scientific reports

Check for updates

OPEN

# Cloning, characterization, and functional analysis of acetyl-CoA C-acetyltransferase and 3-hydroxy-3-methylg lutaryl-CoA synthase genes in *Santalum album*

Meiyun Niu[1,2], Haifeng Yan[3], Yuping Xiong[1], Yueya Zhang[1,2], Xinhua Zhang[1], Yuan Li[1], Jaime A. Teixeira da Silva[4] & Guohua Ma[1]✉

**Sandalwood (*Santalum album* L.) is famous for its unique fragrance derived from the essential oil of heartwood, whose major components are santalols. To understand the mechanism underlying the biosynthesis of santalols, in this study, we cloned two related genes involved in the mevalonate pathway in *S. album* coding for acetyl-CoA C-acetyl transferase (AACT) and 3-hydroxy-3-methyglutary-CoA synthase (HMGS). These genes were characterized and functionally analyzed, and their expression profiles were also assessed. An AACT gene designated as *SaAACT* (GenBank accession No. MH018694) and a HMGS gene designated as *SaHMGS* (GenBank accession No. MH018695) were successfully cloned from *S. album*. The deduced SaAACT and SaHMGS proteins contain 415 and 470 amino acids, and the corresponding size of their open-reading frames is 1538 bp and 1807 bp, respectively. Phylogenetic trees showed that the SaAACT protein had the closest relationship with AACT from *Hevea brasiliensis* and the SaHMGS proteins had the highest homology with HMGS from *Siraitia grosvenorii*. Functional complementation of SaAACT and SaHMGS in a mutant yeast strain deficient in these proteins confirmed that *SaAACT* and *SaHMGS* cDNA encodes functional SaAACT and SaHMGS that mediate mevalonate biosynthesis in yeast. Tissue-specific expression analysis revealed that both genes were constitutively expressed in all examined tissues (roots, sapwood, heartwood, young leaves, mature leaves and shoots) of *S. album*, both genes showing highest expression in roots. After *S. album* seedlings were treated with 100 µM methyl jasmonate, the expression levels of *SaAACT* and *SaHMGS* genes increased, suggesting that these genes were responsive to this elicitor. These studies provide insight that would allow further analysis of the role of genes related to the sandalwood mevalonate pathway in the regulation of biosynthesis of sandalwood terpenoids and a deeper understanding of the molecular mechanism of santalol biosynthesis.**

Indian sandalwood (*Santalum album* L.) is a well-studied small hemi-parasitic evergreen tree that is widely distributed in tropical and subtropical regions of India, Indonesia, Australia and the Pacific Islands[1,2]. The scented heartwood of sandalwood and essential oil extracted from it have considerable economic value. Sandalwood is widely used in medicine[3], wooden handicraft[4] and cosmetics[5]. The value of sandalwood mainly depends on the proportion of heartwood and on the concentration and quality of essential oil extracted from the heartwood[6,7]. Indian sandalwood is considered to be one of the most valuable sandalwood species, usually yielding 3–8% essential oil[8]. Over 85% of *S. album* essential oil consists of sesquiterpenoids, including sesquiterpene

[1]Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, The Chinese Academy of Sciences, Guangzhou 510650, China. [2]University of Chinese Academy of Sciences, Beijing 100039, China. [3]Cash Crop Institute of Guangxi Academy of Agricultural Sciences, Nanning 30007, China. [4]Miki-cho Post Office, Miki-cho, Ikenobe, 3011-2, P.O. Box 7, Kagawa-ken 761-0799, Japan. ✉email: magh@scib.ac.cn

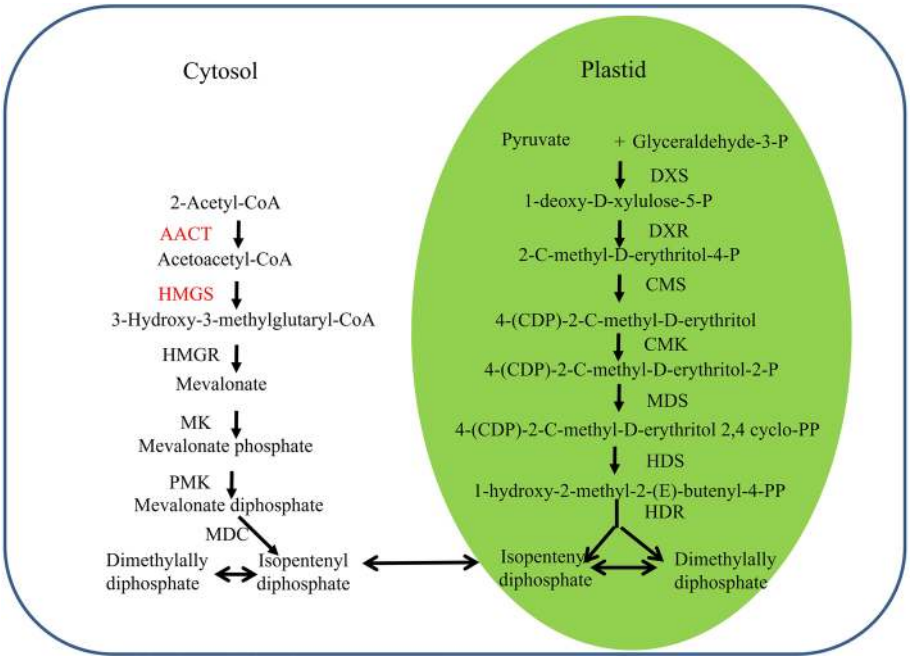


**Figure 1.** Pathway of isoprenoids biosynthesis in plants (DXS, 1-deoxy-Dxylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; CMS, 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase; CMK, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MDS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS, 4-hydroxy-3-methylbut-2-enyl diphosphate synthase; HDR, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase). The Figure was drawn by ourselves after referring to previous studies (Laule et al. 2003; Miziorko et al. 2011; Tholl et al. 2015) and tagged the genes marked red were studied.

alcohols (α-santalol, β-santalol, *trans*-α-bergamotol and epi-β-santalol) and their corresponding sesquiterpenes (α-santalene, β-santalene, *trans*-α-bergamotene and epi-β-santalene)[8–12].

Like all other terpenoids, sesquiterpenoids are derived from two building blocks, isopentenyl diphosphate (IPP) and its allylic isomer, dimethylallyl diphosphate (DMAPP). IPP and DMAPP are formed via two pathways. One is the mevalonate (MVA) pathway, which is located in the cytoplasm, and the other is the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway, which is located in plastids (Fig. 1)[13–16]. The MEP pathway is mainly involved in the biosynthesis of monoterpenoids, diterpenoids and other terpenoids such as hormones, plant pigments, and ubiquitins. The MVA pathway chiefly synthesizes sesquiterpenoids and triterpenoids, for example sterol, as well as other polyterpenoids[17–19]. Since the main components in sandalwood essential oils are sesquiterpenoids, the separation and identification of genes coding for MVA pathway-related enzymes and their functions are essential for further studies of the molecular mechanism of heartwood formation.

The MVA pathway generates IPP and DMAPP in six steps, as shown in Fig. 1[20–22]. Acetyl coenzyme A (CoA) acetyltransferase (AACT) catalyzes the condensation of two molecules of acetyl-CoA to form acetoacetyl-CoA through Claisen condensation, and AACT is a regulatory enzyme involved in the biosynthesis of terpenoids under salt stress, promoting the production of squalene[23]. 3-Hydroxy-3-methyglutaryl-CoA synthase (HMGS) transforms acetyl-CoA and acetoacetyl-CoA to 3-hydroxy-3-methylglutaryl-CoA[24], which is further converted into MVA by 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR). The analysis of HMGS in *S. album* is important to investigate the biosynthesis of α-, β-, epi-β-santalols and bergamotol. In view of their importance in the biosynthesis of terpenoids, the genes related to the MVA pathway have been successfully cloned from a number of plants, such as *Arabidopsis thaliana*[20,25–28,30], *Salvia miltiorrhiza*[31], *Aquilaria sinensis*[32], *Hevea brasiliensis*[33–35] and *Ginkgo biloba*[36–38]. In contrast, research related to sandalwood terpenoids has mainly focused on terpene synthase, which cyclizes farnesyl pyrophosphate into various cyclic sesquiterpenes such as endo-bergamotene, α-santalene, epi-β-santalene and β-santalene[39–42], and cytochrome P450 oxygenase, which oxidizes α-, β-, epi-β-santalene and bergamotene to produce α-, β-, epi-β-santalols and bergamotol, respectively[43,44]. To identify the spatial patterns of sesquiterpenoid biosynthesis and to clone genes that encode the enzymes involved in sesquiterpene biosynthesis, transcriptomic analyses of *S. album* have been performed[42]. Srivastava et al. described the cloning and functional characterization of five genes encoding two sesquisabinene synthases (SaSQS1, SaSQS2), bisabolene synthase (SaBS), santalene synthase (SaSS), and farnesyl diphosphate synthase (SaFDS). There are no reports on the upstream genes of the MVA pathway in the biosynthesis of sesquterpenoids in the genus *Santalum*, including *S. album*.

In the present study, we cloned two related genes, named as *SaAACT* and *SaHMGS*, which are involved in the MVA pathway, by RACE technology. This, the first such study for Indian sandalwood, provides more detailed insight into the molecular biology of the santalol biosynthetic pathway. The structure and function of *SaAACT* and *SaHMGS* were analyzed by bioinformatics analyses, functional complementation was performed in yeast, and

| Proteins | Molecular weight (kD) | Theoretical isoelectric point | Number of acidic amino acids | Number of basic amino acids | Instability index | Aliphatic index | Total average hydropathicity |
|---|---|---|---|---|---|---|---|
| SaAACT | 43.1326 | 8.71 | 37 | 42 | 32.06 | 95.01 | 0.139 |
| SaHMGS | 51.9982 | 6.29 | 54 | 50 | 37.79 | 75.13 | -0.242 |

**Table 1.** Physicochemical properties of deduced proteins in *Santalum album*.

| a: SaAACT | | |
|---|---|---|
| **Species** | **Accession number** | **Identity (%)** |
| *Camellia oleifera* | GU594059.1 | 80 |
| *Catharanthus roseus* | JF739870.1 | 80 |
| *Populus trichocarpa* | XM002308719.2 | 79 |
| *Euphorbia helioscopia* | KP995935.1 | 79 |
| *Hevea brasiliensis* | JN036531.1 | 78 |
| **b: SaHMGS** | | |
| **Species** | **Accession number** | **Identity (%)** |
| *Camellia sinensis* | JQ390224.1 | 82 |
| *Panax ginseng* | GU565098.1 | 82 |
| *Hevea brasiliensis* | JN036533.1 | 82 |
| *Siraitia grosvenorii* | HQ128555.1 | 82 |
| *Platycodon grandiflorus* | KC439366.1 | 81 |

**Table 2.** Nucleotide sequences of target genes and similarity to genes from other plant species.

their expression level in tissues was analyzed by fluorescence quantitative PCR. These studies provide guidance for further analyses of the roles of genes related to the sandalwood MVA pathway in the regulation of biosynthesis of sandalwood terpenoids and a deeper understanding of the molecular mechanism of santalol biosynthesis.

## Results

**Molecular cloning and characterization of the cDNA of *SaAACT* and *SaHMGS*.** After sequencing PCR products, it was shown that full-length *SaAACT* is 1538 bp long and contains a 1248 bp open reading frame (ORF) that encodes 415 deduced amino acid residues while full-length *SaHMGS* is 1807 bp long and contains a 1413 bp ORF that encodes 470 deduced amino acid residues. A BLASTn search of *SaAACT* and *SaHMGS* with other plant species showed that both genes are highly homologous to *AACT* and *HMGS* genes from other plant species. *SaAACT* nucleotide sequences showed 80%, 80%, 79%, 79%, and 79% identity with *Camellia oleifera*, *Catharanthus roseus*, *Populus trichocarpa*, *Euphorbia helioscopia* and *H. brasiliensis*, respectively (Table 1). *SaHMGS* nucleotide sequences showed 82% identity with *C. sinensis*, *Panax ginseng*, *H. brasiliensis* and *Siraitia grosvenorii* and 81% similarity with *Platycodon grandiflorus* (Table 1). These results reveal that *SaAACT* and *SaHMGS* belong to the *AACT* and *HMGS* gene families, respectively. Therefore, these genes were designated as *SaAACT* (GenBank accession No. MH018694) and *SaHMGS* (GenBank accession No. MH018695).

**Bioinformatics analysis of the deduced SaAACT and SaHMGS proteins.** The relative molecular weight, theoretical isoelectric point, instability index, aliphatic index and grand average of hydropathicity of the deduced SaAACT and SaHMGS proteins, which were predicted by ExPASy, are shown in Table 2. The relative molecular weight of SaAACT is 43 kDa, similar to previously reported AACT[38], which had a molecular weight of 41.5 kDa. The relative molecular weight of SaHMGS is 52 kDa, similar to previously reported HMGS[45], with a molecular weight of 53 kDa. The theoretical isoelectric point of SaAACT and SaHMGS are 8.71 and 6.29, respectively. SaAACT is a stable hydrophobic protein and SaHMGS is a stable hydrophilic protein. Well, SaAACT and SaHMGS have no transmembrane domain or signal peptide. SaAACT has a thiolase (like) domain from aa 19-403 and acetyl-CoA acyltransferase activity from aa 15-406 (Fig. 2a). Residues of two cystines (Cys101, Cys391) and one histidine (His361) are present in SaAACT[46], and these are highly conserved in AACT among the thiolases from different sources, and are important for its catalytic activity (Fig. 3A, marked with an asterisk). One highly conserved domain (NVHGGAVSIGHPIGCSG) (aa 351-367) at the C-terminal end (Fig. 3a, marked with a red box), which is also present in SaAACT, is a characteristic sequence of thiolase. SaHMGS has significant HMGS activity from aa 4-407 (Fig. 2b). Residues of one cystine (Cys119), one tyrosine (Tyr296) and one aspartic acid (Asn329) are present in SaHMGS[47,48]. They are highly conserved in HMGS from different sources, and are important for its catalytic activity (Fig. 3b, marked with an asterisk). One conserved motif (NxD/NE/VEGI/VDx (2) NACF/YxG) (aa 108-123), which was present in SaHMGS, is a characteristic sequence of HMGS[37,49]. These findings confirm that SaAACT and SaHMGS have similar catalytic functions to the corresponding AACT and HMGS from other plants.

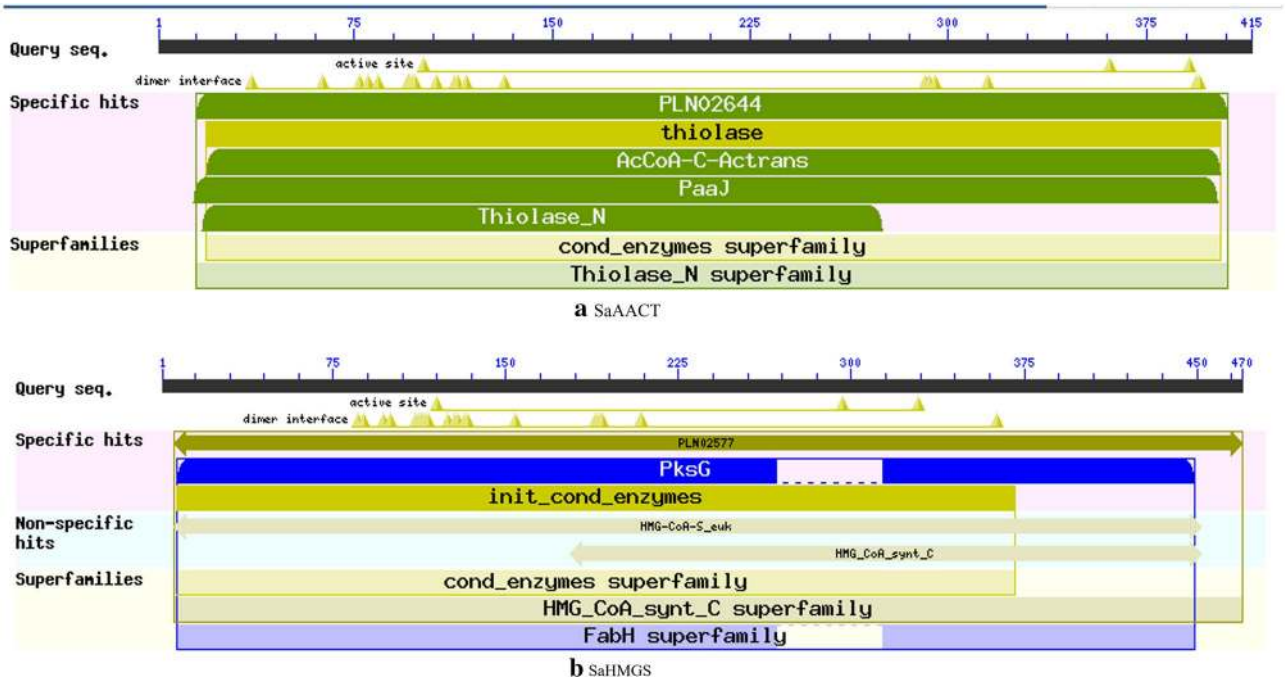


**Figure 2.** Conserved domain of SaAACT (**a**) and SaHMGS (**b**) in *S. album*. Which was predicted by the Conserved Domains database in NCBI using my own gene sequences.

**Phylogenetic analysis of the deduced SaAACT and SaHMGS proteins.** To better understand the evolutionary relationships between deduced SaAACT and SaHMGS proteins with other AACTs and HMGSs from angiosperms, gymnosperms, fungi, and bacteria, MEGA 7 was used to construct two phylogenetic trees with the neighbor-joining (NJ) method. The first phylogenetic tree revealed that angiosperm AACTs were clustered in one group where SaAACT exhibited the highest homology with AACT from *H. brasiliensis*, while AACT of *G. biloba* formed a distinct group of gymnosperms, and the AACT gene from fungi and bacteria were clustered as a different group (Fig. 4a). The second phylogenetic tree indicated that angiosperm HMGSs were clustered in one group where SaHMGS exhibited the closest homology with HMGS from *S. grosvenorii*, while HMGS of *Pinus sylvestris* and *Taxus* × *media* formed a distinct group of gymnosperms and the *AACT* gene from fungi and bacteria were clustered as a separate group (Fig. 4b). These results suggest that SaAACT and SaHMGS share a common evolutionary base with other plant AACT and HMGS proteins based on their conserved structure and sequence characteristics.

**Sub-cellular localization of SaAACT and SaHMGS proteins.** Predicted sub-cellular localization of SaAACT and SaHMGS proteins by PSORT showed that SaAACT has the highest probability scores for peroxisomes (0.78), followed by the cytosol (0.11) and mitochondria (0.11) whereas SaHMGS exists mainly in the cytosol (0.61) but is also distributed in the nucleus (0.22) and mitochondria (0.17). To further verify the sub-cellular localization of SaAACT and SaHMGS protein, sub-cellular localization of SaAACT-YFP (yellow fluorescent protein) and SaHMGS-YFP were studied using a modified polyethylene glycol method to transform *Arabidopsis thaliana* protoplasts with SaAACT-YFP and SaHMGS-YFP constructs. In *A. thaliana*, AACT2, which is involved in the MVA pathway, is localized in the cytosol and nucleus whereas AACT1, which may be involved in fatty acid degradation, is located in peroxisomes[50,51]. We found that SaAACT proteins were located in the nucleus (Fig. 5). Sub-cellular localization of SaHMGS-YFP showed that SaHMGS proteins were located in the cytosol (Fig. 5), like BjHMGS1 of *Brassica juncea*[52]. Our results suggest that SaAACT and SaHMGS cloned in this study may be involved in MVA pathway in *S. album*.

**Functional complementation of *SaAACT* and *SaHMGS* in yeast.** In yeast, the MVA pathway is a biosynthetic pathway that is essential for survival, and disrupting MVA pathway genes in yeast strains can be fatal[53,54]. The disrupted strains with empty pYES2 could not grow on either YPG expression medium or YPD non-expression medium (Fig. 6a). YPL028W harbored pYES2-*SaAACT* and YML126C harbored pYES2-*SaHMGS*, which grew well on YPG medium. However, neither YPL028W, which harbored pYES2-*SaAACT*, nor YML126C, which harbored pYES2-*SaHMGS*, could grow on YPD medium (Fig. 6b). These results indicate that *SaAACT* and *SaHMGS* have AACT and HMGS activity, respectively.

**Tissue-specific and MeJA treatment expression analysis of *SaAACT* and *SaHMGS*.** To explore the tissue-specific expression pattern of *SaAACT* and *SaHMGS* genes in *S. album*, total RNA was isolated from different tissues, including roots, heartwood, sapwood, young leaves, mature leaves and shoots. quantitative real-

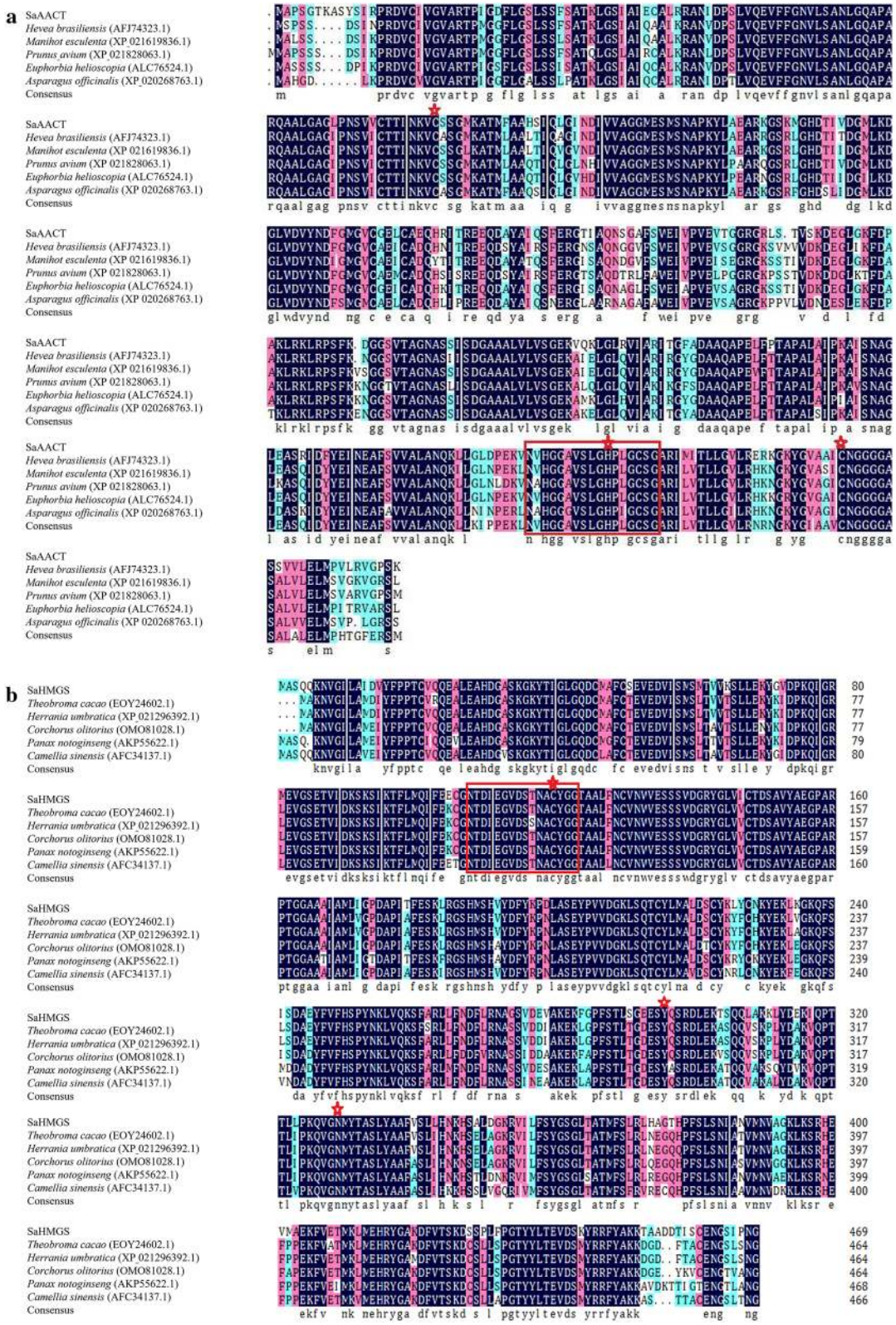


**Figure 3.** Multiple alignments of SaAACT (**a**) and SaHMGS (**b**) deduced amino acid sequences with other corresponding homologous proteins. The figures were a comparison diagram and cluster analysis of my own gene sequences with related sequences of other species by CLUSTALX 2.0 and MEGA.

time PCR (qRT-PCR) was also performed. The results of qRT-PCR are shown in Fig. 7. *SaAACT* and *SaHMGS* were constitutively expressed in all tissues. The *SaAACT* gene was differentially expressed in various tissues and exhibited the highest level of expression in roots, followed by mature leaves and heartwood, approximately

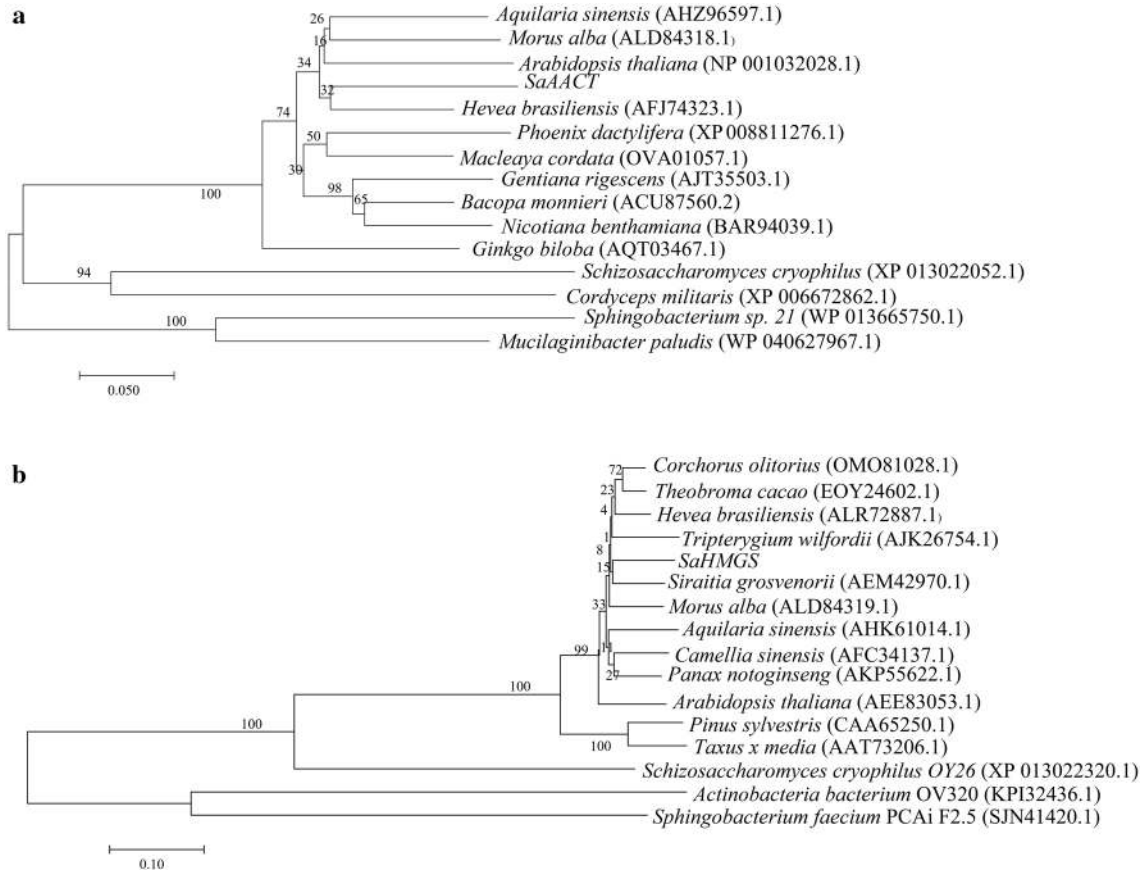


**Figure 4.** Phylogenetic analysis of target *S. album* proteins (**a**: SaAACT; **b**: SaHMGS).

13.44-, 3.66- and 3.55-fold higher than young leaves, respectively. A similar expression pattern of the *AACT* gene was found in *Bacopa monnieri* and *G. biloba*, in which the *BmAACT* and *GbAACT* genes were highly expressed in roots, followed by stems and leaves[38,55]. As shown in Fig. 7, *SaHMGS* was highly expressed in roots, approximately 4.22-fold more than in young leaves. The expression level of *SaHMGS* in mature leaves, sapwood, heartwood, shoots and young leaves showed few differences. In contrast, in *Taxus* × *media*, *TmHMGS* was expressed in needles and stems at a similar level, but no expression was detected in roots[45].

Methyl jasmonate (MeJA) is a small signaling molecule that can regulate secondary metabolism in plants when applied exogenously[56]. MeJA is closely related to terpene metabolism[57]. MeJA treatment can effectively induce the expression of *AACT* and *HMGS* genes and terpenoid biosynthesis in *S. miltiorrhiza*[58], *G. biloba*[38] and *Tripterygium wilfordii*[59]. In this study, we measured the transcript level of *SaAACT* and *SaHMGS* in *S. album* roots, shoots and leaves after treatment with 100 μM MeJA (Fig. 8). *SaAACT* and *SaHMGS* expression were significantly induced by MeJA. The trend of the change in transcript level of *SaAACT* in *S. album* roots, shoots and leaves after MeJA treatment was consistent, all increasing gradually and peaking at 24 h compared with control seedlings in which *SaAACT* transcript level decreased slowly after MeJA treatment. With regard to the *SaHMGS* gene, the trends in the change of transcript level of *SaAACT* in *S. album* roots, shoots and leaves after MeJA treatment were different. In roots, expression level increased rapidly after seedlings were treated with MeJA for 6 h, then gradually decreased. In shoots, expression increased slowly then peaked at 48 h after MeJA treatment. In leaves, expression level gradually increased and peaked at 12 h after MeJA treatment, then gradually decreased.

## Discussion

The commercial value of sandalwood lies mainly in its fragrant heartwood. The final purpose of planting sandalwood is to harvest high quality heartwood in great quantities. However, slow growth rates, susceptibility to diseases and variation in sandalwood oil yield hamper sandalwood production[43]. Researchers have attempted to synthesize santalol by chemical approaches[60–62], but yield is very low and industrial production is not economic. In 1998, sandalwood was listed on the red list of threatened plants as a vulnerable plant by the International Union for Conservation Nature and Natural Resources[63]. Currently there are no reports for *S. album* on the genes upstream of the MVA pathway involved in sesquiterpenoid biosynthesis.

This study is the first to clone and characterize the *SaAACT* and *SaHMGS* genes encoding AACT and HMGS of the MVA pathway and analyze their tissue expression levels. A 1538 bp full-length cDNA of the *SaAACT* gene was isolated from *S. album*. The deduced SaAACT protein contained 415 amino acids and weighed 43 kDa. In addition, an 1807 bp full-length cDNA of the *SaHMGS* gene was isolated from *S. album*. The deduced SaHMGS

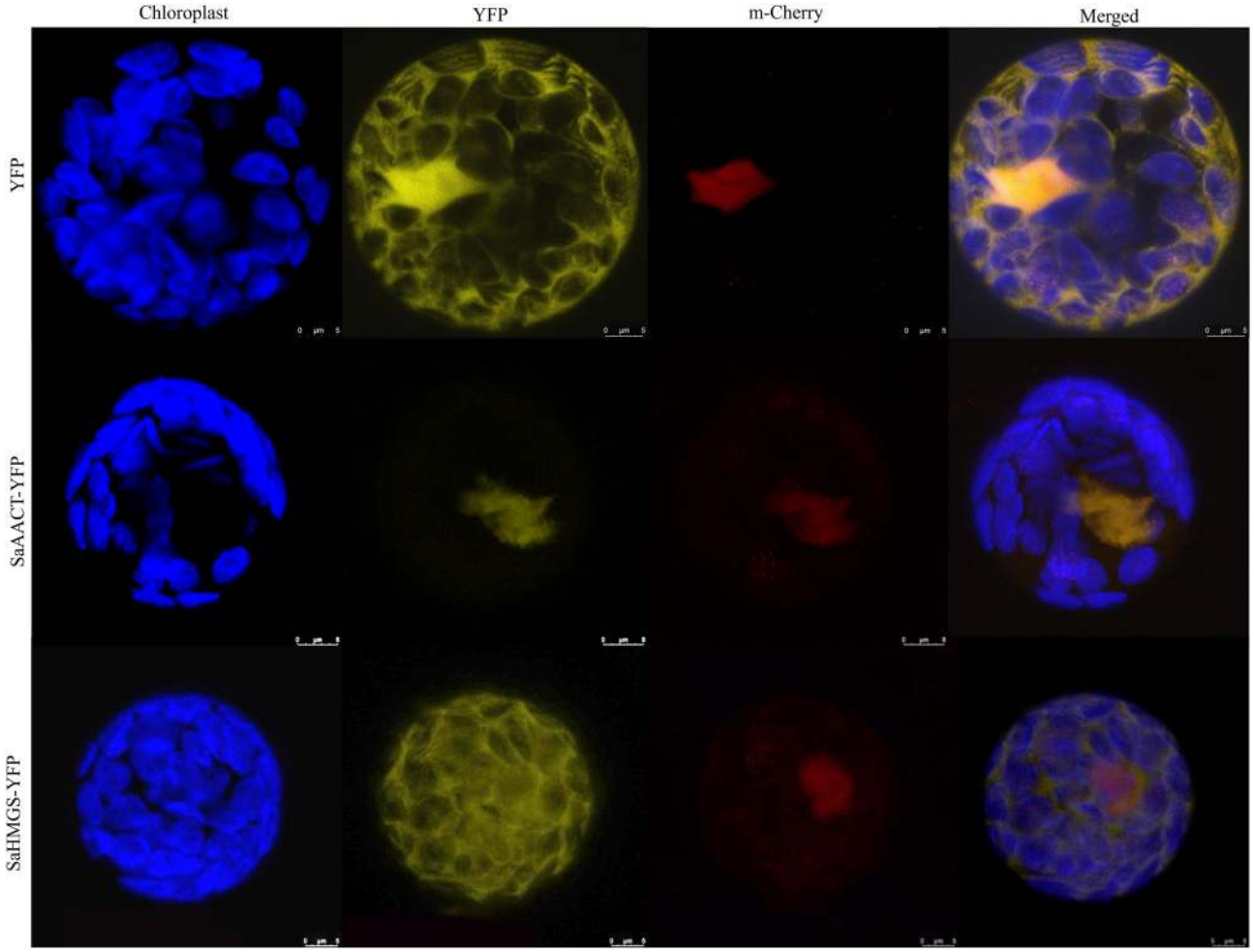


**Figure 5.** Subcellular localization of SaAACT and SaHMGS proteins in *A. thaliana* mesophyll protoplasts. Yellow fluorescence indicates SaAACT-YFP and SaHMGS-YFP fusion protein signal. Blue signal indicates chlorophyll (Chl) autofluorescence and red signal indicates m-Cherry fluorescence. The merged images represent a digital combination of Chl autofluorescence, YFP fluorescent and m-Cherry protein fluorescence images. Fluorescence was excited for YFP at 514 nm, for Chl at 543 nm and for m-Cherry at 587 nm. Scale bar of SaHMGS-YFP and YFP = 5 μm and scale bar of SaAACT-YFP = 8 μm.

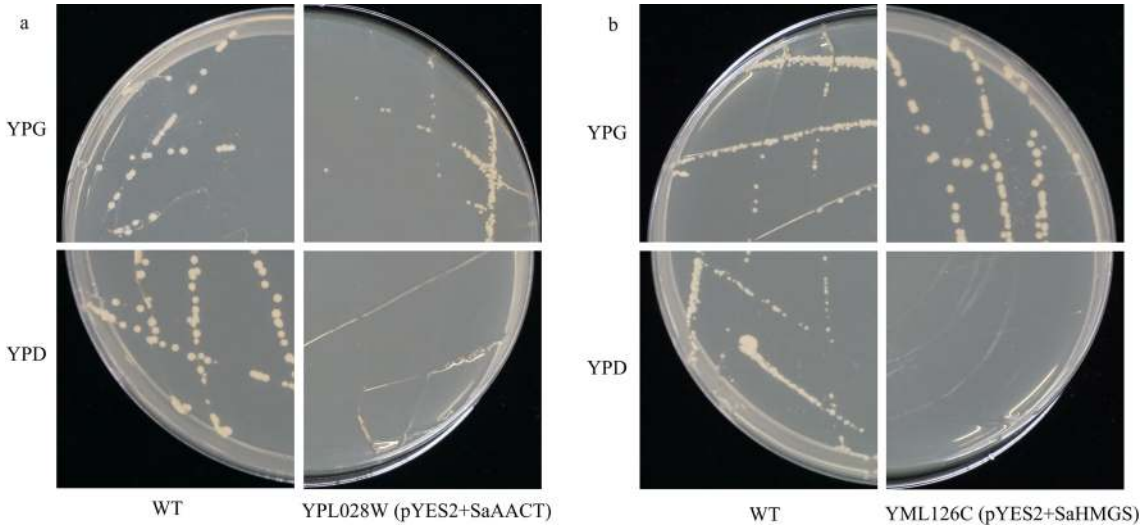


**Figure 6.** Functional complementation of two *Santalum album* genes (**a**: *SaAACT*; **b**: *SaHMGS*) in yeast (*Saccharomyces cerevisiae*). WT, wild type.

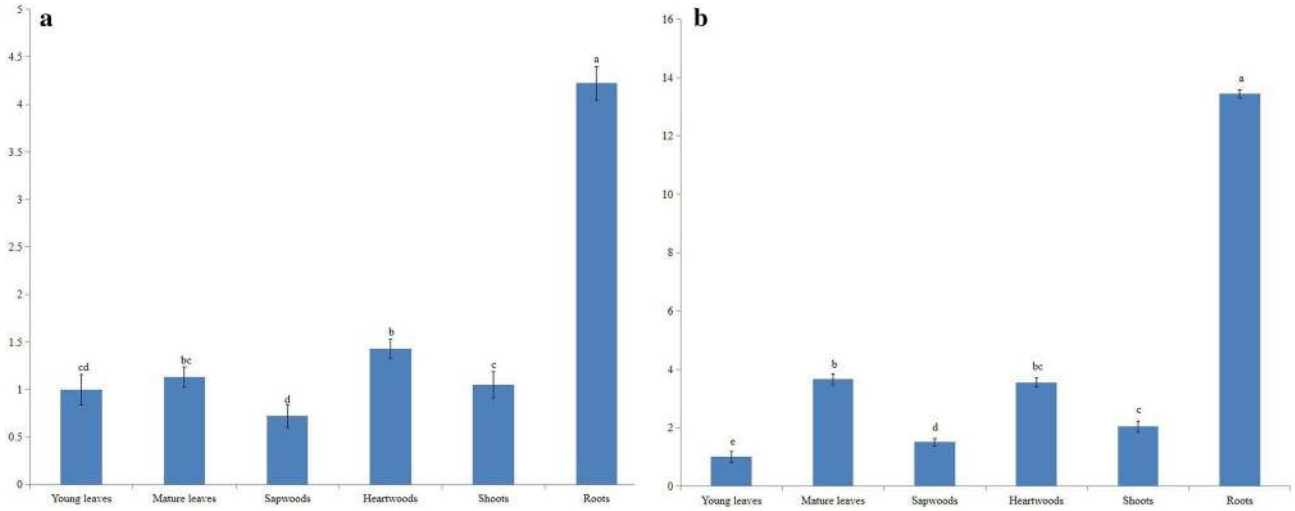


**Figure 7.** Gene expression levels of *SaAACT* (**a**) and *SaHMGS* (**b**) in different organs of *Santalum album* using qRT-PCR. The gene expression level of *SaAACT* and *SaHMGS* in the young leaves was set to 1. Data from qRT-PCR are means ± SD (standard deviation) from triplicate experiments (n = 3). Different letters indicate significant differences ($p < 0.05$) following one-way ANOVA, using Duncan's multiple range test.

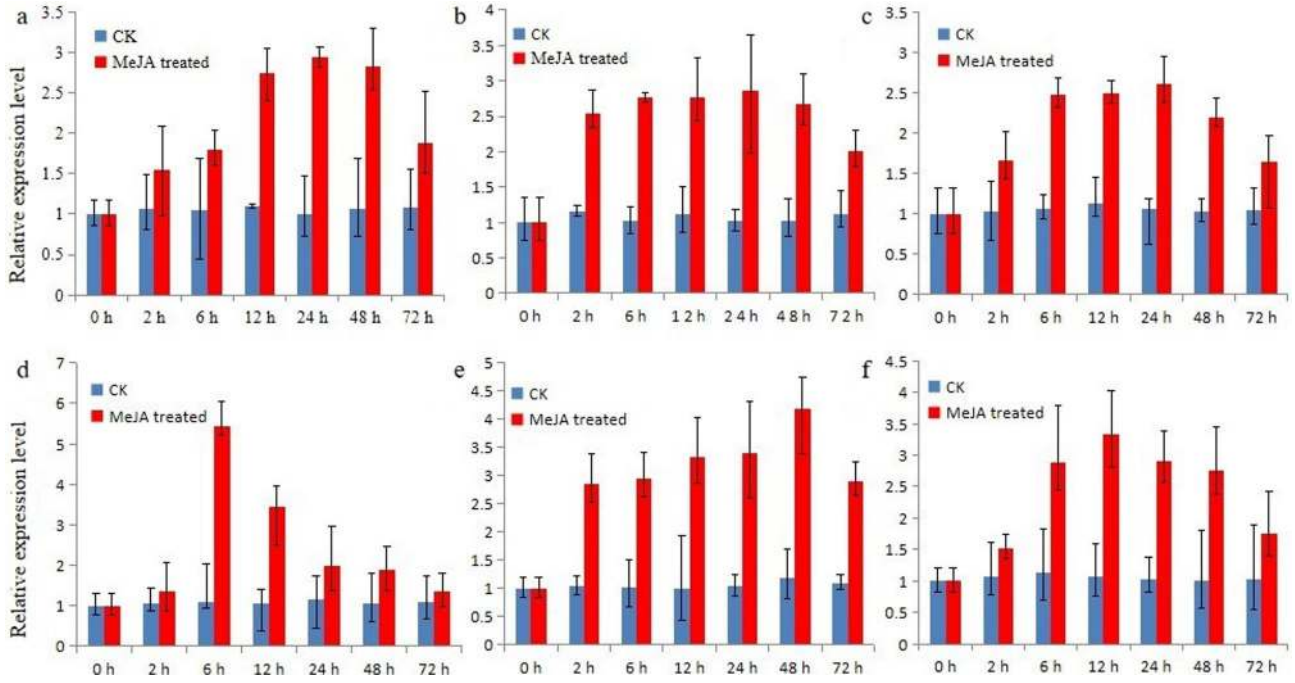


**Figure 8.** Level of *SaAACT* (**a**,**b**,**c**) and *SaHMGS* (**d**,**e**,**f**) transcripts in *Santalum album* roots (**a**,**d**), shoots (**b**,**e**) and leaves (**c**,**f**) after induction by 100 mM MeJA. The expression level of *SaAACT* and *SaHMGS* genes in the untreated control (CK) was set to 1. Data from qRT-PCR are means ± SD (standard deviation) from triplicate experiments (n = 3).

protein contained 470 amino acids and weighed 52 kDa. Bioinformatics analysis showed that *SaAACT* and *SaHMGS* genes have high similarity with other corresponding genes in other plants. The deduced amino acids of SaAACT and SaHMGS have typical motifs and domains of AACT and HMGS proteins. Multiple alignments and phylogenetic analyses showed that SaAACT and *H. brasiliensis* AACT were located in the same clade, while SaHMGS and *S. grosvenorii* HMGS were clustered together in a separate clade, indicating a close genetic relationship between these protein pairs. These results suggest that SaAACT and SaHMGS are conserved across different plants based on their conserved structure and sequence characteristics. Functional complementation experiments in yeast showed that *SaAACT* and *SaHMGS* encode functional AACT and HMGS, respectively. The expression of *HMGS* genes is related to the accumulation of terpenoids[64]. Overexpression of *AACT* and *HMGS* genes can increase the level of sesquiterpenoids in many plants. Previous studies showed that the expression patterns of

*AACT* and *HMGS* genes vary considerably in different plants[50,60]. The qRT-PCR results showed that *SaAACT* and *SaHMGS* genes were constitutively expressed in all the tested tissues but were differentially expressed in various tissues. The transcript level of *SaAACT* in roots was higher than in other tissues followed by mature leaves and heartwood, and the lowest expression level was in young leaves, which is consistent with a previous study[55]. The transcript level of *SaHMGS* in roots was higher than in other tissues, and the lowest expression level was in sapwood, which contrasts to *SmHMGS* reported previously[45,58]. Our results revealed that *SaAACT* and *SaHMGS* were expressed in all tissues, but at higher levels in roots. Many studies have shown that the yield of terpenoids is positively correlated with the expression of *AACT* and *HMGS* genes. Previous studies demonstrated that the accumulation of terpenoids increased in *AACT* overexpressing transgenic plants[65,66]. In *Ganoderma lucidum*, triterpene content increased in Gl-aact overexpressing transformants[65]. Overexpression of *A. thaliana AACT* in *Taraxacum brevicorniculatum* latex increased sterol levels by about five-fold[66]. An *Escherichia coli* strain that was transformed with a codon-optimized *HMGS* gene exhibited significantly more bisabolene production than control bacteria[67]. *BjHMGS* overexpressing transgenic plants significantly increased the total sterol content in leaves and seedlings[68]. Our results revealed that the expression levels of *SaAACT* and *SaHMGS* were highest in roots followed by heartwood. This trend is consistent with the chemical composition analysis in earlier reports[69,70], which showed that roots and heartwood are used to extract essential oils. These results support the relationship between transcription levels of *SaAACT* and *SaHMGS* and the content of sesquiterpenoids, indicating that *SaAACT* and *SaHMGS* may play an important role in the production of sesquiterpenoids in *S. album*. Our results show that the transcript levels of *SaAACT* and *SaHMGS* in roots, shoots and leaves of *S. album* seedlings increased after MeJA treatment, implying that the putative *SaAACT* and *SaHMGS* were responsive to the elicitor, MeJA, and could be induced, at least at the transcriptional level. Studying the biosynthesis of terpenoids at the molecular level is an important way to understand the mechanism of heartwood formation. The expression profiles of *AACT* and *HMGS* genes, which code for key enzymes in the biosynthesis of sesquiterpenes in plants, suggest that MeJA treatment may be an effective way to induce a high yield of sesquiterpenes in *S. album*. However, the regulatory mechanism of *AACT* and *HMGS* genes in the biosynthesis of sesquiterpenes in *S. album* needs to be clarified through additional studies.

The findings of our study are not only helpful to further understand the biosynthesis of santalols, but also provide a theoretical basis for further studies on the prokaryotic expression of related proteins. This study provides a molecular resource for increasing the content of santalols by genetic engineering.

## Materials and methods

Young leaves of five-year-old sandalwood trees (*S. album*) growing in South China Botanical Garden, Guangzhou, were used to isolate the *SaAACT* and *SaHMGS* genes. The young and mature leaves, shoots, heartwood, sapwood and roots were harvested to test the tissue-specific expression of *SaAACT* and *SaHMGS* genes by qRT-PCR. All collected samples were wrapped in tin foil, frozen immediately in liquid nitrogen, and then stored at −70 °C until further use.

**Total RNA extraction from sandalwood leaves and first-strand cDNA synthesis.** Total RNA of sandalwood leaves was extracted using Column Plant RNAOUT (Tiandz, Beijing, China) according to the manufacturer's protocol. The quantity and quality of isolated RNA was measured with a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, NC, USA). The integrity of isolated RNA was detected on a 1% agarose gel. High quality RNA was stored in DEPC-treated water at – 70 °C for future use. First strand cDNA were obtained with the PrimeScript first-strand cDNA synthesis kit (Takara Bio Inc., Dalian, China). Synthesized cDNA was stored at – 20 °C and served as the template for downstream reactions. For qRT-PCR, total RNA was isolated from different tissues (i.e.: roots, heartwood, sapwood, young leaves, mature leaves, and shoots) and first strand cDNA was synthesized as described above.

**Isolation and cloning of full-length *SaAACT* and *SaHMGS* cDNA by RACE.** The full-length cDNA sequence of *SaAACT* and *SaHMGS* genes was isolated by 5′ and 3′ rapid amplification of cDNA ends (RACE)-PCR with the SMARTer RACE cDNA Amplification Kit (Clontech, Palo Alto, CA, USA) according to the manufacturer's protocol. Gene-specific primers for 3′ RACE-PCR were designed on the basis of the initial data of AACT and HMGS unigenes in the *S. album* transcriptome[71]. Based on data from the partial *SaAACT* and *SaHMGS* sequence generated from 3′ RACE-PCR, gene-specific primers were designed for 5′ RACE-PCR to obtain the remaining sequences of *SaAACT* and *SaHMGS*. PCR products were purified by gel DNA purification kits (Tiangen, Beijing, China) and ligated into the pMD18-T vector (Takara Bio Inc., Dalian, China). Recombined plasmids were transformed into DH5α-competent *E. coli* cells (Takara Bio Inc., Dalian, China) and sequenced.

**Bioinformatics analysis of *SaAACT* and *SaHMGS*.** The nucleotide and amino acid sequences of *SaAACT* and *SaHMGS* were analyzed by bioinformatics methods, and their physical and chemical characteristics, transmembrane domain, signal peptide, and subcellular localization were predicted by corresponding bioinformatics software. Sequence assembly was performed with DNAstar (http://www.dnastar.com). Nucleotide sequences, deduced amino acid sequences and ORFs were analyzed, and sequences were compared through a BLAST database search (http://www.ncbi.nlm.nih.gov). Protein molecular weight and theoretical isoelectric point, instability index, aliphatic index and grand average of hydropathicity were calculated by ExPASy (http://www.expasy.ch/tools/). Protein conserved domains and active sites were predicted by the Conserved Domains database in NCBI (http://www.ncbi.mlm.nih.gov/Structure/cdd/wrpsb.cgi) and by Prosit (http://prosite.expasy.org/) in ExPASy. Transmembrane domains were analyzed on the TMHMM Server (http://www.cbs.-dtu.dk/

| Primer purpose | Primer name | Primer sequence (5′→3′) |
|---|---|---|
| 5′ RACE | AACT-5(1)/(2) | CACAAACACCCATGCCAAAATCA/TGCTTTCCATGCCACCAGCTACA |
| | HMGS-5(1)/(2) | AGTCTCGCTGCCTACTTCCATCC/TCCAGTGCTTCCTGTTGAACACA |
| 3′ RACE | AACT-3(1)/(2) | GAGGAAGCTCCGACCAAGTTTTA/GGAGCTGTATCTCTGGGACATC |
| | HMGS-3(1)/(2) | CTCAGTCAGCATGCCTAAACCT/CATTGTTGCCGGCTCTGTTC |
| ORF | AACT-O(F)/(R) | ATGGCTCCATCCGGGACGAAAGC/TAGCTTTGAAGGTCCTACTCTCA |
| | HMGS-O(F)/(R) | ATGGCTTCGCAGCAGAAGAAT/TTAATGGCCATTGGGAATTGAACCA |
| qRT-PCR | q-AACT-F/R | CGGGACGAAAGCCTCTTATT/AGGGAACCAAGAAAGTCACC |
| | q-HMGS-F/R | TCTCGCCATTGATGTCTACTTT/CCCAATGGTGTACTTCCCTTTA |
| Functional complementation | pYES2-AACT-F | CAGTGTGCTGGAATTCATGGCTCCATCCGGGACG |
| | pYES2-AACT-R | CATGCTCGAGCGGCCGCTAGCTTTGAAGGTCCTACTCTCAAG |
| | pYES2-HMGS-F | CAGTGTGCTGGAATTCATGGCTTCGCAGCAGAAGA |
| | pYES2-HMGS-R | CATGCTCGAGCGGCCGCATGGCCATTGGGAATTGAACCA |
| Subcellular localization | YFP-AACT-F | CCGGAATTCATGGCTCCATCCGGGACGAAAGC |
| | YFP-AACT-R | CGCGGATCCTAGCTTTGAAGGTCCTACTCTCA |
| | YFP-HMGS-F | CCGGAATTCATGGCTTCGCAGCAGAAGAA |
| | YFP-HMGS-R | CGCGGATCCATGGCCATTGGGAATTGAACCA |

**Table 3.** Specific primers used in this study.

services/TMHMM/). Signal peptides were analyzed by SignalP (http://www.cbs.dtu.dk/services/SignalP/). Protein subcellular localization was predicted by PSORT II (http://psort.org/).

**Molecular evolution analyses of SaAACT and SaHMGS proteins.** A phylogenetic tree of SaAACT and SaHMGS proteins from *S. album* and other plants, including angiosperms, gymnosperms, fungi, and bacteria, was constructed by the NJ method[72] with CLUSTALX 2.0 (Conway Institute, University College Dublin, Dublin, Ireland) and MEGA 7[73]. The calculation of evolutionary distance was based on the Poisson correction method with 1000 bootstrap repeats.

**Subcellular localization of SaAACT and SaHMGS.** A vector pSAT6-EYFP containing the ORF of enhanced yellow fluorescent protein (EYFP) was used in this study. The cDNA encoding *SaAACT* and *SaHMGS* were amplified with two pairs of primers, YFP-AACT-F and YFP-AACT-R, and YFP-HMGS-F and YFP-HMGS-R, respectively (Table 3). The PCR products were digested by *EcoR* I and *BamH* I restriction enzymes. The digested fragment was ligated into *EcoR* I- and *BamH* I-digested pSAT6-EYFP vector to generate pSAT6-EYFP-SaAACT and pSAT6-EYFP-SaHMGS fusion constructs. The fusion expression vectors and the pSAT6-EYFP vector were transformed into *A. thaliana* mesophyll protoplasts followed a method described previously[74]. After 16–24 h of incubation at 22 °C, YFP fluorescence in transformed protoplasts of *A. thaliana* was observed using a confocal laser-scanning microscope (Leica TCS SP8 STED 3X, Wetzlar, Germany).

**Functional complementation of *SaAACT* and *SaHMGS* in yeast.** To determine the function of *SaAACT* and *SaHMGS*, two ergosterol auxotrophic strains (Dharmacon, Chicago, IL, USA) of *Saccharomyces cerevisiae* that lacked the AACT or HMGS allele, named YPL028W (ΔERG10) and YML126C (ΔERG13), respectively, were used for the experiment. The pYES2 vectors, which contain a yeast galactose-dependent promoter, were used as carriers for target genes in this study. The coding regions of *SaAACT* and *SaHMGS* were amplified with two pairs of primers: pYES2-AACT-F and pYES2-AACT-R, and pYES2-HMGS-F and pYES2-HMGS-R, respectively (Table 3). The forward primers contained the *EcoR* I restriction site, and the reverse primers contained the *Not* I restriction site. The pYES2 vector was digested with *EcoR* I and *Not* I, and then PCR products and the digested fragment of the pYES2 vector were ligated into recombined vector pYES2-SaAACT and pYES2-SaHMGS by the In-Fusion HD Cloning Kit (Clontech, Palo Alto, CA, USA). The two constructed plasmids, pYES2-*SaAACT* and pYES2-*SaHMGS*, were extracted and transformed into YPL028W (ΔERG10) and YML126C (ΔERG13) with the Frozen-EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA, USA). *S. cerevisiae* strain ΔERG10, which lacks the AACT allele, and the *S. cerevisiae* strain ΔERG13, which lacks the HMGS allele, are both haploid yeast strains[59,75]. The transformed heterozygous diploid cells were forced to sporulate on YPG medium (1% yeast extract, 2% bacto-peptone, and 2% galactose)[38], thereby obtaining viable transformed haploid cells by dissecting tetrads. Haploid transformed cells bearing both the disrupted allele and the plasmid-borne MVA pathway genes were selected on minimal medium SC (-Ura) (6.7% yeast nitrogen base without amino acid, 2% galactose)[37]. The transformed diploid cells were induced to sporulate and subsequently formed haploid cells containing pYES2-*SaAACT* and pYES2-*SaHMGS*. To further observe their growth, transformed haploid strains YPL028W and YML126C were grown separately on YPD (1% yeast extract, 2% bacto-peptone, 2% glucose)[76] and YPG.

**Expression profiling of *SaAACT* and *SaHMGS* in different tissues and their induction by MeJA.** The expression levels of *SaAACT* and *SaHMGS* in different tissues (root, sapwood, heartwood, young

leaves, mature leaves, and shoots) and expression profiles after MeJA (Aladdin, Shanghai, China) treatment was detected by qRT-PCR. Two-month-old young seedlings (6–8 leaves) of *S. album* were sprayed with 100 μM MeJA until the leaf surfaces were wet. Samples were then collected at 0 h, 2 h, 6 h, 12 h, 24 h, 48 h and 72 h after treatment. About 1.0 μg of total RNA was reverse transcribed into first-strand cDNA using the PrimeScript RT Reagent Kit (Takara Bio Inc., Dalian, China) according to the manufacturer's protocols. The reaction was performed by ABI7500 fluorescence quantitative PCR (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) using the iTaq Universal SYBR Green Supermix as buffer (Applied Biosystems). Primer design, the reaction system and the reaction procedure were performed according to the manufacturer's instructions. The housekeeping gene, *β-actin*, was selected as the internal control[77]. 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. Melting curve analyses were conducted. All experiments were performed in triplicate and the mean value was analyzed. Gene expression analysis using the $2^{-\Delta\Delta CT}$ method[78] was used to normalize the relative gene expression of the transcripts in different tissues. Significant differences ($p < 0.05$) were assessed by one-way ANOVA, using Duncan's multiple range test. The results were represented by different letters.

## Data availability

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

Received: 20 July 2020; Accepted: 14 December 2020
Published online: 13 January 2021

## References

1. Harbaugh, D. T. & Baldwin, B. G. Phylogeny and biogeography of the sandalwoods (*Santalum*, Santalaceae): repeated dispersals throughout the Pacific. *Am. J. Bot.* **94**, 1028–1040 (2007).
2. Teixeira da Silva, J. A. *et al.* Sandalwood: basic biology, tissue culture, and genetic transformation. *Planta* **243**, 847–887 (2016).
3. Benencia, F. & Courrèges, M. C. Antiviral activity of sandalwood oil against *Herpes simplex* viruses-1 and -2. *Phytomed.* **6**, 119–123 (1999).
4. Solle, H. R. L. & Semiarti, E. Micropropagation of sandalwood (*Santalum album* L.) endemic plant from East Nusa Tenggara. *Indonesia. AIP Conf Proceed* **1744**, 020026 (2016).
5. Birkbeck, A. A. The synthesis of fragrant natural products from *Santalum album* L.: (+)-(*Z*)-α-santalol and (-)-(*Z*)-β-santalol. *Chimia* **71**, 823–835 (2017).
6. Brand, J. E., Fox, J. E. D., Pronk, G. & Cornwell, C. Comparison of oil concentration and oil quality from *Santalum spicatum* and *S. album* plantations, 8–25 years old, with those from mature *S. spicatum* natural stands. *Aust. For.* **70**, 235–241 (2007).
7. Subasinghe, S. M. C. U. P., Samarasekara, S. C., Millaniyage, K. P. & Hettiarachchi, D. S. Heartwood assessment of natural *Santalum album*, populations for agroforestry development in Sri Lanka. *Agrofor. Syst.* **91**, 1157–1164 (2017).
8. Jones, C. G., Plummer, J. A. & Barbour, E. L. Non-destructive sampling of Indian sandalwood (*Santalum album* L.) for oil content and composition. *J. Essent. Oil Res.* **19**, 157–164 (2007).
9. Howes, M. J., Simmonds, M. S. J. & Kite, G. C. Evaluation of the quality of sandalwood essential oils by gas chromatography-mass spectrometry. *J. Chromatogr. A* **1028**, 307–312 (2004).
10. Demole, E., Demole, C. & Enggist, P. A chemical investigation of the volatile constituents of east Indian sandalwood oil (*Santalum album* L.). *Helvet. Chim. Acta* **59**, 737–747 (2010).
11. Baldovini, N., Delasalle, C. & Joulain, D. Phytochemistry of the heartwood from fragrant *Santalum* species: a review. *Flav. Frag. J.* **26**, 7–26 (2011).
12. Hasegawa, T. *et al.* Isolation of new constituents with a formyl group from the heartwood of *Santalum album* L. *Flav. Frag. J.* **26**, 98–100 (2015).
13. Lichtenthaler, H. K. The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Ann. Rev. Plant Biol.* **50**, 47–65 (1999).
14. Rohmer, M. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat. Prod. Rep.* **16**, 565–574 (1999).
15. Lange, B. M., Rujan, T., Martin, W. & Croteau, R. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Nat. Acad. Sci. USA* **97**, 13172–13177 (2000).
16. Kuzuyama, T. Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units. *Biosci. Biotech. Biochem.* **66**, 1619–1627 (2002).
17. Estévez, J. M., Cantero, A., Reindl, A., Reichler, S. & León, P. 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J. Biol. Chem.* **276**, 22901–22909 (2001).
18. Dubey, V. S., Bhalla, R. & Luthra, R. An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *J. Biosci.* **28**, 637–646 (2003).
19. 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).
20. Laule, O. *et al.* Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc. Nat. Acad. Sci. USA* **100**, 6866–6871 (2003).
21. Miziorko, H. M. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* **505**, 131–143 (2011).
22. Tholl, D. Biosynthesis and biological functions of terpenoids in plants. *Adv. Bioch. Engin. Biotech.* **148**, 63–106 (2015).
23. Soto, G. *et al.* Acetoacetyl-CoA thiolase regulates the mevalonate pathway during abiotic stress adaptation. *J. Exp. Bot.* **62**, 5699–5711 (2011).
24. Luskey, K. L. & Stevens, B. Human 3-hydroxy-3-methylglutaryl coenzyme A reductase. Conserved domains responsible for catalytic activity and sterol-regulated degradation. *J. Biol. Chem.* **260**, 10271–10277 (1985).
25. Caelles, C., Ferrer, A., Balcells, L., Hegardt, F. G. & Boronat, A. Isolation and structural characterization of a cDNA encoding *Arabidopsis thaliana* 3-hydroxy-3-methylglutaryl coenzyme-A reductase. *Plant Mol. Biol.* **13**, 627–638 (1990).
26. Riou, C., Tourte, Y., Lacroute, F. & Karsta, F. Isolation and characterization of a cDNA encoding *Arabidopsis thaliana*, mevalonate kinase by genetic complementation in yeast. *Gene* **148**, 293–297 (1994).
27. Montamat, F., Guilloton, M., Karst, F. & Delrot, S. Isolation and characterization of a cDNA encoding *Arabidopsis thaliana* 3-hydroxy-3-methylglutaryl-coenzyme A synthase. *Gene* **167**, 197–201 (1995).
28. Lluch, M. A., Masferrer, A., Arró, M., Boronat, A. & Ferrer, A. Molecular cloning and expression analysis of the mevalonate kinase gene from *Arabidopsis thaliana*. *Plant Mol. Biol.* **42**, 365–376 (2000).

29. Carrie, C., Murcha, M. W., Millar, A. H., Smith, S. M. & Whelan, J. Nine 3-ketoacyl-CoA thiolases (KATs) and acetoacetyl-CoA thiolases (ACATs) encoded by five genes in *Arabidopsis thaliana* are targeted either to peroxisomes or cytosol but not to mitochondria. *Plant Mol. Biol.* **63**, 97–108 (2007).
30. Ahumada, I. *et al.* Characterisation of the gene family encoding acetoacetyl-CoA thiolase in *Arabidopsis*. *Func. Plant Biol.* **35**, 1100–1111 (2008).
31. Dai, Z. B., Cui, G. H., Zhou, S. F., Zhang, X. N. & Huang, L. Q. Cloning and characterization of a novel 3-hydroxy-3-methylglutaryl coenzyme A reductase gene from *Salvia miltiorrhiza* involved in diterpenoid tanshinone accumulation. *J. Plant Physiol.* **168**, 148–157 (2011).
32. Xu, Y. H. *et al.* Identification of genes related to agarwood formation: transcriptome analysis of healthy and wounded tissues of *Aquilaria sinensis*. *BMC Genom.* **14**, 227 (2013).
33. Chye, M. L., Tan, C. T. & Chua, N. H. Three genes encode 3-hydroxy-3-methylglutaryl-coenzyme A reductase in *Hevea brasiliensis*: *hmg1* and *hmg3* are differentially expressed. *Plant Mol. Biol.* **19**, 473–484 (1992).
34. Sirinupong, N., Suwanmanee, P., Doolittle, R. F. & Suvachitanont, W. Molecular cloning of a new cDNA and expression of 3-hydroxy-3-methylglutaryl-CoA synthase gene from *Hevea brasiliensis*. *Planta* **221**, 502–512 (2005).
35. Sando, T. *et al.* Cloning and characterization of mevalonate pathway genes in a natural rubber producing plant, *Hevea brasiliensis*. *Biosci. Biotechnol. Biochem.* **72**, 2049–2060 (2008).
36. Shen, G. A. *et al.* Cloning and characterization of a root-specific expressing gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase from *Ginkgo biloba*. *Mol. Biol. Rep.* **33**, 117–127 (2006).
37. Tao, T. 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).
38. Chen, Q. W. *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).
39. Jones, C. G. *et al.* Isolation of cDNAs and functional characterisation of two multi-product terpene synthase enzymes from sandalwood, *Santalum album* L. *Arch. Biochem. Biophys.* **477**, 121–130 (2008).
40. 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).
41. Rani, A., Ravikumar, P., Reddy, M. D. & Kush, A. Molecular regulation of santalol biosynthesis in *Santalum album* L. *Gene* **527**, 642–648 (2013).
42. Srivastava, P. L. *et al.* Functional characterization of novel sesquiterpene synthases from Indian sandalwood, *Santalum album*. *Sci. Rep.* **5**, 10095 (2015).
43. Diaz-Chavez, M. L. *et al.* Biosynthesis of sandalwood oil: *Santalum album* CYP76F cytochromes P450 produce santalols and bergamotol. *PLoS ONE* **8**, e75053 (2013).
44. Celedon, J. M. *et al.* Heartwood-specific transcriptome and metabolite signatures of tropical identity sandalwood (*Santalum album*) reveal the final step of (*Z*)-santalol fragrance biosynthesis. *Plant J.* **86**, 289–299 (2016).
45. Kai, G. *et al.* Molecular cloning and expression analyses of a new gene encoding 3-hydroxy-3-methylglutaryl-CoA synthase from *Taxus× media*. *Biol. Plant.* **50**, 359–366 (2006).
46. Anderson, V. E., Bahnson, B. J., Wlassics, I. D. & Walsh, C. T. The reaction of acetyldithio-CoA, a readily enolized analog of acetyl-CoA with thiolase from *Zoogloea ramigera*. *J. Biol. Chem.* **265**, 6255–6261 (1990).
47. Price, A. C. *et al.* Inhibition of β-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin structure and mechanism. *J. Biol. Chem.* **276**, 6551–6559 (2001).
48. Theisen, M. J. *et al.* 3-hydroxy- 3-methylglutaryl-CoA synthase intermediate complex observed in "real-time". *Proc. Nat. Acade. Sci. USA* **101**, 16442–16447 (2004).
49. Campobasso, N. *et al. Staphylococcus aureus* 3-hydroxy-3-methylglutaryl-CoA synthase. *J. Biol. Chem.* **279**, 44883–44888 (2004).
50. Ahumada, I. *et al.* Characterization of the gene family encoding acetoacetyl-CoA thiolase in *Arabidopsis*. *Func. Plant Biol.* **35**, 1100–1111 (2008).
51. Nagegowda, D. A. Plant volatile terpenoid metabolism: Biosynthetic genes, transcriptional regulation and subcellular compartmentation. *FEBS Lett.* **584**, 2965–2973 (2010).
52. Nagegowda, D. A., Ramalingam, S., Hemmerlin, A., Bach, T. J. & Chye, M. L. *Brassica juncea* HMG-CoA synthase: localization of mRNA and protein. *Planta* **221**, 844–856 (2005).
53. Trocha, P. J. & Sprinson, D. B. Location and regulation of early enzymes of sterol biosynthesis in yeast. *Arch. Biochem. Bioph.* **174**, 45–51 (1976).
54. Servouse, M. & Karst, F. Regulation of early enzymes of ergosterol biosynthesis in *Saccharomyces cerevisiae*. *Biochem. J.* **240**, 541–547 (1986).
55. Vishwakarma, R. K. *et al.* Molecular cloning, biochemical characterization, and differential expression of an acetyl-CoA C-acetyltransferase gene (AACT) of brahmi (*Bacopa monniera*). *Plant Mol. Biol. Rep.* **31**, 547–557 (2013).
56. Pauwels, L. *et al.* Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured Arabidopsis cells. *Proc. Nati. Acad. Sci. USA* **105**, 1380–1385 (2008).
57. Robert-Seilaniantz, A., Grant, M. & Jones, J. D. G. Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. *Ann. Rev. Phytopath.* **49**, 317–343 (2011).
58. Zhang, L. *et al.* Molecular cloning and expression analysis of a new putative gene encoding 3-hydroxy-3-methylglutaryl-CoA synthase from *Salvia miltiorrhiza*. *Acta Physiol. Plant.* **33**, 953–961 (2011).
59. Liu, Y. J. *et al.* Cloning and characterisation of the gene encoding, 3-hydroxy-3-methylglutaryl-CoA synthase in *Tripterygium wilfordii*. *Molecules* **19**, 19696–19707 (2014).
60. Brocke, C., Eh, M. & Finke, A. Recent developments in the chemistry of sandalwood odorants. *Chem. Biodiv.* **5**, 1000–1010 (2008).
61. Muratore, A., Clinet, J. C. & Dunach, E. Synthesis of new exo- and endo-3,8-dihydro-β-santalols and other norbornyl-derived alcohols. *Chem. Biodiv.* **7**, 623–638 (2010).
62. Zhang, Y. Y. *et al.* Multiple strategies for increasing yields of essential oil and obtaining sandalwood terpenoids by biotechnological methods in sandalwood. *Trees* **32**, 21–29 (2018).
63. Oldfield, S., Lusty, C., MacKinven, A. The World List of Threatened Trees, WCMC, IUCN, 632 (1998).
64. Dooley, K. A., Millinder, S. & Osborne, T. F. Sterol regulation of 3-hydroxy-3-methylglutaryl-coenzyme A synthase gene through a direct interaction between sterol regulatory element binding protein and the trimeric CCAAT-binding factor/nuclear factor Y. *J. Biol. Chem.* **273**, 1349–1356 (1998).
65. Fang, X. *et al.* The cloning, characterization and functional analysis of a gene encoding an acetyl-CoA acetyltransferase involved in triterpene biosynthesis in *Ganoderma lucidum*. *Mycoence* **54**, 100–105 (2013).
66. Pütter, K. M., Deenen, N. V., Unland, K., Prüfer, D. & Gronover, C. S. Isoprenoid biosynthesis in dandelion latex is enhanced by the overexpression of three key enzymes involved in the mevalonate pathway. *BMC Plant Biol.* **17**, 88 (2017).
67. Peralta-Yahya, P. P. *et al.* Identification and microbial production of a terpene-based advanced biofuel. *Nat. Comm.* **2**, 483 (2011).
68. Wang, H., Nagegowda, A. D., Rawat, R., Bouvier-Navé, P., Guo, D. J., Bach, T. J. & Chye, M. L. Overexpression of *Brassica juncea* wild-type and mutant HMG-CoA synthase 1 in *Arabidopsis* up-regulates genes in sterol biosynthesis and enhances sterol production and stress tolerance. Plant Biotech. J. 31–42 (2012).
69. Burdock, G. A. & Carabin, I. G. Safety assessment of sandalwood oil (*Santalum album* L.). *Food Chem. Toxicol.* **46**, 421–432 (2008).

70. Zhang, X. H., Teixeira da Silva, J. A., Jia, Y. X., Yan, J. & Ma, G. H. Essential oil composition from roots of *Santalum album* L.. *J Essent Oil Bear. Plants* **15**, 1–6 (2012).
71. Zhang, X. H. *et al.* Physiological and transcriptomic analyses reveal a response mechanism to cold stress in *Santalum album* L. leaves. *Sci. Rep.* **7**, 42165 (2017).
72. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
73. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
74. Yoo, S. D., Cho, Y. H. & Sheen, J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* **2**, 1565–1572 (2007).
75. Hiser, L., Basson, M. E. & Rine, J. ERG10 from *Saccharomyces cerevisiae* encodes acetoacetyl-CoA thiolase. *J. Biol. Chem.* **269**, 31383–31389 (1994).
76. Albers, E. & Larsson, C. A comparison of stress tolerance in YPD and industrial lignocellulose-based medium among industrial and laboratory yeast strains. *J. Indus. Microbiol. Biotech.* **36**, 1085–1091 (2009).
77. Zhang, X. H. *et al.* RNA-Seq analysis identifies key genes associated with haustorial development in the root hemiparasite *Santalum album*. *Front. Plant Sci.* **6**, 661 (2015).
78. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **3**, 1101–1108 (2008).
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).


## Author contributions

M.Y.N., Y.P.X., H.F.Y., Y.Y.Z., X.H.Z., Y.L., J.A.TdS. and G.H.M. designed the experiment and provided guidance for the study. MYN and YPX prepared samples for all analyses. H.F.Y. and Y.Y.Z. participated in statistical analyses. M.Y.N. and J.A.Td.S. co-wrote the manuscript. G.H.M. interpreted the data and experimental results. All authors read and approved the manuscript for publication.

## 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, 31270720 and 31100498) and a Guangdong key Science and Technology project (2015B020231008).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to G.M.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.