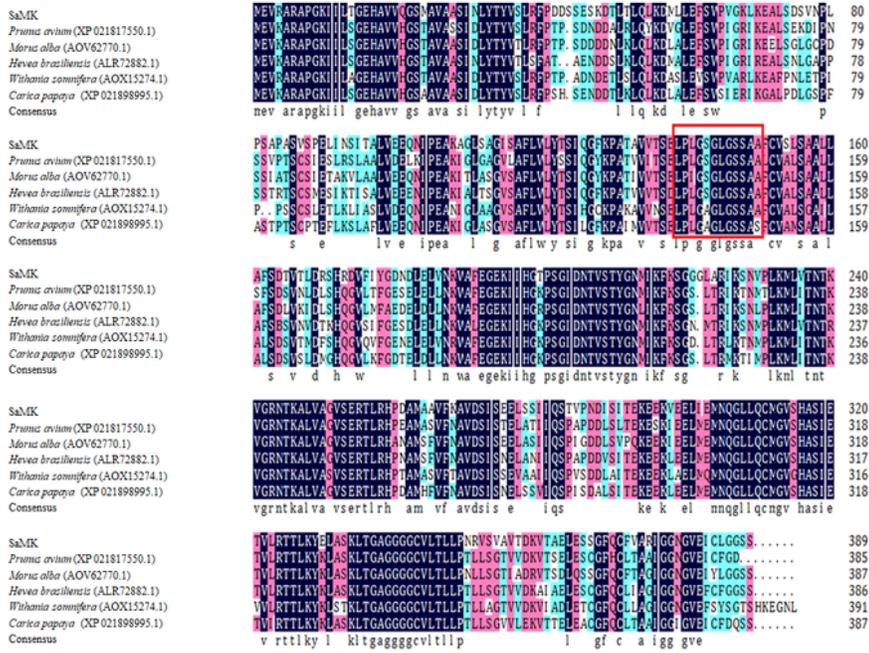


Figure 2

Conserved domains of SaMK (a) and SaPMK (b) proteins in *Santalum album*.

a



b

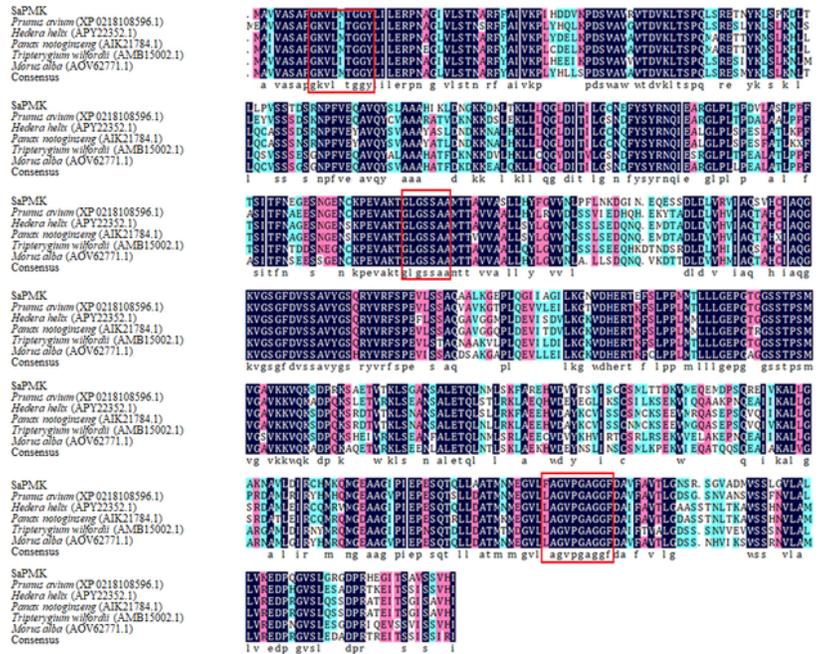


Figure 3

Multiple alignments of SaMK (a) and SaPMK (b) deduced amino acid sequences with other corresponding homologous proteins. The red rectangles marked the conserved motifs.

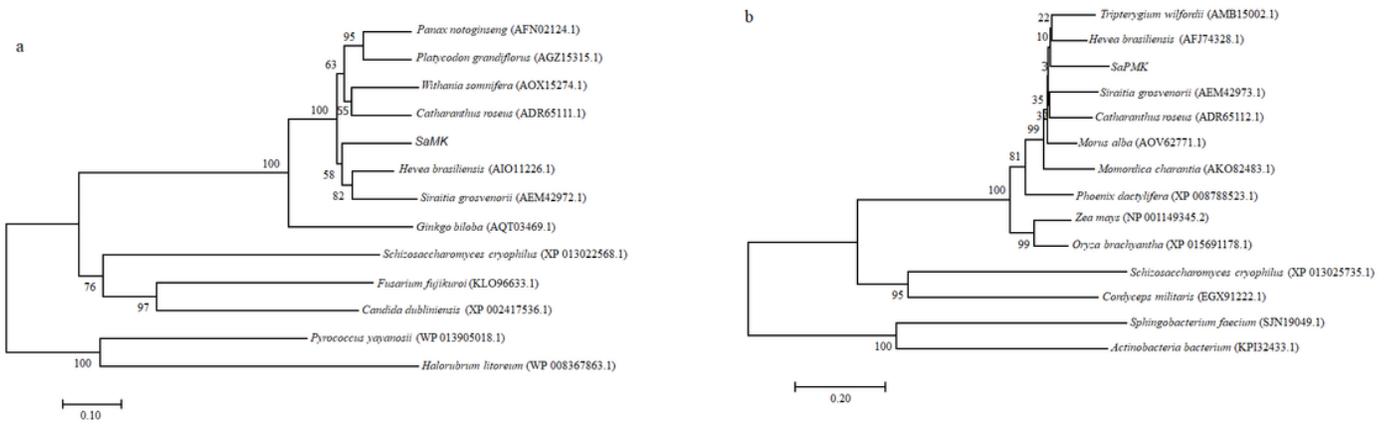


Figure 4

Phylogenetic trees of SaMK and SaPMK proteins. (a) SaMK; (b) SaPMK.

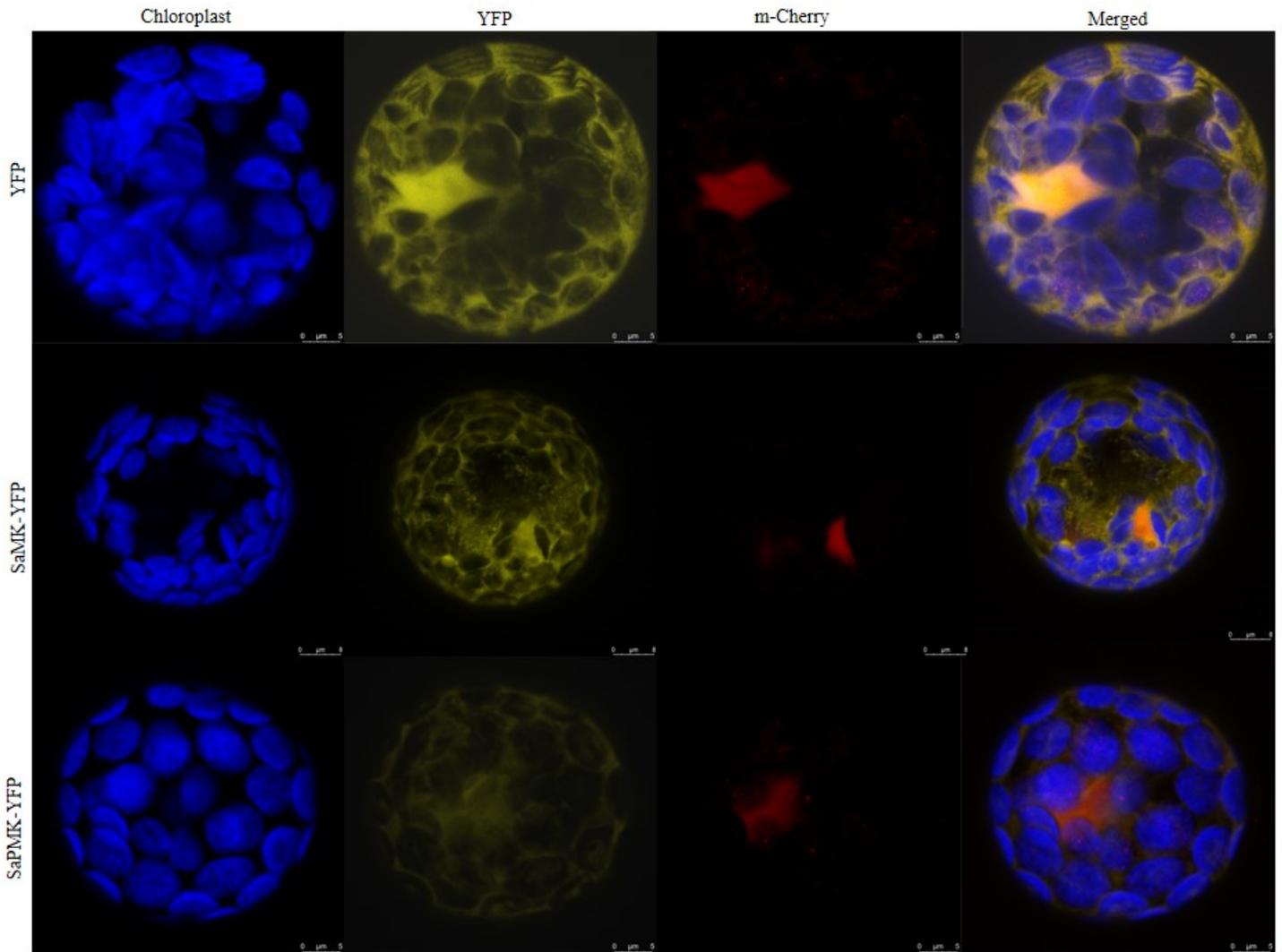


Figure 5

Subcellular localization of SaMK and SaPMK. Blue fluorescence indicates chlorophyll (Chl) autofluorescence, Yellow fluorescence indicates YFP and fusion vectors signal and red signal indicates m-Cherry fluorescence. The merged images represent a digital combination of Chl autofluorescence, YFP fluorescent and m-Cherry protein fluorescence. Fluorescence was excited for YFP at 514 nm, for Chl at 543 nm and for m-Cherry at 587 nm. Scale bar of SaPMK-YFP and YFP= 5 μ m and scale bar of SaMK-YFP = 8 μ m.

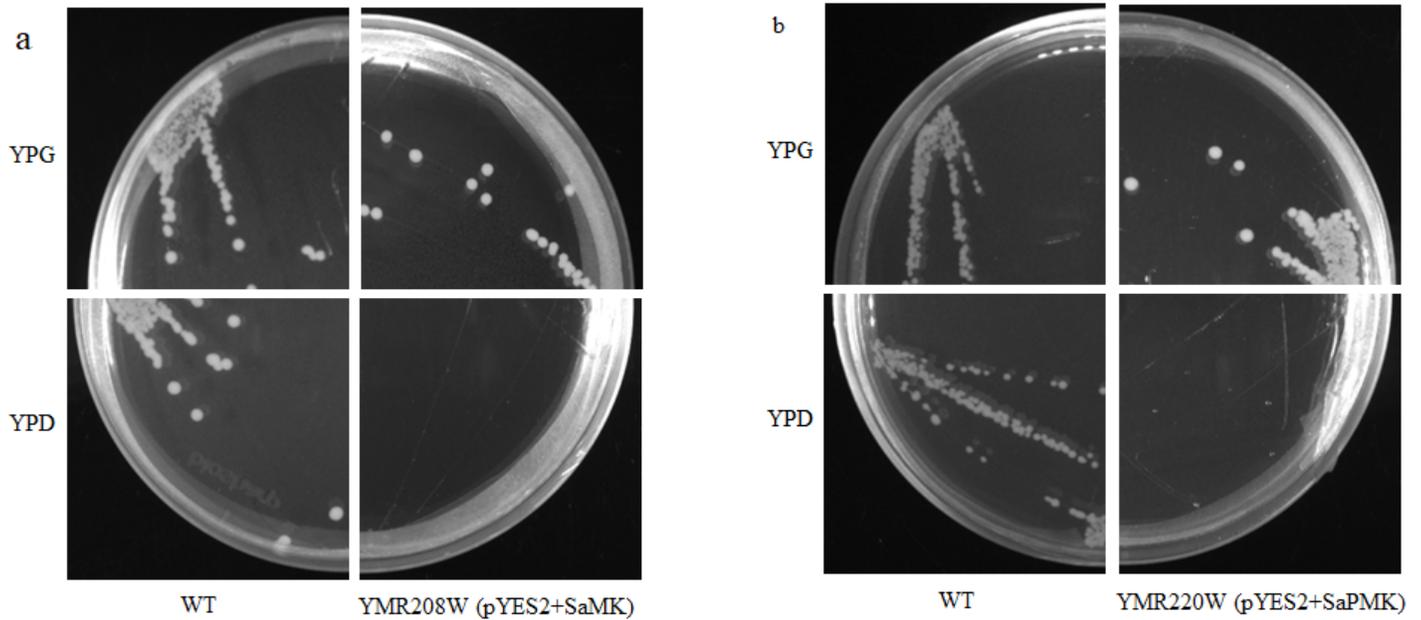


Figure 6

Functional complementation of SaMK (a) and SaPMK (b) genes.

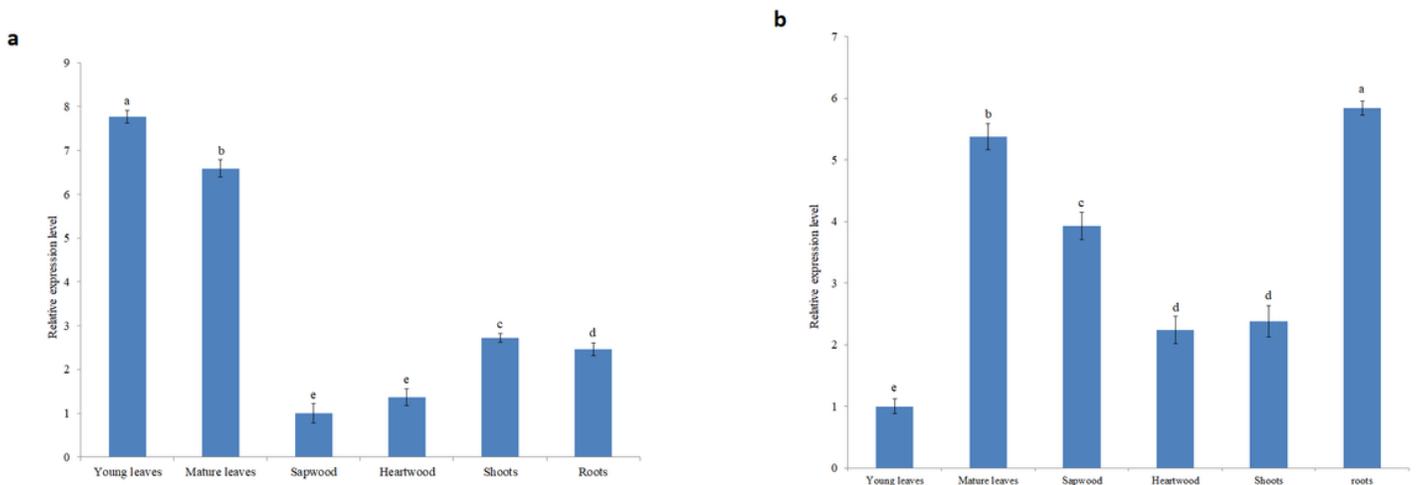


Figure 7

Tissue expression analysis of SaMK (a) and SaPMK (b) genes. The gene expression level of SaMK and SaPMK in young leaves was set to 1. Data from qRT-PCR are means \pm SD (standard deviation) from triplicate experiments (n = 3). Different letters indicate significant differences ($p < 0.05$) according to Duncan's multiple range test.

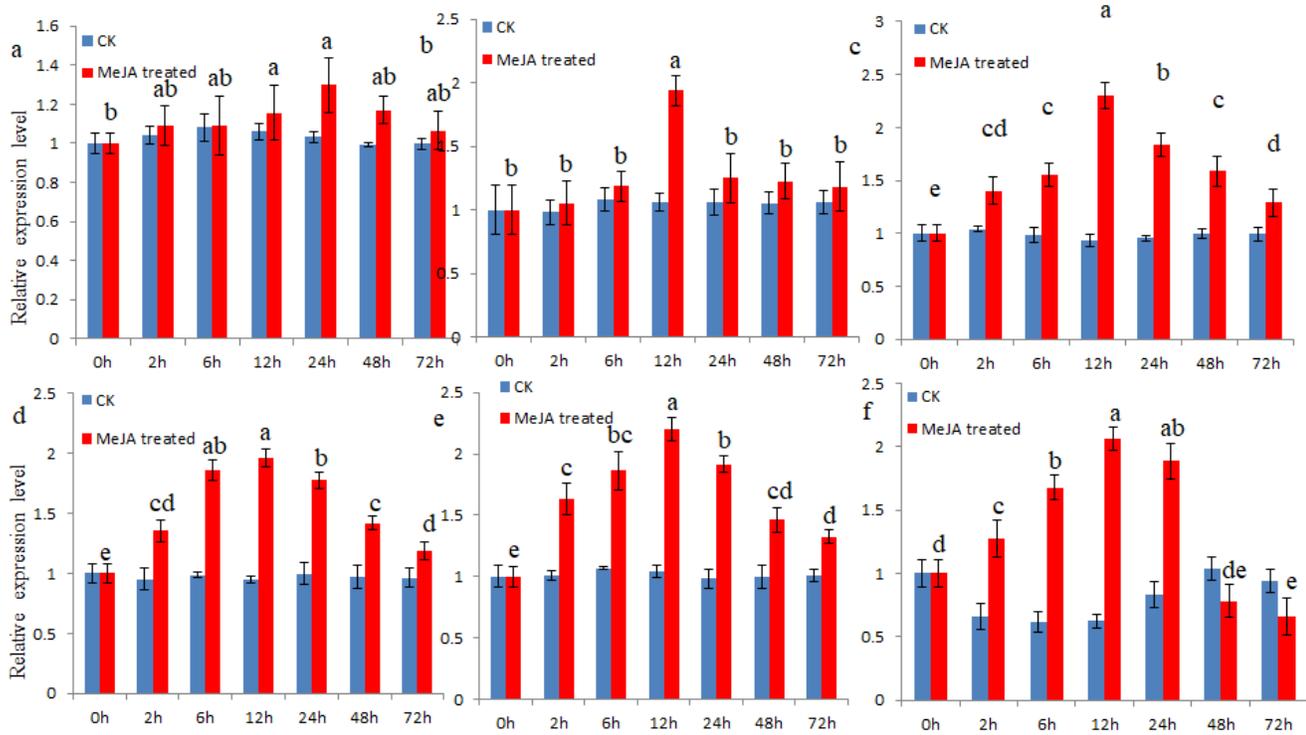


Figure 8

Tissue expression analysis of SaMK (a) and SaPMK (b) genes. The gene expression level of SaMK and SaPMK in young leaves was set to 1. Data from qRT-PCR are means \pm SD (standard deviation) from triplicate experiments (n = 3). Different letters indicate significant differences ($p < 0.05$) according to Duncan's multiple range test.