CLONING, MOLECULAR CHARACTERIZATION AND FUNCTIONAL ANALYSIS OF THE CIS-MUUROLADIENE SYNTHASE (*SgCMS*) GENE FROM LEAVES OF *SALVIA GUARANITICA* PLANT

Mohammed Ali
Genetics and Cytology Unit, Department of Genetic Resources, Desert Research Center, El-Matareya, Cairo, Egypt
E-mail: mohammedalidrc@gmail.com

In terms of its nutritional and medical value, *Salvia guaranitica* is ranked among the most important medicinal and aromatic plants because of the abundance of essential active components it contains. The most important and abundant active ingredients are terpenoid molecules, particularly monoterpenes (C10) and sesquiterpenes (C15). Terpenes play a variety of tasks and possess beneficial biological traits in plants. Considering these characteristics, the current research was done to clone *S. guaranitica* gene for Cis-muuroladiene synthase (*SgCMS*, EC: 4.2.3.67). The 600 amino acid protein coded by *SgCMS* has an open reading frame of 1.802 base pairs. Moreover, *SgCMS* protein has four terpene synthase family domains which belong to terpenes and terpenoid synthase domains. We altered and overexpressed *SgCMS* gene in *Arabidopsis thaliana* to investigate the gene's function. Compared to the wild-type plants, the transgenic *A. thaliana* plants showed a decrease in leaf diameter, while the flowering stems started growth earlier when compared to the GUS control. Ultimately, the results of the transgenic plants' gas chromatography-mass spectrometry examination revealed that *SgCMS* was in charge of producing several terpene types, particularly sesquiterpene.

Keywords: *Salvia*, cis-muuroladiene synthase, *Arabidopsis*, cloning, molecular characterization

INTRODUCTION

More than 1,000 species of woody fragrant shrubs make up the genus *Salvia* (Lamiaceae), some of which, like *S. epidermindis, S. miltiorrhiza, S. officinalis, S. aureus, S. allagospadonopsis, S. przewalskii,* and *S. japonica,* are economically significant and are grown all over the world for their numerous medical benefits and the manufacture of their essential oils (EOs). Several *Salvia* plant species are found around the planet, although they are primarily found in Central and South America (500 species), East Asia (100 species), and West Asia (200 species), respectively (Ali et al., 2017; Sarrou et

al., 2017 and Ali et al., 2018). In contrast, the other *Salvia* species are found all over the globe. *Salvia* species EOs have recently emerged as an essential source for pharmacological and aromatic research to find and characterize physiologically active chemicals. *Salvia* Eos have significant bioactivities, such as antibacterial, anticancer, anti-inflammatory, choleretic, antioxidant, and ant-mutagenic properties (Li et al., 2015; Wang et al., 2015a and Zhenqing et al., 2018).

Terpenoid compounds, which have been identified in the plant kingdom with more than 40,000 different structural variations, are widely recognized as the biggest group of natural products and a class of secondary metabolites (Ali et al., 2018; 2022a and b). The structural basis for terpenoids comes from the compound isopentenyl diphosphate (IPP), which has five carbon atoms (C5) (Xi et al., 2016 and Abbas et al., 2019). However, *Pistacia terebinthus* tree, from which all of these were recognized as distinct terpene compounds, is where these various structures got their original names. The terpene unit's structure was changed and illustrated by Ruzicka (Ruzicka 1953 and 1973). Some terpenes, such as the carotenoid pigments, phytol side chains of chlorophyll, gibberellin plant hormones, and phytosterols of cellular membranes, are connected to the fundamental plant metabolism and are crucial for plant growth, flowering, and development (Gershenzon, 1999 and Trapp and Croteau, 2001). However, most of the identified terpenes fall under the category of secondary metabolites and are crucial to how plants interact with their surroundings. Both nonvolatile and volatile terpenes have a part in processes like defense against herbivore predators, photo-oxidative stress, pollinator attraction, and natural defense against microorganisms and insects (Dorothea et al., 2006). Numerous investigations are being conducted to fully comprehend how terpene and terpenoid functions work (Ali et al., 2021; 2022b and c).

It is common knowledge that *Salvia* species contain a high proportion of essential oil; this fragrant essential oil mainly contains a lot of monoterpenes and sesquiterpenes (Fateme et al., 2013 and Ali et al., 2017). According to Ali et al. (2018), the primary sesquiterpenes in the Chinese cultivar of *S. guaranitica* are germacrene-A, isocaryophyllene, humulene, -caryophyllene, pi-a-muurolene, caryophyllene oxide, and d-cadinene. These previous sesquiterpenes compounds have various biological and physiological functions, and most genes responsible for synthesizing these sesquiterpenes compounds in *Salvia* are still obscure. In this investigation, the Cis-muuroladiene synthase (*SgCMS*, EC: 4.2.3.67) gene from *S. guaranitica* was cloned and functionally characterized. Additionally, the recombinant *SgCMS* catalyzed the transformation of farnesyl pyrophosphate (FPP) into a single product with several forms of terpene, particularly sesquiterpene.

MATERIALS AND METHODS

1. Plant Materials

S. guaranitica plantlets were obtained from the Wuhan Botanical Garden (WBG), Chinese Academy of Sciences, Hubei, Wuhan, China. And then, the plantlets were grown in the greenhouse at National Research Centre, Cairo, Egypt. Six leaves from four-year-old plants were collected and immediately frozen in liquid nitrogen for 10 min and then stored at -80°C for RNA isolation and gene cloning.

2. Sequence Characterization of *SgCMS*

The sequence of the *SgCMS* gene was selected based on the highest sequence similarity found with the known plant sesquiterpene syntheses genes. The physical and chemical property of the *SgCMS* was assessed using PROTPARAM Server (http://web.expasy.org/protparam/). The putative transit peptide for open-reading-frames (ORF) of *SgCMS* was analyzed by the iPSORT prediction tool (http://ipsor t.hgc.jp/). NCBI BLASTX tool was used for comparative protein sequence analysis of *SgCMS* (http://blast.ncbi.nlm.nih.gov/). The evolutionary relationships of *SgCMS* protein and other plant TPS proteins were inferred using the PhyML Server without changing the tool parameters (http://www.phylogeny.fr/) (Dereeper et al., 2008 and Mehmood et al., 2021).

3. Putative Tissue Expression Pattern of *SgCMS* Gene and Its Subcellular Localization

Tissue-specific expression data from -six tissues were analyzed using *Arabidopsis thaliana* transcript expression database. *Arabidopsis* eFP browsers (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) were used to generate expression profiles. While the putative sub-cellular localizations of *SgCMS* gene from *S. guaranitica* was analyzed using the Cell-eFP browsers (http://bar.utoronto.ca/cell_efp/cgi-bin/cell_efp.cgi) as described by Makhadmeh et al. (2022a and 2022b). The color boxes represent the expression scale (e.g. red color= high expression and yellow color = low expression).

4. RNA Extraction and cDNA Synthesis

Six tissues with three replicates from four-year-old *S. guaranitica* were used for RNA extraction using TransZol Reagent (Focus Bioscience, Australia) for gene cloning and qRT-PCR. Also, leaves from wild and transgenic *A. thaliana* were used for RNA extraction for semi-quantitative RT-PCR. For cDNA Synthesis, 1 μg from each RNA was used to synthesize the first-strand cDNA using reverse transcriptase TransScript® First-Strand cDNA Synthesis Super Mix kit as described by Ali et al. (2017) and Hussain et al. (2017).

5. Isolation of Full-Length of *SgCMS* Gene

The entire *SgCMS* cDNA was used as a template to magnify the full-length by short-gene-specific forward (5'- TCACAACAGAATGTCAACATATAGC-3') and reverse (5'-AATGGGTATTGGATTGA TAAAGAG-3') primers and TaKaRa Taq DNA Polymerase enzyme and amplification were done at 5 min at 95°C; 36 cycles for 12 s at 97°C; 30 s at 58°C; 95s at 68°C, and then 11 min at 68°C. The first PCR product was used as a model for the second PCR using long-gene-specific forward (5-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTCACAACAGAATGTCAAC-3) and reverse (5-GGGGACCACTTTGTACAAGAAAGCTGGGTAATGGGTATTGGATTGAT-3) primers with the same polymerase enzyme and PCR program conditions to clone into Gateway vectors. The successful amplicon was purified and cloned into a pDONR221-vector (Invitrogen, Carlsbad, CA, USA), then a pB2GW7-vector (Invitrogen, Carlsbad, CA, USA) and sent for sequencing as depicted by Ali et al. (2017 and 2018).

6. Growth Conditions of *Arabidopsis* Plants and Transformation Using *Agrobacterium*

A. thaliana seeds from ecotype Columbia-0 (Col-0) were grown in our Lab growth chamber as reported by Ali et al. (2021). After 2 months from the growth, healthy plants at the pre-flowering stage were selected for the floral-dip transformation experiment. Moreover, to characterize the function of *SgCMS* gene, the vector pB2GW7-*SgCMS* and pB2GW7 (empty vector) were transformed into *Agrobacterium tumefaciens* strain GV101. *A. tumefaciens* GV101 containing pB2GW7-*SgCMS* was grown in selective solid LB media supplied with rifampin (Rif) and spectinomycin (Spc) as antibiotics. One positive colony was selected and inoculated into 0.8 ml of liquid-LB-media with Rif and Spc, after one day from the incubation period at 28°C, the bacterial culture was sub-cultured to a conical flask containing about 60 ml LB media and incubation at an incubator with shaker until the optical density of *Agrobacterium* cells reached to 0.75 (OD 600) according to Ali et al. (2017, 2018) and Darwish et al. (2022). The next day, cell suspension was collected and the bacteria was re-suspended in fresh-inoculation-medium (5.2% sucrose and 0.055% Silwet). The plasmids pB2GW7-*SgCMS* and pB2GW7 (control) were introduced separately into *A. thaliana* plant by directly immersing the axis of bud flowers in the fresh floral-dip medium with a gentle press to ensure the intake of *A. tumefaciens* GV101 harbouring the pB2GW7-*SgCMS* and pB2GW7 (control) plasmids into the flower gynaecium (Aharoni et al., 2003 and Su-Fang et al., 2014). A total of 12 *A. thaliana* transgenic lines were obtained and survived the successive subculture process under BASTA resistance. The previous lines' leaf morphology, flowering time and terpene metabolic were assessed

7. Semi-Quantitative RT-PCR (Sqrt-PCR) Analysis

To ensure the success of the gene transfer process, qRT-PCR was implemented using a PCR system from Biometra. The *At-B-actin* gene forward primer 5'-GGCTGAGGCTGATGATATTC-3' and reverse primer 5'-CCTTCTGGTTCATCCCAAC-3' was used as housekeeping with 155 bp, and *SgCMS* forward primer 5'-GTAACTCGCATCGCACAT-3' and reverse primer 5'- GAAGCAAGAGACGCTATCAG-3' with 145 bp length. We used IDTdna online website (http://www.idtdna.com/scitools/Applications /RealTimePCR/) to design our target primers. The Biometra Thermocycler T-Gradient Thermo Block conditions program as follows: 96°C for 6 min, 35 cycles of 94°C for 35 s, 60°C for 30 s, and 72°C for 1.2 min, 72°C for 11 min. PCR products screened on 1.6 % agarose gel to determine our gene expression levels.

8. Terpenoids Extraction and GC-MS Analysis

For a rapid survey of terpenoids in transgenic and non-transgenic *A. thaliana* plant lines, intact leaves of various lines were frozen in liquid nitrogen (LN_2), powdered with a ceramic mortar and pestle, and directly inundation in n-hexane as a solvent for 75 h for ensuring complete removal of terpenoids contents as previously described (Ali et al., 2022a; b and c). Approximately 1 µl aliquot of each extract was analysed by the Shimadzu-GC-MS system with three replicates. Terpenoid component identification was made by reference to Wiley GC/MS Library, the Volatile Organic Compounds (VOC) Analysis S/W software, and the NIST Library as previously described (Ali et al., 2017 and 2018).

9. Quantitative Real-Time PCR (qRT-PCR) Analysis

To analyze the expression of *SgCMS* in various *S. guaranitica* tissues (e.g., young leaves, stems, old leaves, roots, bud flowers, and flowers), tissues were collected with three biological replicates. QRT-PCR was performed to represent the scale of *SgCMS* transcript involved in the production of Cis-muuroladiene synthase. Fragments of 166 and 145 bp in the 3’ region of *SgACTIN* and *SgCMS* were amplified using the following couple of primers: for *SgACTIN* forward 5′-CTGGATTTGCGGGAGATG-3′ and reverse 5′-CCGTGCTCAATTGGATACTT-3′ was used as a housekeeping gene, and *SgCMS* forward 5'-GTAACTCGCATCGCACAT-3' and reverse 5'-GAAGCAAGAGACGCTATCAG-3'. The qRT-PCR experiments were performed using IQTM5 System, SYBR Green, and the cycler program as follows: 96°C for 10 s, 60°C for 30 s, and 72°C for 20 s, then 66°C for 5 s and 94°C for 6 s. The expression levels were enumerated by comparing our target gene cycle thresholds (CTs) with the housekeeping gene *SgACTIN* using the $^{2-\Delta\Delta}$Ct method (Anders and Huber, 2010; Hussain et al., 2017 and Rehman et al., 2017). Values were offered as means ± SE of three different RNA pool replicates.

RESULTS

1. Full-Length Isolation of Cis-Muuroladiene Synthase (*SgCMS*) Gene and Sequence Characterization

The full-length-ORF of *SgCMS* gene with 1802 bp encoded a 600 amino acid protein with a predicted theoretical isoelectric point (pI) of 9.46, aliphatic index 93.26, and molecular mass of 65.26 kDa. The amino acids (aa) of SgCMS have the same length signal peptide with monoterpene synthases (600–650 aa) and matched with many sesquiterpene synthases of 550–580 aa in the thirty amino acids (MRFFFHSYELTTLRTYTKYTRTFSNKFINP) that exist at N-terminal sequence. Using 'iPSORT' program revealed that *SgCMS* is localized in the mitochondria or chloroplast, where the FPP is originated, and their biosynthesis takes place. NCBI-BLASTX analysis in Table (1) indicated that *SgCMS* shared identity between 96.01% to ⩽ 52.89% with its homologue sesquiterpene synthase protein from *Salvia splendens* and other plants, respectively. The sequence alignment of *SgCMS* gene with putative and recognized TPS genes from Lamiaceae and other plants aided the prediction of its putative function. Based on this prediction, *SgCMS* protein are specified by the InterPro database. So, *SgCMS* protein has four terpene synthase family domains, such as terpenoid cyclases/protein prenyltransferase alpha-alpha toroid (IPR008930: from 74 - 277), isoprenoid synthase domain (IPR008949: from 279 - 605), terpene synthase, metal-binding domain (IPR005630: from 284 - 547) and terpene synthase, N-terminal domain (IPR001906: from 75 - 253) (Fig. 1). Some of these previous domains are predominant in similar sesquiterpene synthases (Fahmy et al., 2022; Ali et al., 2022a and b). The *SgCMS* was tabulated to the TPS-g subfamily of Angiosperm sesquiterpene synthases based on the phylogenetic analysis results (Fig. 2). This subfamily is primarily responsible for synthesizing mono-, sesquit- and diterpene synthase (Ali et al., 2017; 2018 and Jiang et al., 2019).

2. Putative Tissue Expression and Subcellular Localization of *SgCMS* Gene

To examine the putative tissue expression pattern of *SgCMS* in the *A. thaliana* genomics, a BlastP search against *A. thaliana* genomics at Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html#!search?show=BLAST&method=Org_Athaliana) was conducted with the protein sequence of *SgCMS* as a query. This research identified several proteins closely related to the *SgCMS* sequences, especially (AT1G31950), with a high BLAST identities % and e-value (91% and 2.84e-2), respectively. The tissue expression of *SgCMS* gene in *Arabidopsis* uncovered by these data were analysed across fifty six tissues using BAR database (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) and the

Arabidopsis Electronic Fluorescent Pictograph Browsers (eFP browsers (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) (Fig. 3a, b, c and d). The *Arabidopsis* eFP browsers clearly showed that (*SgCMS* gene: AT1G31950) was presented in most of the tissues with highly expression in root (217.6), seeds stage 5 w/ siliques (37.77), seeds stage 6 w/o siliques (28.72), seeds stage 7 w/o siliques (22.38), imbibed seed, 24 h (18.57) and leaf 7, proximal half (12.08) (Fig. 3a and Table S1). Also, *SgCMS* gene was highly expressed in all tissue specific trichomes (Fig. 3b and Table S1). On the other hand, *SgCMS* gene was remarked and highly expressed in tissue-specific stem epidermis, especially in the stem epidermis, top of stem (10.46), followed by whole stem, top of stem (8.59), stem epidermis, bottom of stem (7.68) and whole stem, bottom of stem (5.62) (Fig. 3c and Table S1). Moreover, putative subcellular localization built using ePlant and cell eFP (http://bar.utoronto.ca/cell_efp/cgi-bin/cell_efp.cgi) for (*SgCMS* gene: AT1G31950) gene was present with different expression levels at the fourteen cell organelles see (Fig. 3d).

3. Screening the Expression of *SgCMS* Gene Using qRT-PCR

QRT-PCR was used to inspect the transcription levels of *SgCMS* at various tissues of *S. guaranitica* (e.g., young leaves, stems, old leaves, roots, bud flowers, and flowers) (Fig. 4). From the qRT-PCR analysis results, it was found that the highest expression levels were observed in young leaves followed by old leaves, flowers, bud flowers, stems and roots (Fig. 4).

4. Functional Expression of *SgCMS* Gene in Non- and Transgenic *A. thaliana* Plants

To study the effect of *SgCMS* gene on *A. thaliana* plants phenotypes after 37 days of growth, this gene was cloned from *S. guaranitica.* After that, *A. thaliana* was utilized as a transient expression system to overexpress the isolated *SgCMS* gene. *A. tumefaciens* strain GV101 harbouring the vector pB2GW7- *SgCMS* was used and controlled by 35S promoter to generate transgenic *A. thaliana* plants overexpressing *SgCMS* gene constitutively (Fig. 5a). The positive transformants were further verified using BASTA reagent and sqRT-PCR of the genomic cDNA (Fig. 5b). Transgenic *A. thaliana* plants showed a decrease in leaf diameter, while the flowering stems start growth earlier when compared to the GUS control.

Table (1). BLASTX analysis of *SgCMS* compared with the NCBI protein database for gene identification purposes.

NCBI Accession	[a]Descriptiona	Organism	*E* value	Identity (%)	Accession length
XP_042020998.1	Gamma-cadinene synthase-like	*Salvia splendens*	0.0	96.01	550
XP_042008943.1	Cis-muuroladiene synthase-like	*Salvia splendens*	0.0	85.95	555
XP_042063765.1	Gamma-cadinene synthase-like isoform X1	*Salvia splendens*	0.0	83.03	542
XP_042047639.1	Cis-muuroladiene synthase-like	*Salvia splendens*	0.0	79.67	551
XP_042059247.1	Cis-muuroladiene synthase-like	*Salvia splendens*	0.0	72.14	542
XP_042059101.1	Cis-muuroladiene synthase-like isoform X1	*Salvia splendens*	0.0	71.03	542
XP_042063774.1	Gamma-cadinene synthase-like isoform X2	*Salvia splendens*	0.0	72.51	483
XP_042059105.1	Gamma-cadinene synthase-like isoform X2	*Salvia splendens*	0.0	68.27	527
Q5W283.1	RecName: Full=Cis-muuroladiene synthase	*Mentha x piperita*	0.0	66.97	551
Q5SBP5.1	RecName: Full=Gamma-cadinene synthase	*Ocimum basilicum*	0.0	65.99	540
ADC92564.1	(E)-beta farnesene synthase	*Mentha arvensis*	0.0	65.57	550
O48935.1	RecName: Full=Beta-farnesene synthase	*Mentha x piperita*	0.0	64.29	550
CAH10289.1	(E)-beta-farnesene synthase	*Mentha x piperita*	0.0	64.29	550
AEA49039.1	(E)-beta-farnesene synthase 2	*Mentha asiatica*	0.0	64.10	550
AEK32002.1	(E)-B-farnesene synthase	*Mentha x piperita*	0.0	63.92	550
U3LW50.1	RecName: Full=Tau-cadinol synthase	*Lavandula angustifolia*	0.0	60.22	555
XP_012844573.1	PREDICTED: cis-muuroladiene synthase	*Erythranthe guttata*	0.0	55.95	555
XP_012844616.1	PREDICTED: cis-muuroladiene synthase	*Erythranthe guttata*	0.0	57.88	545
XP_011102272.2	Gamma-cadinene synthase	*Sesamum indicum*	0.0	57.46	542
PIN23179.1	(+)-gamma-cadinene synthase	*Handroanthus impetiginosus*	0.0	54.84	544
PIN22639.1	Cis-muuroladiene synthase	*Handroanthus impetiginosus*	0.0	52.89	552

[a]Description—homology search using BLASTX

5. Overexpression of *SgCMS* Gene Altered the Terpene Profiles in Transgenic *A. thaliana* Leaves

To examine the effect of overexpressing *SgCMS* gene in *A. thaliana* leaves, the changes of terpene metabolism were analysed in transgenic *A. thaliana* leaves using GC-MS. The results revealed that a variety of terpenes significantly increased in transgenic *A. thaliana* leaves overexpressing *SgCMS* gene in parallel with control as shown in Fig. (6a) and Table (2). In leaves of *A. thaliana* plants overexpressing *SgCMS*, cis-muurol-5-en-4-α-ol appeared as the main sesquiterpene compound (14.81%) (Fig 6b), followed by Cis-muurola-3,5-diene (8.68%) and γ-muurolene (8.67%); whereas, 2α,9-dihydroxy-1,8-cineole (0.73%) and 17-α-methyl-5-β-androstan-3-β,17-β-dioll compound (6.74%) were observed as the significant monoterpene and diterpene compounds, respectively.

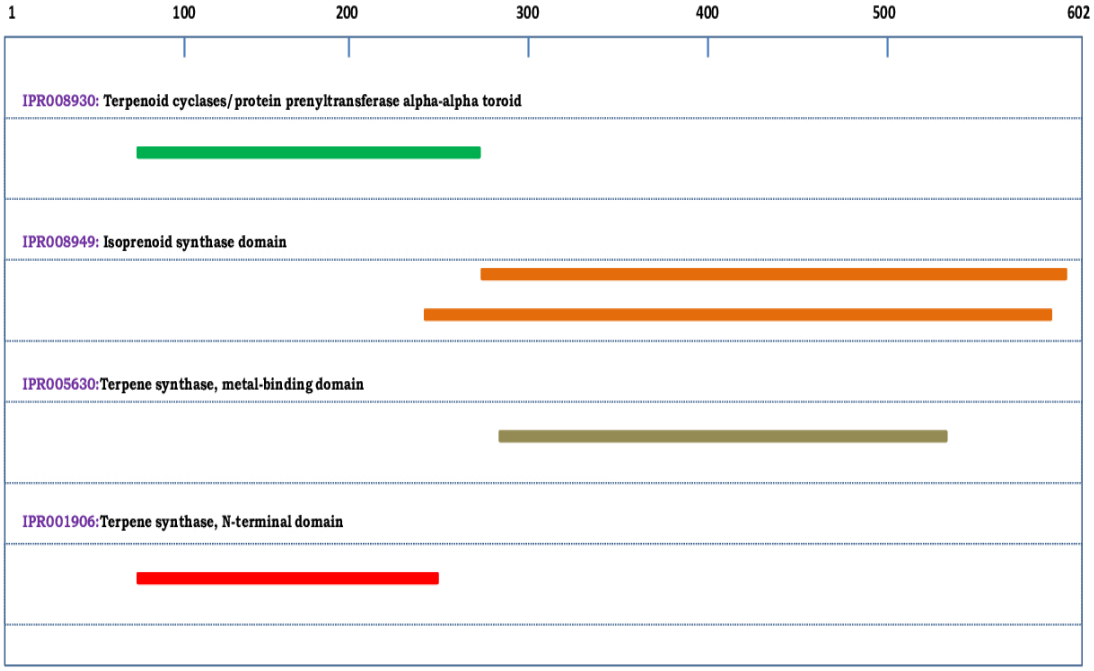

Fig. (1). Putative domain analysis of SgCMS using the InterPro protein sequence analysis and classification (https://www.ebi.ac.uk/interpro/) database. SgCMS protein sequence have four protein family domains.

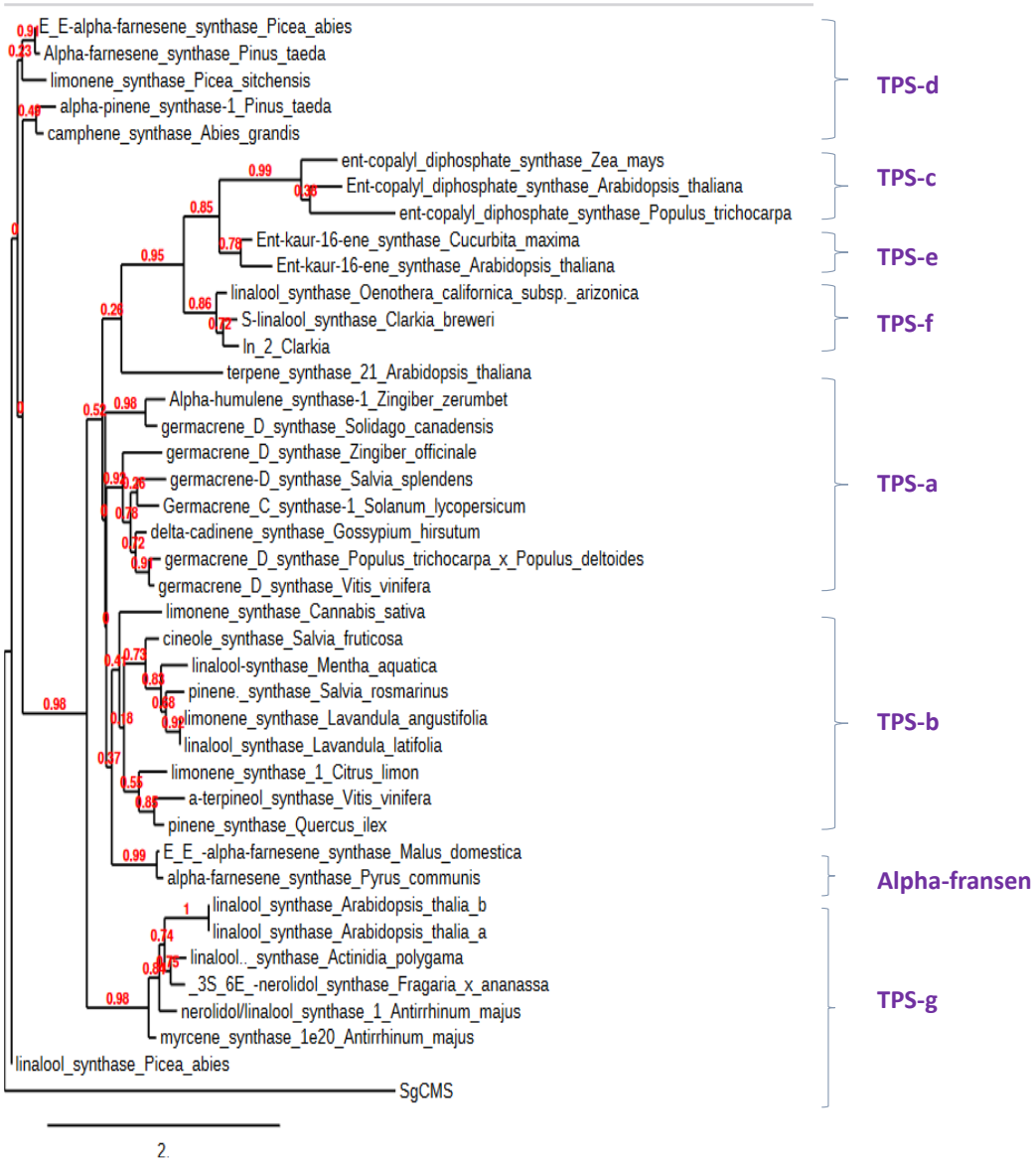

Fig. (2). Phylogenetic tree of *SgCMS* with selected terpene synthases from other plants. Seven previously identified TPS subfamilies (Tps-a to Tps-g) were chosen based on Bohlmann et al. (1998); Fahmy et al. (2022) and Elsherbeny et al. (2022). The alignment was performed

using the PhyML server. The numbers indicated are the actual bootstrap values of the branches.

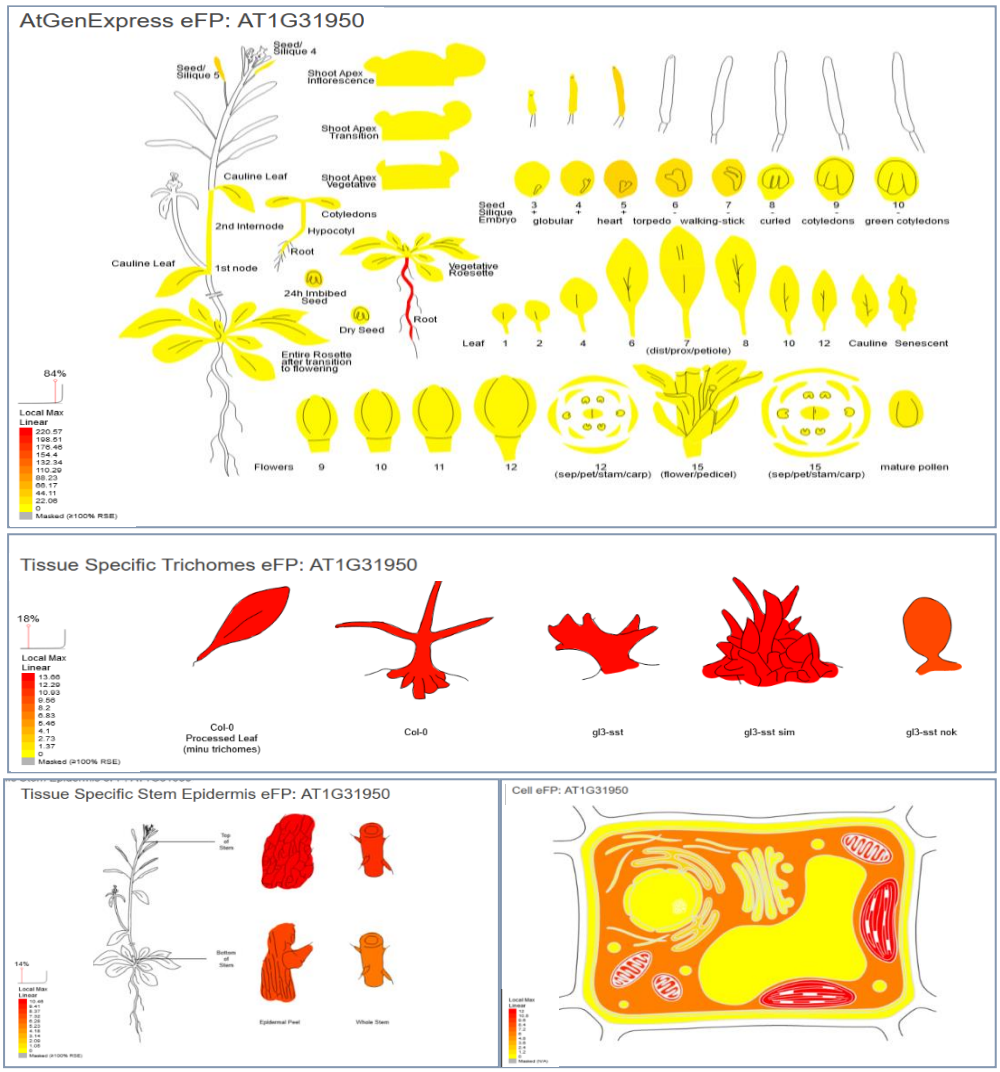

Fig. (3). Visualization of the putative "electronic fluorescent pictograph" browser for exploring the putative tissue expression and cell localization of *SgCMS* (AT1G31950) gene, based on *Arabidopsis* gene expression and protein localization at different tissues and cell organs. **a.** Expression data at different tissues from seedling to flowering stages. b Expression data tissue specific trichomes. **c.** Expression data of tissue specific stem epidermis at top and bottom. **d.** Expression data at different cell organs. The color box represents the expression scale (the more intense red color, the more gene expression).

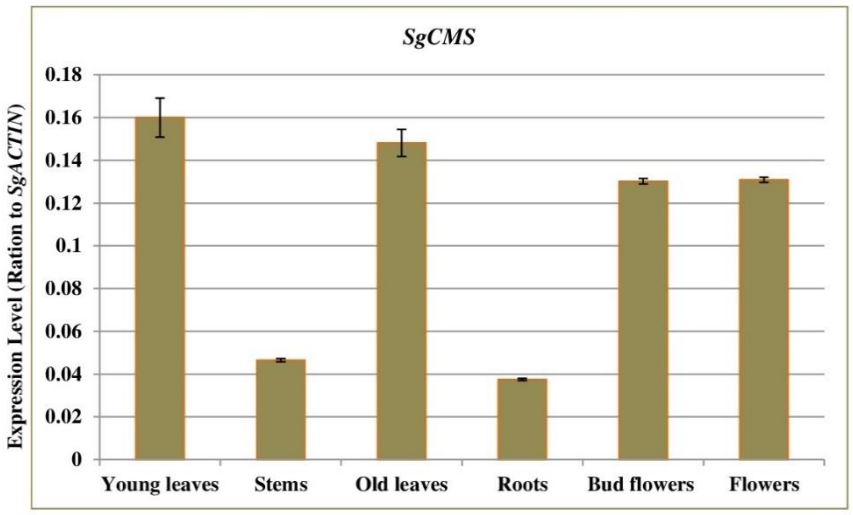

Fig (4). Quantitative RT-PCR validation of expression of *SgCMS* gene from *S. guaranitica*. Total RNAs were extracted from young leaves, stems, old leaves, roots, bud flowers and flower samples and the expression of *SgCMS* gene was analysed using quantitative real-time. *SgACTIN* was used as the internal reference. The values are means±SE of three biological replicates.

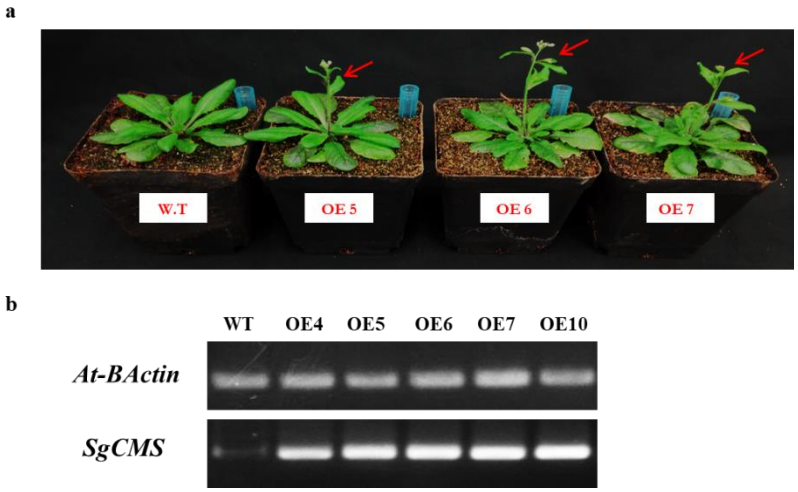

Fig. (5). Overexpression of *SgCMS* gene in transgenic *Arabidopsis*. **a.** Comparison of the phenotypes of the transgenic *A. thaliana* and wild type *A. thaliana*. The red arrow (↘) indicates the flowers at the transgenic *A. thaliana*. **b.** Semiquantitative RT-PCR to confirm the expression of terpenoid genes.

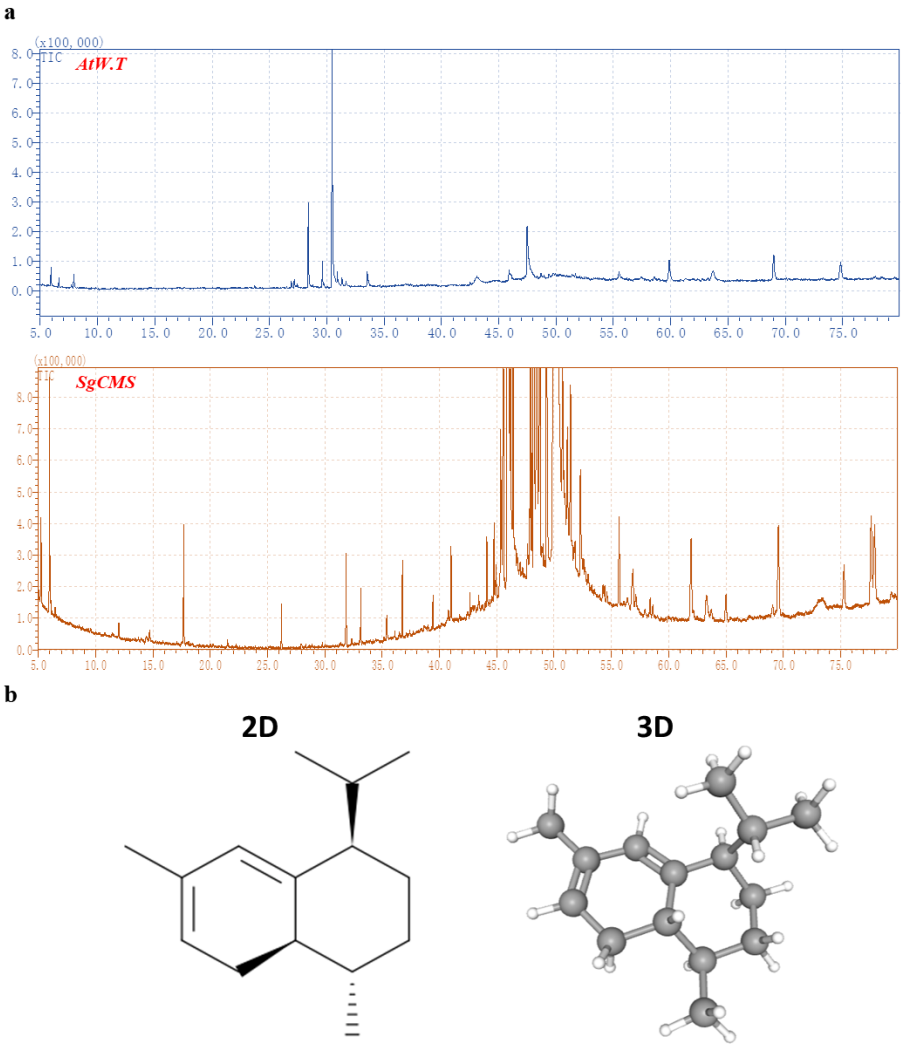

Fig (6). Typical GC-MS mass spectrographs for terpenoids from leaf of *A. thaliana* plants. **a.** GC-MS Peak of the essential oil, **b.** 2D and 3D chemical structure of the cis-muuroladiene.

Table (2). The major terpenoid compositions in transgenic *A. thaliana* leave over-expressing of *SgCMS*.

No.	Compound name	R.T. (min)	Formula	Molecular mass (g mol^{-1})	Type of terpene	% Peak area	
						AtW.T	*SgCMS*
1	o-Methyltoluene	5.233	C_8H_{10}	106.165	Aromatic	0.00	0.40
2	Alpha-Pinene	5.938	$C_{10}H_{16}$	136.234	Monoterpene	1.60	0.00
3	(8)Annulene	5.977	C_8H_8	104.1491		0.00	0.99
4	Utyl cellosolve	6.465	$C_6H_{14}O_2$	118.1742		0.00	0.09
5	Camphene	6.616	$C_{10}H_{16}$	136.234	Monoterpene	0.31	0.00
6	Beta-pinene	7.933	$C_{10}H_{16}$	136.234	Monoterpene	0.96	0.00
7	N-Decane	12.022	$C_{10}H_{22}$	142.2817	Alkane	0.00	0.10
8	Acetaldehyde, phenyl-	14.717	C_8H8O	120.1485		0.00	0.15
9	N-Tridecane	17.687	$C_{13}H_{28}$	184.3614	Alkane	0.00	0.53
10	Tar camphor	21.509	$C_{10}H_8$	128.1705	Aromatic	0.00	0.05
11	Isocaryophyllene	23.699	$C_{15}H_{24}$	204.3511	Sesquiterpene	0.11	0.00
12	Dodecamethylcyclo-hexasiloxane	26.212	$C_{12}H_{36}O_6Si_6$	444.9236		0.00	0.16
13	Humulene	27.17	$C_{15}H_{24}$	204.3511	Sesquiterpene	1.52	0.00
14	1(10),4(14),5-Germacratriene	28.394	$C_{15}H_{24}$	204.3511	Sesquiterpene	6.84	0.00
15	Gamma-elemene	29.611	$C_{15}H_{24}$	204.3511	Sesquiterpene	2.16	0.00
16	Palmitaldehyde	30.455	$C_{16}H_{32}O$	240.4247	Sesquiterpene	40.52	0.00
17	Tetradecamethylcyclo-heptasiloxane	31.852	$C_{14}H_{42}O_7Si_7$	519.0776	Sesquiterpene	0.00	0.35
18	Topanol; Stavox	33.135	$C_{15}H_{24}O$	220.3505	Sesquiterpene	0.00	0.21
19	Germacrene D-4-ol	33.53	$C_{15}H_{26}O$	222.37	Sesquiterpene	3.03	
20	Cyclohexyl isopropylphosphonofluori-date	35.402	$C_9H_{18}FO_2P$	208.2102		0.00	0.15
21	Hexadecamethylcyclooct-asiloxane	36.777	$C_{16}H_{48}O_8Si_8$	593.2315		0.00	0.29
22	Myristaldehyde	39.436	$C_{14}H_{28}O$	212.3715		0.00	0.10
23	Octadeamethyl-cyclononasiloxane	41.008	$C_{18}H_{54}O_9Si_9$	667.3855		0.00	0.28
24	Palmitaldehyde	42.665	$C_{16}H_{32}O$	240.4247		0.00	0.24
25	Diethyl[(Z)-1-ethyl-2-(methoxymethyl)-1-butenyl]borane	43.075	$C_{12}H_{25}BO$	196.14		6.52	0.00
26	Pentadecylic acid	43.423	$C_{15}H_{30}O_2$	242.3975		0.00	0.13
27	Linolenic acid	44.134	$C_{18}H_{30}O_2$	278.4296	Fatty acid	0.00	0.26
28	Hexadecamethylheptasi-loxane	44.766	$C_{16}H_{48}O_6Si_7$	533.1472		0.00	0.29

Table (2). Cont.

29	α-Linolenic acid	45.343	$C_{18}H_{30}O_2$	278.4296	Fatty acid	0.00	1.01
30	Retinol, acetate, all-trans-	45.556	$C_{22}H_{32}O_2$	328.4883		0.00	2.8
31	Levomenthol	45.906	$C_{10}H_{20}O$	156.2652	Monoterpene	3.35	0.00
32	N-Hexadecanoic acid	46	$C_{16}H_{32}O_2$	256.4241		0.00	14.36
33	All-trans-Retinol acetate	46.218	$C_{22}H_{32}O_2$	328.4883		0.00	1.24
34	Trans-Vitamin A acetate	46.426	$C_{22}H_{32}O_2$	328.4883		0.00	2.73
35	Benzoic acid, 2-(2-oxo-2-piperidin-1-yl-ethylsulfanyl)-	47.472	$C_{14}H_{17}NO_3S$	279.36		13.42	0.00
36	Trans-Muurol-3,5-diene	47.852	$C_{15}H_{24}$	204.3511	Sesquiterpene	0.00	0.25
37	Caryophyllene	47.947	$C_{15}H_{24}$	204.3511		0.00	1.74
38	All-trans-Geranylgeraniol	48.064	$C_{22}H_{36}O_2$	332.52	Diterpene	0.00	0.50
39	Cis-muurola-3,5-diene	48.221	$C_{15}H_{24}$	204.3511	Sesquiterpene	0.00	8.68
40	γ-Muurolene	48.473	$C_{15}H_{24}$	204.3511	Sesquiterpene	0.00	8.67
41	α-Farnesene	48.55	$C_{15}H_{24}$	204.3511	Sesquiterpene	0.00	2.49
42	Dihydro-α-ionone;	48.675	$C_{13}H_{22}O$	194.3132		0.42	0.73
43	Longifolene	48.774	$C_{15}H_{24}$	204.3511	Sesquiterpene	0.00	3.44
44	17-α-Methyl-5-β-androstan-3-β,17-β-dioll	49.345	$C_{20}H_{34}O_2$	306.4828	Diterpene	0.00	6.74
45	4,8,13-Duvatriene-1,3-Diol	49.913	$C_{20}H_{34}O_2$	306.4828	Diterpene	0.00	3.54
46	Linoleic	50.019	$C_{18}H_{32}O_2$	280.4455	Fatty acid	0.00	1.52
47	Cis-Muurol-5-en-4-α-ol	50.179	$C_{15}H_{26}O$	222.3663	Sesquiterpene	0.00	14.81
48	Oleic acid	50.309	$C_{18}H_{34}O_2$	282.461	Fatty acid	0.00	8.44
49	Stearic acid	50.73	$C_{18}H_{36}$	284.477	Fatty acid	0.00	1.47
50	Cycloartenol	51.163	$C_{30}H_{50}O$	426.7174	Triterpene	0.00	0.49
51	Octadeamethyl-cyclononasiloxane	51.436	$C_{18}H_{54}O_9Si_9$	667.3855		0.00	0.84
52	17-Pentatriacontene	51.827	$C_{35}H_{70}$	490.9303		0.00	0.37
53	Cis-muurola-4(14),5-diene	52.301	$C_{15}H_{24}$	204.3511	Sesquiterpene	0.00	0.71
54	8-Heptadecene	52.979	$C_{17}H_{34}$	238.4519		0.00	0.22
55	Docosanol	54.28	$C_{22}H_{46}O$	326.6	Fatty alcohol	0.00	0.38
56	Phthalic acid, heptyl oct-3-yl ester	55.504	$C_{23}H_{36}O_4$	376.5295		1.76	0.00
57	Octadeamethyl-cyclononasiloxane	55.675	$C_{18}H_{54}O_9Si_9$	667.3855		0.00	0.60
58	Ledane	56.375	$C_{15}H_{26}$	206.3669		0.00	0.23
59	4,8,13-Duvatriene-1,3-Diol	56.87	$C_{20}H_{34}O_2$	306.4828	Diterpene	0.00	0.72
60	Palmitic acid 3-[(trimethylsilyl)oxy]propyl ester	57.44	$C_{22}H_{46}O_3Si$	386.7		0.96	0.00
61	Oleamide	57.928	$C_{18}H_{35}NO$	281.4766		0.00	0.06
62	Nonadec-1-ene	58.379	$C_{19}H_{38}$	266.505		0.00	0.26
63	Cucurbitacin b, 25-desacetoxy-	58.556	$C_{30}H_{44}O_6$	500.7		1.11	0.00

Table (2). Cont.

64	N,N-Dimethylformamide dicyclohexyl acetal	59.861	$C_{15}H_{29}NO_2$	255.396		4.0	0.00
65	Octadeamethyl-cyclononasiloxane	61.954	$C_{18}H_{54}O_9Si_9$	667.3855		0.00	0.48
66	Nonadec-1-ene	63.302	$C_{19}H_{38}$	266.505		0.00	0.63
67	Eicosane, 2-methyl-	63.676	$C_{21}H_{44}$	296.5741		3.04	0.00
68	Phthalic acid dioctyl ester;	65	$C_{24}H_{38}O_4$	390.5561		0.00	0.31
69	Geranyl linallol	68.993	$C_{20}H_{34}O$	290.4834	Diterpene	4.19	0.00
70	1-Nonadecene	69.053	$C_{19}H_{38}$	266.505		0.00	0.16
71	Octadeamethyl-cyclononasiloxane	69.575	$C_{18}H_{54}O_9Si_9$	667.3855		0.00	0.75
72	2α,9-dihydroxy-1,8-cineole	73.439	$C_{10}H_{18}O_3$	186.2481	Monoterpene	0.00	0.73
73	6,11-Dimethyl-2,6,10-dodecatrien-1-ol	74.819	$C_{14}H_{24}O$	208.34		4.18	
74	N-Tetracontane	75.293	$C_{40}H_{82}$	563.0791		0.00	0.55
75	Octadeamethyl-cyclononasiloxane	77.672	$C_{18}H_{54}O9Si_9$	667.3855		0.00	0.42
76	Benzophenone-12	77.963	$C_{21}H_{26}O_3$	326.4293		0.00	0.92
77	Erucylamide	79.421	$C_{22}H_{43}NO$	337.5829		0.00	0.24
	Total % of monoterpene					6.22	0.73
	Total % of sesquiterpene					13.66	39.26
	Total % of diterpene					4.19	11.50
	Total % of triterpene					0.00	0.49
	Total % of aromatic					0.00	0.45
	Total % of alkane					0.00	0.63
	Total % of fatty acid					0.00	12.70
	Total % of fatty alcohol					0.00	0.38

DISCUSSION

1. Cloning and Sequence Analysis of *SgCMS* Gene from *S. guaranitica*

The full-length cDNA of *SgCMS* gene was recognized and isolated from leaves of *S. guaranitica*, based on the highly conserved sequence similarity between the query sequence and the other sequences detected in different plant species such as *Salvia splendens*, *Mentha x piperita*, *Ocimum basilicum, Mentha arvensis, Mentha asiatica, Lavandula angustifolia, Erythranthe guttata, Sesamum indicum* and *Handroanthus impetiginosus*. Compared to other sesquiterpene synthases, SgCMS protein has four domains specified by the InterPro database. The first domain is terpenoid cyclases/protein prenyltransferase alpha-alpha toroid (IPR008930: from 74 - 277), and the second is isoprenoid synthase domain (IPR008949: from 279 - 605), while the other domains are terpene synthase, metal-binding domain (IPR005630: from 284 - 547) and terpene synthase, N-terminal domain (IPR001906: from 75 - 253) as in Fig. (1).

A comparison between the putative SgCMS protein sequence and other plant sesquiterpene synthase revealed various highly conserved domains that are responsible for coordination of divalent metal ion co-factors and substrate binding (Ali et al., 2018; Abbas et al., 2019 and Rebecca et al., 2020) (Fig. 2). Moreover, they are also known for their role in binding a tri-nuclear-magnesium cluster, two-magnesium-ions and one-magnesium-ion (Christianson, 2006; Lima et al., 2013 and Abbas et al., 2019). This magnesium cluster binds and interacts with the diphosphate moiety of farnesyl diphosphate (FPP), hence, catalyzing the C15-substrate-FPP formation at the hydrophobic substrate binding pocket (Su-Fang et al., 2014; Abbas et al. 2019 and Rebecca et al. 2020). Eventually, each protein sequence belonging to the terpene synthase family has one or two of these conserved domains (Ali et al., 2017, 2018 and 2022a). A phylogenetic tree was generated to analyse the evolutionary relationship between *SgCMS* and other plant sesquiterpene synthase genes, and the evolutionary tree was constructed well by the neighbor-joining method parameters. Based on our classification results, the SgCMS protein was classified into the TPS-g subfamily that can encode mono-, sesqui- and diterpene, which explains the ability of SgCMS to produce various types of terpenes as previously mentioned by Ali et al. (2017 and 2018) (Fig. 2).

2. Putative Tissue Expression Pattern and Subcellular Localizations of *SgCMS* Gene

To identify the physiological roles of *SgCMS*, its putative expression was explored in fifty six tissues. This was aided by the high resemblance between *SgCMS,* and AT1G31950 gene from *A. thaliana*. *SgCMS,* gene was recognized in the tested tissues, and this result was in line up with Ali et al. (2017, 2018 and 2022a), who reported that most TPS genes (e.g. *SoFLDH, SgTPSV, SgGERIS, SoLINS2, GmTPS21, SgFARD, SoNEOD* and *SoHUMS*) from *S. officinalis*, *Glycine max*, and *S. guaranitica* showed an increased expression in leaves, roots and seeds. Furthermore, the putative subcellular localization for SgCMS protein revealed that the gene under study is present at most in the cytosol, then plastid, mitochondria and nucleus. The gained results are similar to Wang et al. (2015b), Chen et al. (2018) and Makhadmeh et al. (2022a), who found that various TPSs genes were localized in the plastid, mitochondria, and nucleus (Fig. 3).

3. Effectiveness of qRT-PCR for Analysis of the *SgCMS* Gene Expression

qRT-PCR was used to measure the expressed levels of *SgCMS*. Therefore, qPCR data revealed that *SgCMS* is highly expressed in young leaves followed by old leaves, flowers, bud flowers, stems, and roots. This fits with the previously provided information that found various TPS syntheses genes were highly expressed in young leaves, bud flowers, old leaves, and stems (Croteau et al., 1981; Sabin et al., 2012; Ali et al., 2018). The low expression of *SgCMS* gene in stems and roots is probably due to many reasons

such as gene-regulatory mechanisms (possible posttranscriptional and/or post-translational) (Fig. 4). It is also possible that the expression levels of *SgCMS* is associated with tissue-developmental-stages, which strongly influences the expression of terpene synthase in *Salvia* (Croteau et al., 1981; Sabin et al., 2012 and Ali et al., 2018).

4. Overexpression of *SgCMS* Gene Changed the *A. thaliana* Plants Phenotypes

To evaluate the function of *SgCMS* in *A. thaliana* (Col-0: Columbia-0) plants. *SgCMS* was overexpressed in *A. thaliana* through *Agrobacterium* bacteria harboring the overexpression vector pB2GW7-*SgCMS*. After that, sqRT-PCR confirmed the expression of the target gene in positive transgenic lines (Fig. 5). The transgenic lines showed a higher expression level of the *SgCMS* gene in parallel with the wild type, which indicated the existence of the target gene in the transgenic plants. After that, three transgenic plants (named; OE-*SgCMS*-5, OE-*SgCMS*-6, and OE-*SgCMS*-7) were chosen for terpene analysis. The morphological analysis showed that the previous transgenic plants had an accelerated rate of flowering stem formation, unlike the wild-type plants **(**Fig. 5). These earn positive results are in line with Ali et al. (2017, 2018 and 2022a and b), where the overexpression of TPS synthesis and terpenoids genes, such as *SgGPS, SgLINS, SgFPPS, SoTPS6, SoCINS, SoLINS, SoFLDH, SoSABS, SgGPS* and *SoNEOD* from *S. guaranitica* and *S. officinalis* in *A. thaliana* and *Nicotiana tabacum*, accelerated the growth and flower formation when compared to wild type plants. In previous results, numerous TPSs family genes played a crucial role in different cell-specific processes, such as: 1, 8-cineole, Z-γ-bisabolene, rhizathalene, β-amyrin and thalianol synthesis as a mono-,sesqui-, di- and triterpene, respectively (Ro et al., 2006; Vaughan et al., 2013 and Wang et al., 2016). This means these genes have the ability to co-expressed initially in diverse cells, tissues, and organs for producing a different plant phenotype and confirming the role of TPSs-genes in plant growth, flowering, and development (Kampranis et al., 2007; Field and Osbourn, 2008; Chen et al., 2011). GC-MS-system analyzed the metabolites to recognize the specific terpenes synthesized after introducing the *SgCMS* gene into *Arabidopsis* plants and generating over-expressing transformant lines. The mono-, sesqui-, and diterpene peaks were easily visible; the percentage of peak area (% peak area) demonstrated the type and amount of the compounds. To identify these terpenes in the transgenic *Arabidopsis* plants, the libraries of mass spectra and the previous extracts of wild-type *Arabidopsis* was used as a reference, which produce different quantitative and qualitative terpenoids. The results shown in Fig. (6) and Table (2), an apparent alteration was observed in the transgenic plants, and a new peak at retention time (50.179) was detected. This peak was identified as cis-muurol-5-en-4-α-ol, based on the matched mass with the Wiley GC/MS, NIST

Library, and VOC Analysis S/W software. On the other hand, the production of sesquiterpene by the overexpression of sesquiterpene synthase genes in *A. thaliana* was described formerly by Su-Fang et al. (2014) and Ali et al. (2018). Various terpene synthase genes are known to synthesize various products in unison, e.g. carene, (±)-linalool, cineole, myrcene, β-amyrin and terpinolene synthases (Yoko et al., 2004; Shimada et al., 2005 and Xi et al., 2016). It is believed that *SgCMS* was responsible for the production of cis-muurol-5-en-4-α-ol via isoprenoid pathway which is common in sesquiterpene biosynthesis.

CONCLUSIONS

S. guaranitica is a valuable Chinese medicinal herb with distinctive pharmacological effects. As a result, cloning and characterizing several genes involved in secondary metabolic pathways will help metabolic engineering in *S. guaranitica* and other therapeutic plants. In this study, *SgCMS* gene was cloned as a plant cis-muuroladiene biosynthetic from *S. guaranitica*. *SgCMS* overexpression in *A. thaliana* enhanced the accelerated of flowering axis formation in the *OE-SgCMS-5*, *OE-SgCMS-6*, and *OE-SgCMS-7* transgenic lines. Compared to the control, these previous lines produced a lot of cis-muuroladiene. cis-muuroladiene in these transgenic lines demonstrates that *A. thaliana* plants may synthesize the same product via the common mevalonate pathway (MVK) of sesquiterpene biosynthesis. Additionally, *SgCMS* gene was highly expressed in different plant tissues and was primarily localized in the mitochondria and cytosol, underscoring the potential role of this gene in producing various types of terpenes, particularly the sesquiterpene cis-muuroladiene. This study showed that the sesquiterpene gene might be studied using *A. thaliana* as a model plant, which can be used to improve the essential oil composition in *S. guaranitica* and other plant species through metabolic engineering.

ACKNOWLEDGEMENT

The author thanks Professor Dr. Osama Ezzat Elsayed and Dr. Ahmed Ali for proofreading the manuscript. He also thanks Dr. Wael Moussa for his constructive comments and help. The author also thanks all Molecular Genetics Lab (Prof. Dr. Osama Ezzat Elsayed lab) members at National Research Centre, Dokki, Cairo, Egypt, for facilitating the practical work at all levels of the experiments. The author would like to thank the Desert Research Center labs.

REFERENCES

Abbas, F., K. Yanguo, Y. Rangcai and F. Yanping (2019). Functional characterization and expression analysis of two terpene synthases

involved in floral scent formation in *Lilium* 'Siberia.' Planta, 249: 71–93.

Aharoni, A., A.P. Giri, S. Deuerlein, F. Griepink, W.J. de Kogel, F.W. Verstappen, H.A. Verhoeven, M.A. Jongsma, W. Schwab and H.J. Bouwmeester (2003). Terpenoid metabolism in wild-type and transgenic *Arabidopsis* plants. Plant Cell, 15 (12): 2866-2884.

Ali, M., P. Li, G. She, D. Chen, X. Wan and J. Zhao (2017). Transcriptome and metabolite analyses reveal the complex metabolic genes involved in volatile terpenoid biosynthesis in garden sage (*Salvia officinalis*). Sci Rep., 7 (1): 16074.

Ali, M., R.M. Hussain, N.U. Rehman, G. She, P. Li, X. Wan, L. Guo and J. Zhao (2018). *De novo* transcriptome sequencing and metabolite profiling analyses reveal the complex metabolic genes involved in the terpenoid biosynthesis in Blue Anise Sage (*Salvia guaranitica* L.). DNA Res., 25 (6): 597-617.

Ali, M., L. Miao, Q. Hou, D.B. Darwish, S.S. Alrdahe, A. Ali, V.A. Benedito, M. Tadege, X. Wang and J. Zhao (2021). Overexpression of terpenoid biosynthesis genes from garden sage (*Salvia officinalis*) modulates rhizobia interaction and nodulation in soybean. Front. Plant Sci., 12: 783269.

Ali, M., D. Alshehri, A.M. Alkhaibari, N.A. Elhalem and D.B.E. Darwish (2022a). Cloning and characterization of 1,8-cineole synthase (*SgCINS*) gene from the leaves of *Salvia guaranitica* plant. Front. Plant Sci., 13: 869432.

Ali, M., E. Nishawy, W.A. Ramadan, M. Ewas, M.S. Rizk, A.G.M. Sief-Eldein, M.A.S. El-Zayat, A.H.M. Hassan, M. Guo, G.W. Hu, S. Wang, F.A. Ahmed, M.H. Amar and Q.F. Wang (2022b). Molecular characterization of a novel NAD+-dependent farnesol dehydrogenase *SoFLDH* gene involved in sesquiterpenoid synthases from *Salvia officinalis*. PLoS One, 17 (6): e0269045.

Ali, M., L. Miao, F.A. Soudy, D.B.E. Darwish, S.S. Alrdahe, D. Alshehri, V.A. Benedito, M. Tadege, X. Wang and J. Zhao (2022c). Overexpression of terpenoid biosynthesis genes modifies root growth and nodulation in soybean (*Glycine max*). Cells, 11 (17): 2622.

Anders, S. and W. Huber (2010). Differential expression analysis for sequence count data. Genome Biol., 11 (10): R106.

Bohlmann, J., G. Meyer-Gauen and R. Croteau (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. USA, 95 (8): 4126–4133.

Chen, F., D. Tholl, J. Bohlmann and E. Pichersky (2011). The family of terpene synthases in plants: A mid-size family of genes for

specialized metabolism that is highly diversified throughout the kingdom. Plant J., 66: 212–229.

Chen, X., H. Chen, J.S. Yuan, T.G. Köllner, Y. Chen, Y. Guo, X. Zhuang, X. Chen, Y.J. Zhang, J. Fu, A. Nebenführ, Z. Guo and F. Chen (2018). The rice terpene synthase gene *OsTPS19* functions as an (S)-limonene synthase in planta, and its overexpression leads to enhanced resistance to the blast fungus *Magnaporthe oryzae*. Plant Biotechnol. J., 16 (10): 1778–1787.

Christianson, D.W. (2006). Structural biology and chemistry of the terpenoid cyclases. Chemical Reviews, 106 (8): 3412–3442.

Croteau, R., M. Felton, F. Karp and R. Kjonaas (1981). Relationship of camphor biosynthesis to leaf development in sage *Salvia officinalis*. Plant Physiol., 67: 820–824.

Darwish, D.B.E, M. Ali, A.M. Abdelkawy, M. Zayed, M. Alatawy and A. Nagah (2022). Constitutive overexpression of *GsIMaT2* gene from wild soybean enhances rhizobia interaction and increase nodulation in soybean (*Glycine max*). BMC Plant Biol., 22 (1): 431.

Dereeper, A., V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J.F. Dufayard, S. Guindon, V. Lefort, M. Lescot, J.M. Claverie and O. Gascuel (2008). Phylogeny. fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res., 36: W465–W469.

Dorothea, T., B. Wilhelm, H. Armin, L. Francesco, S.R.R. Ursula and S. J. Peter (2006). Practical approaches to plant volatile analysis. The Plant Journal, 45: 540–560.

Elsherbeny, E.A., M. Ali, F.A. El-Ramah and M.K. Ahmed (2022). Molecular cloning and 571 characterization of terpene synthase 4 (SgTPS4) gene from *Salvia guaranitica* plant. Egypt. J. Genet. Cytol., 51: 1-20.

Fahmy A.E., M. Ali, E.A. Elsherbeny and M.K. Ahmed (2022). Molecular cloning and characterization of 575 beta-amyrin synthase (SoAMYS) gene from *Salvia officinalis* plant. Egyptian J. Desert Res., 72 (1): 27-45.

Fateme, A.M., H.F. Mohammad, R. Abdolhossein, Z. Ali and S. Maryam (2013). Volatile Constituents of *Salvia compressa* and *Logochilus macranthus*, two Labiatae herbs growing wild in Iran. Res. J. Recent Sci., 2: 66–68.

Field, B. and A.E. Osbourn (2008). Metabolic diversification–independent assembly of operon-like gene clusters in different plants. Science, 320: 543–547.

Gershenzon, J. and W. Kreish (1999). Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes, Diterpenes, Sterols, Cardiac Glycosides and Steroid Saponins. In: 'Wink, M. Ed.' Biochemistry of Plant Secondary Metabolism. Florida, CRC Press., pp. 222–299.

Hussain, R.M., M. Ali, X. Feng and X. Li (2017). The essence of NAC gene family to the cultivation of drought-resistant soybean (*Glycine max* L. Merr.) cultivars. BMC Plant Biol., 17 (1): 55.

Jiang, S.Y., J. Jin, R. Sarojam and S. Ramachandran (2019). A Comprehensive Survey on the Terpene Synthase Gene Family Provides New Insight into Its Evolutionary Patterns. Genome Biology and Evolution, 11 (8): 2078–2098.

Kampranis, S.C., D. Ioannidis, A. Purvis, W. Mahrez, E. Ninga, N.A. Katerelos, S. Anssour, J.M. Dunwell, J. Degenhardt, A.M. Makris, P.W. Goodenough and C.B. Johnson (2007). Rational conversion of substrate and product specificity in a *Salvia* monoterpene synthase: structural insights into the evolution of terpene synthase function. Plant Cell, 19 (6): 1994-2005.

Li, D., F. Shao and S. Lu (2015). Identification and characterization of mRNA-like noncoding RNAs in *Salvia miltiorrhiza*. Planta, 241: 1131–1143.

Lima, A.S., S. Jette, L. Brigitte, N. Johannes, G.B. Jose´, A.C. Figueiredo, G.P. Luis, D. Jo¨rg, and T. Helena (2013). Genomic characterization, molecular cloning and expression analysis of two terpene synthases from *Thymus caespititius* (Lamiaceae). Planta, 238: 191–204.

Makhadmeh, I.M., S.G. Thabet, M. Ali, B. Alabbadi, A. Albalasmeh and A.M. Alqudah (2022a). Exploring genetic variation among Jordanian *Solanum lycopersicon* L. landraces and their performance under salt stress using SSR markers. J. Genet. Eng. Biotechnol., 20 (1): 45.

Makhadmeh, I., A.A. Albalasmeh, M. Ali, S.G. Thabet, W.A. Darabseh, S. Jaradat and A.M. Alqudah (2022b). molecular characterization of tomato (*Solanum lycopersicum* L.) accessions under drought stress. Horticulturae, 8 (7): 600.

Mehmood, N., Y. Yuan, M. Ali, M. Ali, J. Iftikhar, C. Cheng, M. Lyu and B. Wu (2021). Early transcriptional response of terpenoid metabolism to *Colletotrichum gloeosporioides* in a resistant wild strawberry *Fragaria nilgerrensis*. Phytochemistry, 181: 112590.

Rebecca, S.W., M.A. Ayelign, B. Lina, N. Elaheh and S.M. Soheil (2020). Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia*. Planta, 251: 41.

Rehman, N.U., M. Ali, M.Z. Ahmad, G. Liang and J. Zhao (2017). Strigolactones promote rhizobia interaction and increase nodulation in soybean (*Glycine max*). Microb Pathog., 114: 420-430.

Ro, D.K., J. Ehlting, C.I. Keeling, R. Lin, N. Mattheus and J. Bohlmann (2006). Microarray expression profiling and functional characterization of *AtTPS* genes: Duplicated *Arabidopsis thaliana* sesquiterpene synthase genes *At4g13280* and *At4g13300* encode

root-specific and wound-inducible (Z)-γ-bisabolene synthases. Arch. Biochem. Biophys., 448: 104–116.

Ruzicka, L. (1953). The isoprene rule and the biogenesis of terpenic compounds. Experientia, 9: 357-367.

Ruzicka, L. (1973). In the borderland between bioorganic chemistry and biochemistry, Annu. Rev. Biochem., 42: 1-20.

Sabine, G.G., S. Corinna, S. Ralf and N. Johannes (2012). Seasonal influence on gene expression of monoterpene synthases in *Salvia officinalis* (Lamiaceae). J. Plant Physiol., 1 (69): 353–359.

Sarrou, E., G. Ioannis, X. Aliki, M. Domenico, M. Stefan, M. Panagiotis, M. Athanasios and C. Paschalina (2017). Genetic diversity and metabolic profile of *Salvia officinalis* populations: implications for advanced breeding strategies. Planta, 246: 201–215.

Shimada, T., T. Endo, H. Fujii, M. Hara and M. Omura (2005). Isolation and characterization of (E)-beta-ocimene and 1, 8 cineole synthases in *Citrus unshiu* Marc. Plant Sci., 168: 987–995.

Su-Fang, E., Z.A. Mohamed-Hussein, R. Othman, N.A. Shaharuddin, I. Ismail and Z. Zainal (2014). Functional characterization of sesquiterpene synthase from *Polygonum minus*. The Scientific World Journal, DOI: 10.1155/2014/840592.

Trapp, S. and R. Croteau (2001). Defensive resin biosynthesis in conifers. Annu. Rev. Plant Physiol. Plant Mol. Biol., 52: 689-724.

Vaughan, M.M., Q. Wang, F.X. Webster, D. Kiemle, Y.J. Hong, D.J. Tantillo, R.M. Coates, A.T. Wray, W. Askew, C. O'Donnell, J.G. Tokuhisa and D. Tholl (2013). Formation of the unusual semivolatile diterpene rhizathalene by the *Arabidopsis* class I terpene synthaseTPS08 in the root stele is involved in defense against belowground herbivory. Plant Cell, 25: 1108–1125.

Wang, B., W. Sun, Q. Li, Y. Li, H. Luo, J. Song, C. Sun, J. Qian, Y. Zhu, A. Hayward, H. Xu, and S. Chen (2015a). Genome-wide identification of phenolic acid biosynthetic genes in *Salvia miltiorrhiza*. Planta, 241 (3): 711–725.

Wang, B., S. Wei, L. Qiushi, L. Ying, L. Hongmei, S. Jingyuan, S. Chao, Q. Jun, Z. Yingjie, H. Alice, X. Haibin and C. Shilin (2015b). Genome-wide identification of phenolic acid biosynthetic genes in *Salvia miltiorrhiza*. Planta, 241: 711–725.

Wang, Q., M. Jia, J.H. Huh, A. Muchlinski, R.J. Peters and D. Tholl (2016). identification of a dolabellane type diterpene synthase and other root-expressed diterpene synthases in *Arabidopsis*. Front. Plant Sci., 7: 1761.

Xi, i., R. Lorenzo, L. Xiuli and X. De-Yu (2016). Overexpression of a synthetic insect–plant geranyl pyrophosphate synthase gene in *Camelina sativa* alters plant growth and terpene biosynthesis. Planta, 244: 215–230.

Yoko, I., R.G. David, F. Eyal, L. Efraim and P. Eran (2004). Characterization of geraniol synthase from the peltate glands of sweet basil. Plant Physiology, 134: 370–379.

Zhenqing, B., L. Wenrui, J. Yanyan, Y. Zhiyong, J. Jie, H. Wenli, X. Pengguo and L. Zongsuo (2018). The ethylene response factor *SmERF6* co-regulates the transcription of *SmCPS1* and *SmKSL1* and is involved in tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots. Planta, 248: 243–255.

الاستنساخ والتوصيف الجزيئي والتحليل الوظيفي لجين Cis-muuroladiene synthase (*SgCMS*) من أوراق نبات السالفيا جوارانيتيكا

محمد علي
وحدة الوراثة والسيتولوجي، قسم الأصول الوراثية، مركز بحوث الصحراء، المطرية، القاهرة، مصر

من حيث القيمة الغذائية والطبية، تم تصنيف سالفيا جوارانيتيكا (*Salvia guaranitica*) من بين أهم النباتات الطبية والعطرية بسبب وفرة المكونات النشطة الأساسية التي تحتوي عليها. تعتبر مركبات التربينويد، وخاصة المونوتربين (C_{10}) والسيسكيتيربين (C_{15})، أهم هذه المكونات النشطة وأكثرها وفرة. تلعب التربينات مجموعة متنوعة من المهام وتمتلك سمات بيولوجية مفيدة في النباتات. مع الأخذ في الاعتبار هذه الخصائص، تم إجراء البحث الحالي لاستنساخ جين الـ Cis-muuroladiene synthase (*SgCMS*); EC: 4.2.3.67 من نبات *S. guaranitica*. يحتوي بروتين الأحماض الأمينية البالغ ٦٠٠ حامض أميني والمشفر بواسطة *SgCMS* على إطار قراءة كامل مفتوح من ١.٨٠٢ زوجًا من القواعد. علاوة على ذلك، يحتوي بروتين *SgCMS* على خمس مجالات (domains) من عائلة terpene synthase تنتمي إلى مجالات terpenes و terpenoid synthase. تم التعديل والإفراط في التعبير عن جين *SgCMS* في نبات ال *Arabidopsis thaliana* من أجل التحقق من وظيفة الجين. مقارنة بالنباتات البرية، أظهرت نباتات ال *A. thaliana* المحورة جينيًا انخفاضًا في قطر الورقة، بينما تبدأ السيقان المزهرة في النمو مبكرًا مقارنةً بمجموعة الكنترول (GUS). في النهاية، كشفت نتائج الفحص اللوني للغاز- قياس الطيف الكتلي (GC-Mass) للنباتات المعدلة وراثيًا أن جين ال *SgCMS* كان مسؤول عن إنتاج عدة أنواع من التربين، وخاصة السيسكيتيربين.