

Biochemical analysis of the TPS-a subfamily in *Medicago truncatula*

Hannah Hendrickson¹, Monirul Islam², Ghislain F. Wabo³, Sibongile Mafu^{1*}

¹Plant Biology Graduate Program, University of Massachusetts Amherst, United States, ²Department of Biochemistry and Molecular Biology, United States, ³Department of Organic Chemistry, University of Yaoundé I, Cameroon

Submitted to Journal:
Frontiers in Plant Science

Specialty Section:
Plant Metabolism and Chemodiversity

Article type:
Original Research Article

Manuscript ID:
1349009

Received on:
04 Dec 2023

Revised on:
17 Jan 2024

Journal website link:
www.frontiersin.org

Scope Statement

Dear Editors Frontiers in Plant Science We are submitting the manuscript entitled “Biochemical analysis of the TPS-a subfamily in *Medicago truncatula*” for your consideration of publication in Frontiers in Plant Science: Plant Metabolism and Chemodiversity: Special issue Plant Secondary Metabolite Biosynthesis. In this study, we report on our comprehensive biochemical and molecular characterization of the terpene synthases from the TPS-a subfamily in *Medicago truncatula*. Our study defines gene-metabolite relationships in *Medicago* linking knowledge to previous literature but more importantly reveals new enzymes and metabolites. Combined the data defines the chemical landscape and demonstrates continued functional divergence resulting in species specific blends. We believe our findings are important in assessing the terpene chemical diversity in *Medicago* and becomes a useful resource in comparing to other legumes. We thank you for your time and effort in handling our manuscript. We look forward to hearing from you. Sibongile Mafu

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

CRedit Author Statement

Ghislain Fotso Wabo: Formal Analysis, Writing - review & editing. Hannah Hendrickson: Investigation, Methodology, Writing - original draft, Writing - review & editing. Monirul Islam: Investigation, Writing - original draft. Sibongile Mafu: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing - original draft, Writing - review & editing.

Keywords

Medicago truncatula, terpene synthase, Sesquiterpenes, biochemical analysis, Chemical diversity

Abstract

Word count: 122

Terpenes are important mediators of plant chemical response to environmental cues. Here, we describe the genome-wide identification and biochemical characterization of TPS-a members in *Medicago truncatula*, a model legume crop. Genome mining identified thirty-nine full-length terpene synthases with a significant number predicted to produce monoterpenes and sesquiterpenes. Biochemical characterization of the TPS-a subfamily associated with sesquiterpene biosynthesis revealed such compounds, that exhibit substantial biological activity in other plants. Gene expression analysis using qPCR and the *Medicago* gene atlas illustrated distinct tissue and time-based variation in expression in leaves and roots. Together our work establishes the gene-to-metabolite relationships for sesquiterpene synthases in *M. truncatula*. Understanding the biosynthetic capacity is a foundational step to defining the ecological roles of this important family of compounds.

Funding information

Financial support was provided by the Department of Biochemistry and Molecular Biology at the University of Massachusetts -Amherst; Richard and Susan Smith Family Foundation and National Science Award 2300060 to SM. HH was supported by the National Research Service Award T32 GM135096 from the National Institutes of Health.

Funding statement

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

Ethics statements***Studies involving animal subjects***

Generated Statement: No animal studies are presented in this manuscript.

Studies involving human subjects

Generated Statement: No human studies are presented in the manuscript.

Inclusion of identifiable human data

Generated Statement: No potentially identifiable images or data are presented in this study.

In review

Data availability statement

Generated Statement: The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

In review

Title: Biochemical analysis of the TPS-a subfamily in *Medicago truncatula*

Hannah Hendrickson¹, Monirul Islam², Ghislain Fotso Wabo³ and Sibongile Mafu^{1, 2*}

¹Plant Biology Graduate Program, University of Massachusetts-Amherst

²Department of Biochemistry and Molecular Biology, University of Massachusetts – Amherst
240 Thatcher Way, Life Science Laboratories, Amherst Massachusetts 01003, USA

³Department of Organic Chemistry, University of Yaoundé 1, Cameroon

*Corresponding Author: Sibongile Mafu

smafu@umass.edu

Abstract

Terpenes are important mediators of plant chemical response to environmental cues. Here, we describe the genome-wide identification and biochemical characterization of TPS-a members in *Medicago truncatula*, a model legume crop. Genome mining identified thirty-nine full-length terpene synthases with a significant number predicted to produce monoterpenes and sesquiterpenes. Biochemical characterization of the TPS-a subfamily associated with sesquiterpene biosynthesis revealed such compounds, that exhibit substantial biological activity in other plants. Gene expression analysis using qPCR and the *Medicago* gene atlas illustrated distinct tissue and time-based variation in expression in leaves and roots. Together our work establishes the gene-to-metabolite relationships for sesquiterpene synthases in *M. truncatula*. Understanding the biosynthetic capacity is a foundational step to defining the ecological roles of this important family of compounds.

Keywords: *Medicago truncatula*, terpene synthase, sesquiterpenes, biochemical analysis

INTRODUCTION

Plants produce an array of specialized metabolites in response to environmental stressors. Terpenoids are one such class of natural products with unprecedented chemical diversity. Biosynthesis proceeds from simple core steps which initiate from five-carbon units to make varying lengths of backbones (Figure 1). Terpene synthases (TPS) increase chemical diversity by conversion of precursors into hydrocarbons through complex lysis and allylic diphosphate ester bond cleavage [1]. The resulting compounds are structurally related forming variably sized and species-specific gene families [2].

Phylogenetic analysis of TPSs results in the clustering of genes into seven subfamilies based on sequence homology, gene architecture and functional studies in various plants [3]. Generally, sesquiterpene synthases (C15) fall primarily into TPS-a, monoterpene synthases (C10) in TPS-b and TPS-g, and diterpene synthases (C20) in TPS-c and TPS-e/f whereas TPS-d and TPS-h are specific to gymnosperms and lycophytes respectively. This vast chemical diversity results in species-specific complex mixtures of terpenes with broadly varying biological activity. Terpenoids mediate interactions between plants and their environment and are critical in coordinating responses to (a)biotic stress [2]. Induced volatile terpene blends composed mainly of lower molecular weight monoterpenes and sesquiterpenes have additional functions in inter and intra-plant communication [4].

Legumes are important crops providing a significant portion of protein in human and animal diets and play an important role in agricultural sustainability. Early investigations in *M. truncatula*, have demonstrated that differential terpene-enriched volatile blends are induced in response to herbivores [5, 6]. Subsequent studies identified biosynthetic genes underlying these observed metabolites summarized in Figure 1 [6-9]. *MtTPS1* produces a sesquiterpene, (*E*)- β caryophyllene, which in maize is instrumental in indirect defense by attracting parasitoids of herbivores and entomopathogenic nematodes [10, 11]. *MtTPS3* forms nerolidol and geranylinalool common to many plants with defense capabilities. In *M. truncatula*, *MtTPS4* which encodes for a monoterpene β -ocimene, was upregulated in response to methyl jasmonate and feeding insects [8]. β -ocimene has a priming effect that activates a defense response in Chinese cabbage and cucumber [4, 8, 12]. *MtTPS5* is a multiproduct TPS, producing up to twenty-seven different terpene products [7]. More recently, Yadav *et al.* (2019) showed that *MtTPS10*, which catalyzes the formation of a sesquiterpene alcohol, himachalol, was upregulated in the root upon infection by an oomycete pathogen, *Aphanomyces euteiches* [13]. The combined literature illustrates that *M. truncatula* produces a blend of terpene-based compounds upon interactions with biotic stressors as part of its chemical response strategy.

Soybean terpene synthases *GmTPS3*, *GmTPS18* and *GmTPS21* have been characterized and shown to produce β -ocimene, geraniol and α -farnesene respectively, each of which have demonstrated protective roles in plants [14-16]. For example, *GmTPS21* (which encodes for α -farnesene) provided resistance against soybean cyst nematode belowground and aphids aboveground [16, 17]. In cowpea, terpenes were identified as part of the volatiles produced in response to herbivory whereas biochemical characterization of the genes responsible for floral scent was carried out in sweet pea [18, 19].

In grasses, such as maize, rice and wheat, terpenes play an important role in the interactions of plants with their environment [10, 20, 21]. Although terpenes are widespread there still exist distinct chemotypes in species with new chemistry and biological roles which are still being uncovered [2, 22-24]. Comparatively less is known about the composition of the terpenome in legumes.

Here, we define the gene-to-metabolite relationships underlying terpene biosynthesis in *M. truncatula*. Through genome-wide identification of terpene synthases and biochemical characterization of the TPS-a subfamily, we have identified sesquiterpene synthases essential in generating the sesquiterpene-based chemical diversity in *M. truncatula*. Understanding the terpenoid biosynthetic capacity is a foundational step to defining the ecological roles of this important family of compounds.

MATERIALS AND METHODS

MtTPS phylogenetic and sequence analysis

Putative terpene synthases (protein models) from the *Medicago truncatula* cv. Jemalong A17 were collated from the Mt4.0v1 genome in Phytozome (<https://phytozome-next.jgi.doe.gov>) through an annotation-based search for TPSs. The putative TPSs were further manually curated for the presence of catalytic domains namely the Class I site with characteristic DDxxD and secondary NSE/DTE metal binding motif involved in ionization-dependent cyclization. Class II synthases were identified by the DxDD motif involved in protonation-initiated cyclization characteristic of copalyl diphosphate synthases (CPPs) as shown in Figure S1 [1]. For phylogenetic analysis, previously reported and biochemically characterized genes were collated from *S. lycopersicum* ITAG2.4 and *A. thaliana* TAIR10 genomes [25], aligned with *M. truncatula* TPSs using maximum likelihood analysis with 1000 bootstrap repetitions using CLC software and visualized using iTOL [26].

Combinatorial expression in *Escherichia coli*.

Synthetic MtTPSs codon-optimized for expression in *Escherichia coli* were ordered through ThermoFisher and subcloned into pET28b using primers listed in Table S1. Functional analysis of genes was carried out using a previously described modular metabolic engineering system in *E. coli* enhanced for sesquiterpene production [27]. Briefly, TPS-a family genes in pET28b were co-transformed with an (*E, E*)-FPP synthase from *Zea mays* (ZmFPPs) in *E. coli* BL21 DE3-C41 cells (Lucigen). Transformed cultures were grown in 45 mL of Terrific Broth medium to an OD₆₀₀ of ~0.6 at 37°C. Cultures were cooled to 16°C before induction with 1 mM isopropyl-thiogalactopyranoside and supplemented with 1 mM MgCl₂ and 25 mM sodium pyruvate incubation for 72 hours. Dodecane (20 mL) was added to ensure the capture of volatile compounds. In a second flask, enzyme products were extracted with 50 mL of 100% hexane and concentrated under an N₂ stream. Samples were resuspended in 1 mL of *n*-hexane for analysis by gas chromatography–mass spectrometry (GC-MS).

GC-MS analysis

Samples were analyzed on an Agilent 8890B gas chromatograph (GC) coupled to a mass spectrometer (MS) 5977B extractor electron ionization detector at 70 eV. Samples (1 µL) were injected in pulsed splitless mode with the inlet temperature set to 250°C. Separation was achieved

on an HP5MS column (30 m, 0.25 mm i.d., 0.25 μ m film) using helium as the carrier gas at a flow rate of 1.2 mL/min. The initial oven temperature of 70°C was increased after 1 min to 260°C at a rate of 10°C/min and held for 2 min at 260°C. MS data was collected from 60 to 500 mass-to-charge ratio(m/z). Compounds were identified by comparison of mass spectra and retention times with those of the authentic standards, when available, or with the Wiley, National Institute of Standards and Technology (NIST) mass spectral library.

Product purification for NMR analysis.

To obtain sufficient compound for structural elucidation by NMR, cultures were scaled up to 500 mL batches to a total volume of 5 Liters and expressed using the combinatorial expression system described above [27]. The resultant products were extracted with an equal volume of hexanes. The organic extract was recovered using a separatory funnel and then dried by rotary evaporation. The resulting residue was passed through a silica column and eluted using a hexane: ethyl acetate gradient. Fractions of interest were further purified by HPLC using an Agilent 1260 series instrument equipped with an autosampler, fraction collector and diode array UV detector over a ZORBAX Eclipse XDB C18 (4.6 x 150 mm, 5 mm) at a 0.5 mL/min flow rate using a water and acetonitrile gradient as the mobile phase. Purified compounds were resuspended in CDCl₃. Structural analysis was performed using 1D (¹H, ¹³C) and 2D (TOCSY, ROESY, HMB, DQF, and 13CHSQC) NMR experiments. Spectra were acquired on a Bruker Avance 500-MHz TopSpin NMR spectrometer.

Plant material, growth conditions and elicitors treatment method

Germination of *M. truncatula* cv. Jemalong A17 seeds were carried out by scarification using concentrated anhydrous sulfuric acid for 5 min followed by intensive washing steps with Milli-Q water. Seeds were then placed in water in a 50 ml falcon tube at 4°C for 3 days to break dormancy. Seeds were transferred to moist blotting paper and placed on a petri dish incubated at 26°C. Single seedlings were transferred to pots (9 cm diameter) filled with a perlite and sand mixture (5:1), then grown in a growth chamber with 16 h light/8 h dark for 6 weeks. After 6 weeks of growth, plants were treated by foliar application with solutions containing 1 mM methyl jasmonate (MeJA) in 0.01% Triton-X-100 (Sigma-Aldrich, St Louis, MO, USA) or 1.5 mM salicylic acid solution dissolved in 1 % ethanol (SA) (Sigma-Aldrich, St Louis, MO, USA). Plants were treated with 0.01% Triton-X-100 for MeJA control and 1% ethanol for SA. For each plant, 20 ml solution was applied (either MeJA or SA or controls) and at least three biological replicates were conducted for each treatment. The leaf and root tissues were separately collected in time courses from 0h, 2 h, 6 h, 12 h, 24, 48, 72 and 96h and immediately frozen in liquid nitrogen and stored at -80 °C for further RNA extraction.

RNA extraction and cDNA synthesis from root and leaf tissues

Sampling was performed on three biological replicates for both elicitors and controls. The root and leaf tissues were separately ground in liquid nitrogen and 100 mg powder was used for total RNA isolation with an RNeasy Plant Mini Kit (Qiagen), and RNase-Free DNase to remove genomic DNA following the manufacturer's protocol. RNA purity and yield were checked using a NanoDrop 8000 Spectrophotometer (ThermoFisher Scientific), while RNA integrity and quantification were evaluated with an Agilent 2100 bioanalyzer (Agilent Technologies). First-strand cDNA was synthesized using a Superscript IV Reverse transcriptase kit (ThermoFisher Scientific, USA).

Quantitative real time-PCR (qPCR)

Primers were designed using coding sequences from the genome of *M. truncatula* (Mt4.0v1) in Phytozome13 (<https://phytozome-next.jgi.doe.gov/>) (Table S2). A quantitative real-time PCR system was carried out using SYBR green with ROX (pre-mixed) as an internal loading standard performed on an Eppendorf Mastercycler. The reaction mixture was 10 µl and comprised of 5 µl of 2× SYBR™ Green qPCR Master Mix (ThermoFisher Scientific, USA), 100 nM primers (Thermo Fisher Scientific) and 0.5 µl of (1:10 dilution) cDNA. PCR protocol was followed to initiate polymerase activation: 10 min at 95 °C; 40 cycles of 30 s at 95 °C, 60 s at 55 °C and 30 s at 72 °C. Each run of qPCR was followed by a melting curve analysis from 55 to 95 °C. PCR conditions were determined by comparing the threshold values of the RT product (cycle threshold (Ct) value). Relative RNA levels were calibrated and normalized with the level of actin and histone-3 reference genes based on the $2^{-\Delta\Delta C_t}$ method [28, 29]. For each gene, three technical replicates were run for three biological samples. qPCR data were expressed as a fold change with respect to the equivalent time-point in the control groups of MeJA (in 0.01 % Triton-X-100 and SA (in 1 % ethanol).

RESULTS

Genome-wide analysis of *MtTPSs*

We assessed for the genetic potential of *M. truncatula* cv. Jemalong A17 to produce terpenes using the Mt4.0v1 genome in Phytozome. We identified 54 putative *MtTPS* gene models using BLAST and subsequent analysis of characteristic motifs (Table S3 and Figure S1). Further curation through sequence alignments confirmed thirty- nine full-length genes. Twenty-two of the reported fifty-four potential TPSs in *M. truncatula* were previously assigned numbers by Parker et al. 2014 (*MtTPS16* is no longer classified as a putative terpene synthase) [30]. Table S3 is a summary of the previous and newly assigned TPS numbers. A majority of the genes had the Class I DDxxD and secondary NSE/DTE metal binding motif involved in ionization-dependent cyclization [1]. Three Class II synthases were identified by DxDD motifs involved in protonation-initiated cyclization characteristic of copalyl diphosphate synthases (CPPs) involved in diterpene scaffold formation. Phylogenetic analysis of the terpene synthases from *M. truncatula* and the comprehensively biochemically characterized TPS collections of *A. thaliana* and *S. lycopersicum* shows plant species clustering more than enzymes with similar functions (Figure 2) [3, 25, 30, 31]. The genes are distributed into the different clades with fourteen members in TPS-a; five members in TPS-b; three members in TPS-c and four members in TPS e/f. Interestingly there is an expanded TPS-g (thirteen members) subfamily as compared to *A. thaliana* and *S. lycopersicum*.

The biochemical analysis described here focuses on the TPS-a subfamily only, guided by previous studies that showed sesquiterpene compounds are a predominant component of volatile collections of *M. truncatula* in response to herbivory [5, 6]. In addition, subsequent biochemical characterization assigned metabolites for three of fourteen sesquiterpene synthases - namely *MtTPS1*, 5, and 10 - which are in the TPS-a subfamily (Figure 1) [6, 7, 9]. *MtTPS2* showed increased gene expression upon elicitation and was putatively described as germacrene-D synthase based on phylogenetic analysis but there was no corresponding biochemical confirmation [9]. Figure 2 shows little overlap of the members of TPS-a clade amongst plants as they tend to group by species. In *S. lycopersicum* the fifteen TPSs encode for sesquiterpenes whereas *Arabidopsis* TPS-a members are not limited to sesquiterpene biosynthetic genes [25]. Combined, these

observations guided our investigations to determine the functional landscape of TPS-a subfamily members in *M. truncatula* through heterologous expression.

Biochemical analysis of TPS-a subfamily.

To correlate metabolites produced to the identified biosynthetic genes, we assessed for TPS activity and resultant metabolite formation using a modular metabolic engineering system in *E. coli* [27]. Briefly, heterologous expression was performed through co-expression of an (*E, E*)-FPP synthase with uncharacterized full-length TPS-a subfamily genes (*MtTPS2*, 7, 8, 9, 11, 12, 13, 24, 25, 26 and 27). Metabolite production was analyzed using GC-MS (Figure 3). Compounds were compared to authentic standards when available, the NIST library database, or verified through de novo structural elucidation by NMR.

Analysis of resultant metabolites produced by TPS-a subfamily members reveals the production of an assortment of sesquiterpenes. Figure 3 highlights the most abundant products formed by TPSs and products in smaller quantities are described in the supplementary information (Figures S3-S8). *MtTPS7* is a multiproduct enzyme that produces α -copaene (**2**), cyclosativene (**3**) and a hydroxylated compound which matches a compound produced by *MtTPS26* (Figure 3B & S3). *MtTPS8* is a multiproduct synthase that produces farnesene (**4**) as the major compound and exhibits relatively significant turnover to amorphadiene, bergamotene and nerolidol, and production of up to fourteen products in smaller amounts (Figure 3C & S4), *MtTPS11*, produces longicyclene (**5**) as the main product and a smaller amount of longifolene and longipinene (Figure 3D & S5).

Although terpene synthases more commonly form hydrocarbons, the addition of water before terminating deprotonation generates hydroxylated products [1]. *M. truncatula* produces a set of TPSs capable of producing such hydroxylated compounds as supported by a molecular weight of 222. *MtTPS13* produces germacrene-D-4-ol (**6**) as compared to NIST (Figure 3E & S6). The genes, *MtTPS24* and *MtTPS26* are present in the *Mt4.0v1* genome and not identified in earlier genomes [30]. *MtTPS24* produces a compound with a distinct base peak at 151 (**7**) (Figure 3F & S7). However, we were unable to collect NMR data as the compound aggregates in tested deuterated solvents (chloroform, methanol and DMSO) and no further analysis could be performed. Upon expression, *MtTPS26* produces two predominant compounds that required structural verification by NMR due to the unavailability of authentic standards. Structural elucidation by NMR and comparison to literature values showed that compound **8** was eudesm4(14)-en-6 β -ol (Figure 3G) with chemical shift data closely matching previously reported data for 10-*epi*-junenol except for the hydroxyl in a 6 β -position in **8** (Table S5) [32]. Compound **9** was confirmed to be guai-6-en-10 β -ol, matching previously reported chemical shifts as shown in Table S6 [33]. No activity was detected with *MtTPS2*, 9, 12, 25 and 27 upon co-expression with (*E, E*)-FPP synthase using the combinatorial expression system.

The ability of TPSs to make several products is not uncommon, for example, *MtTPS5* was demonstrated to produce up to twenty-seven products [7, 34]. Here, *MtTPS8* is the most promiscuous forming up to twenty compounds (Figure S4). Multiple products produced in smaller quantities are highlighted for all the genes (Figures S3-S8). Similarly, because of the related mechanism of terpene synthases, it is common to observe that some genes will also produce the same compounds in varying amounts *in vitro*. This is observed for *MtTPS7* and *MtTPS26* which

both produce **9**; a similar case is noted for *MtTPS11* which produces longicyclene (**5**) as the main product, but profiles of *MtTPS5* & *10* also produce longicyclene in small amounts [7, 13].

In summary, this biochemical analysis defines gene-to-metabolite relationships in the TPS-a subfamily of *M. truncatula*. We defined the genetic basis for the biosynthesis of farnesene, nerolidol, cyclosativene, copaene and longicyclene, enabling the association of genes responsible for the biosynthesis of several compounds observed in terpene metabolite response to herbivore feeding [5] and identified a set of sesquiterpene alcohols (Figure 4).

Gene expression analysis of TPS-a members in response to treatment with methyl jasmonate and salicylic acid.

Gene expression of the TPS-a subfamily members in response to elicitors was carried out through exogenous application of phytohormones MeJA and SA and analysis by qPCR (Figure 5). Overall, TPS-a genes showed tissue-specific and temporal differences in expression in response to elicitation with MeJA and SA. There was higher expression of TPS-a genes in leaf tissue compared to the root upon elicitation with MeJA. Specifically, our time course evaluations revealed that MeJA application resulted in increased expression (16-20 fold) of *MtTPS1* after 6h of exposure in leaves whereas expression of *MtTPS2* and *MtTPS5* was highest after 12 hours. A 4-8 fold increase in expression was observed amongst the *MtTPS8*, *MtTPS10*, *MtTPS13* and *MtTPS26* genes in leaf tissues after 6 and 12 h of post-treatment of MeJA. *MtTPS11*, 12, 24, 25 & 27 have low expression in both the leaves and roots. There was a consistently low expression for all the TPS-a members in roots under MeJA elicitation, with *MtTPS25* & 27 showing down-regulation.

In contrast, SA induced gene expression occurred mainly in roots. The genes *MtTPS10* and *MtTPS11* showed the highest expression at 24h compared to control with a 20-fold increase whereas a 16-fold increase in expression was observed for genes *MtTPS13*, 24 & 26 (Figure 5). However, comparatively lower expression was observed in both leaf and root tissue for the remaining eight genes (*MtTPS1*, 2, 5, 7, 8, 12, 25 & 27). Overall, the TPS-a genes exhibited differential expression to the phytohormones where MeJA largely affected the expression of TPS genes in leaf tissue and SA in root tissue upon foliar treatment.

Expression patterns of TPS in different tissues and under (a)biotic stress.

We used a publicly available expression data set of *M. truncatula* from the Noble Research Institute legacy Gene Atlas V3 to survey the effect of (a)biotic stress on the TPS-a subfamily [35, 36]. Mining of the reference data set, which is a collation of gene expression in different tissues, shows that terpene synthases *MtTPS1*, 2 & 5 which produce volatile compounds exhibited increased expression in the shoots and leaves. Transcript levels of the *MtTPS13* were increased during root and nodule development, AM symbiosis, and fungal colonization (Figure S23A).

Evaluation of the TPS-a subfamily genes response to abiotic stress highlights increased transcript abundance of *MtTPS1*, 2, 5 and *MtTPS13* under limited nitrogen (N₂), ammonium (NH₄), nitrate (NO₃), drought, and salt stress in shoot and root tissues (Figure S23B). An assessment of varied biotic stressors showed that *MtTPS1* and 2 had increased expression in leaf tissue after aphid infestation whereas *MtTPS13*, *MtTPS24* and *MtTPS25* displayed increased transcript abundance in root tissues under pathogenic fungi, *Rhizoctonia solani* and *Fusarium oxysporum* (Figure S23C). Notably, datasets obtained from Gene Atlas do not capture the expression of six

terpene synthases (*Mt*TPS7,8, 9, 25, 26 & 27). Combined this data supports the variable expression of select terpene synthases dependent on tissue, time and stressor applied.

DISCUSSION

Biochemical analysis reveals functional divergence of TPS-a subfamily.

The catalytic mechanism of sesquiterpenes proceeds via a succession of carbocationic intermediates initiated by dephosphorylation and subsequent deprotonation or water capture [1]. The metabolites of the encoded enzyme of *Mt*TPS1, *Mt*TPS7 and *Mt*TPS8 have characteristic profiles to previously described enzymes, for example, *Mt*TPS8 produces (*E*)- β -farnesene with bergamotene similar to what is observed for *Zm*TPS10 [10, 20, 37]. Notably, *Mt*TPS11 produces longicyclene, a compound commonly found in pine trees [38]. Previous studies have identified a longifolene synthase with smaller amounts of longicyclene and longipinene - here the enzyme *Mt*TPS11 forms longicyclene as the major product. Olefins (compounds 2-5) produced by *M. truncatula* are characteristic in volatile blends of several plants upon herbivore feeding and associated interactions [20, 37]. Here, some of the TPS-a genes quenched the reaction through water capture resulting in sesquiterpene alcohols (compounds 6-9) that have been identified in other plants, but little is known of their biological activity. Our biochemical analysis combined with the previously characterized genes establishes the use of terpenes common to plants generating a species-specific sesquiterpene blend.

A summary of the known gene-to-metabolite relationships in the TPS-a subfamily of *M. truncatula* shows functional divergence from the TPS profiles of the two comprehensively biochemically characterized eudicots (Figure 4). Previously characterized genes from *M. truncatula* mainly fall in the TPS-a clade (Figures 1 & 2). Comparison of the currently defined *M. truncatula* TPS-a subfamily members compared to those of *Arabidopsis* and tomato show a difference in metabolite profiles. In tomato, the fifteen TPSs encode for sesquiterpenes whereas the product outcome of the TPS-a subfamily in *A. thaliana* is not limited to sesquiterpenes (Supplementary Table S4) [25]. Biochemical analysis here revealed that a subset of the metabolites synthesized by *M. truncatula* sesquiterpene synthases was specific to the species, revealing functional divergence.

Differential expression of TPS-a subfamily in response to phytohormone elicitation

In this study, we elicited TPS expression by exogenous application of phytohormones. The genes that produce volatile compounds had increased expression on the early onset of foliar application of MeJA (Figures 1 & 3). Genes corresponding to the production of volatile compounds β -caryophyllene (*Mt*TPS1), putative germacrene-D (*Mt*TPS2) and *Mt*TPS5 a multiproduct synthase with the demonstrated antiherbivore response were upregulated by 16 to 20-fold (Figure 5). These observations correlate in part to a study by Leitner et al. (2005), which showed that herbivore interactions and artificial wounding induced the emission of up to 23 compounds [5, 6]. This observation is supported by data from gene atlas that also showed increased expression of the volatile terpenes predominantly in the leaves under various stresses (Figure S23). However, ten out of fourteen TPS-a subfamily members exhibited moderate upregulation by 4-6-fold using MeJA.

In contrast, elicitation using SA increased the transcript abundance in root tissue of select genes; *Mt*TPS10, *Mt*TPS11, *Mt*TPS13, *Mt*TPS24 and *Mt*TPS26 (Figure 5). In general, SA mediates response to biotrophic and hemibiotrophic pathogens and triggers systemic acquired resistance

[39]. Yadav et al (2019) demonstrated that MtTPS10, which produces himalachol, reduced susceptibility to the oomycete pathogen *Aphanomyces euteiches* [13]. The terpene scaffolds may be elaborated by downstream enzymes resulting in further oxygenated non-volatile sesquiterpenoids. The resultant compounds may have various bioactivity similar to previously identified antifungals and insect deterrents such as germacrene-A based costunolides in lettuce; costic acids derived from selinene and β -macrocarpene based zealexins produced in maize under pathogen attack [40-42].

In *conclusion*, the results reported here establish gene-metabolite relationships of the TPS-a subfamily that encodes for sesquiterpene synthases in *M. truncatula*. Biochemical analysis reveals a blend of sesquiterpenes that highlight the evolution of species-specific chemotypes. Gene expression analysis shows a temporal tissue-specific and stress induced variation. While determining the function of these compounds is beyond the scope of this work, our study further defines the sesquiterpene biosynthetic capacity and provides a foundation to understand the biological roles of the compounds.

Conflict of Interest

There is no conflict of interest.

Author Contributions

HH performed biochemical analysis. MI performed gene expression analysis. GFW performed NMR analysis. SM designed the research, supervised the project and wrote the manuscript with contributions from all authors. All authors reviewed the article and approved the final version before submission.

Funding

Financial support was provided by the Department of Biochemistry and Molecular Biology at the University of Massachusetts – Amherst; Richard and Susan Smith Family Foundation and National Science Award 2300060 to SM. HH was supported by the National Research Service Award T32 GM135096 from the National Institutes of Health.

Acknowledgments

Mass spectral data were obtained at the University of Massachusetts Mass Spectrometry Core Facility, RRID: SCR_019063. NMR data were obtained at Rensselaer Polytechnic Institute (RPI) NMR Core Facilities: NIST Award # 60NANB22D167.

FIGURE LEGENDS

Figure 1. A) Schematic of terpene biosynthesis. Three major substrates are used in the production of most terpene products by TPSs: GPP (C10), (*E, E*)-FPP (C15), and (*E, E, E*)-GGPP (C20). B) Resultant terpene products of previously characterized *Mt*TPSs. *Mt*TPS1 produces β -caryophyllene, *Mt*TPS3 produces both the sesquiterpene product nerolidol and the diterpene product geranyl linalool, *Mt*TPS4 produces β -ocimene, *Mt*TPS5 produces cubebol as a major product but is a multiproduct enzyme and *Mt*TPS10 produces mainly himachalol.

Figure 2. Phylogeny-based analysis of terpene synthases in *Medicago truncatula*, *Arabidopsis thaliana*, and *Solanum lycopersicum*. The sequences were aligned using maximum likelihood analysis with 1000 bootstrap repetitions using CLC software and visualized using iTOL. TPS subfamilies are grouped as follows: TPS-a (green), monoterpenes in TPS-b (blue) and TPS-g (red), and diterpenes in TPS-c (orange) and TPS-e/f (purple).

Figure 3. Functional of the TPS-a subfamily in *Medicago truncatula*. Total ion chromatograms of products resulting from combinatorial expression in *E. coli* of (*E, E*)-FPP synthase and corresponding genes from the TPS-a subfamily. The chromatograms focus on the main products. Detailed analysis of the smaller peaks is depicted in Supplementary Figures S3-8. Compound identification was through similarity to library matches (NIST, Wiley) for B-E and NMR verification for G.

- A) pET28b containing FPP synthase produces H) dephosphorylated FPP (farnesol) (1)
- B) *Mt*TPS7 produces (I & J) cyclosativene (2), α -copaene (3).
- C) *Mt*TPS8 produces K) farnesene-like product (4).
- D) *Mt*TPS11 produces L) longicyclene (5).
- E) *Mt*TPS13 makes a M) germacrene-D-4-ol product (6),
- F) *Mt*TPS24 makes a N) unknown product (7) and
- G) *Mt*TPS26 makes a O & P) eudesm4(14)-en-6 β -ol (8) and a guai-6-en-10 β -ol (9)

Figure 4: Summary of the main products from biochemical characterization of TPS-a subfamily members in *M. truncatula*.

Figure 5: Expression profile of TPS-a terpene synthase genes in leaf and root tissues in response to exogenous application of MeJA and SA in a time course exposure by qRT-PCR. Plants were exposed for 0h, 2 h, 6 h, 12 h, 24 h, 48, 72 and 96 h after foliar application of MeJA (1 mM) or SA (1.5 mM). The data are expressed as the relative level of expression of fold change relative to controls (0.01% Triton-100 for MeJA, or 1% ethanol for SA control) plants, based on the $2^{-\Delta\Delta ct}$ method. Error bars represent the mean $\pm$ standard deviation (SD) levels of the relative abundance of three biological replicates.

REFERENCES

1. Christianson, D.W., *Structural and Chemical Biology of Terpenoid Cyclases*. Chem Rev, 2017. **117**(17): p. 11570-11648.
2. Pichersky, E. and R.A. Raguso, *Why do plants produce so many terpenoid compounds?* New Phytol, 2018. **220**(3): p. 692-702.
3. Chen, F., et al., *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., 2011. **66**(1): p. 212-29.
4. He, J., et al., *Terpene synthases in cucumber (*Cucumis sativus*) and their contribution to herbivore-induced volatile terpenoid emission*. New Phytol, 2022. **233**(2): p. 862-877.
5. Leitner, M., W. Boland, and A. Mithöfer, *Direct and indirect defences induced by piercing-sucking and chewing herbivores in *Medicago truncatula**. New Phytol, 2005. **167**(2): p. 597-606.
6. Arimura, G., et al., *Herbivore-induced terpenoid emission in *Medicago truncatula*: concerted action of jasmonate, ethylene and calcium signaling*. Planta, 2008. **227**(2): p. 453-64.
7. Garms, S., T.G. Köllner, and W. Boland, *A multiproduct terpene synthase from *Medicago truncatula* generates cadalane sesquiterpenes via two different mechanisms*. J Org Chem, 2010. **75**(16): p. 5590-600.
8. Navia-Giné, W.G., et al., **Medicago truncatula* (E)-beta-ocimene synthase is induced by insect herbivory with corresponding increases in emission of volatile ocimene*. Plant Physiol Biochem, 2009. **47**(5): p. 416-25.
9. Gomez, S.K., et al., *Lepidopteran herbivory and oral factors induce transcripts encoding novel terpene synthases in *Medicago truncatula**. Arch Insect Biochem Physiol, 2005. **58**(2): p. 114-27.
10. Rasmann, S., et al., *Recruitment of entomopathogenic nematodes by insect-damaged maize roots*. Nature, 2005. **434**(7034): p. 732-7.
11. Degenhardt, J., et al., *Restoring a maize root signal that attracts insect-killing nematodes to control a major pest*. Proc Natl Acad Sci U S A, 2009. **106**(32): p. 13213-8.
12. Kang, Z.W., et al., *Volatile β -Ocimene Can Regulate Developmental Performance of Peach Aphid*. Front Plant Sci, 2018. **9**: p. 708.
13. Yadav, H., et al., **Medicago* TERPENE SYNTHASE 10 Is Involved in Defense Against an Oomycete Root Pathogen*. Plant Physiol, 2019. **180**(3): p. 1598-1613.
14. Han, T., et al., *Functional Characterization of a* (Int J Mol Sci, 2023. **24**(8).
15. Liu, J., et al., *Genome-wide analysis of terpene synthases in soybean: functional characterization of GmTPS3*. Gene, 2014. **544**(1): p. 83-92.
16. Lin, J., et al., *An (E,E)- α -farnesene synthase gene of soybean has a role in defence against nematodes and is involved in synthesizing insect-induced volatiles*. Plant Biotechnol J, 2017. **15**(4): p. 510-519.
17. Constantino, N., et al., *Soybean Cyst Nematodes Influence Aboveground Plant Volatile Signals Prior to Symptom Development*. Front Plant Sci, 2021. **12**: p. 749014.
18. Steinbrenner, A.D., et al., *Signatures of plant defense response specificity mediated by herbivore-associated molecular patterns in legumes*. Plant J, 2022. **110**(5): p. 1255-1270.
19. Bao, T., et al., *Functional Characterization of Terpene Synthases Accounting for the Volatilized-Terpene Heterogeneity in *Lathyrus odoratus* Cultivar Flowers*. Plant Cell Physiol, 2020. **61**(10): p. 1733-1749.
20. Schnee, C., et al., *The products of a single maize sesquiterpene synthase form a volatile defense signal that attracts natural enemies of maize herbivores*. Proc Natl Acad Sci U S A, 2006. **103**(4): p. 1129-34.

21. Schmelz, E.A., et al., *Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins*. Plant J, 2014. **79**(4): p. 659-78.
22. Ding, Y., et al., *Genetic elucidation of interconnected antibiotic pathways mediating maize innate immunity*. Nat Plants, 2020. **6**(11): p. 1375-1388.
23. Zhang, J., et al., *A (conditional) role for labdane-related diterpenoid natural products in rice stomatal closure*. New Phytol, 2021. **230**(2): p. 698-709.
24. Murphy, K.M. and P. Zerbe, *Specialized diterpenoid metabolism in monocot crops: Biosynthesis and chemical diversity*. Phytochemistry, 2020. **172**: p. 112289.
25. Zhou, F. and E. Pichersky, *The complete functional characterisation of the terpene synthase family in tomato*. New Phytol, 2020. **226**(5): p. 1341-1360.
26. Letunic, I. and P. Bork, *Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation*. Nucleic Acids Res, 2021. **49**(W1): p. W293-W296.
27. Kitaoka, N., et al., *Optimization of recombinant expression enables discovery of novel cytochrome P450 activity in rice diterpenoid biosynthesis*. Appl Microbiol Biotechnol, 2015. **99**(18): p. 7549-58.
28. Rey, T., et al., *A Central Transcriptional Regulator of Symbiotic Nodule Development, Is Also a Determinant of*. Front Plant Sci, 2016. **7**: p. 1837.
29. Zhang, H., et al., *Efficient Generation of CRISPR/Cas9-Mediated Homozygous/Biallelic*. Front Plant Sci, 2020. **11**: p. 294.
30. Parker, M.T., et al., *Comparative genomic and transcriptomic analysis of terpene synthases in Arabidopsis and Medicago*. IET Syst Biol, 2014. **8**(4): p. 146-53.
31. Falara, V., et al., *The tomato terpene synthase gene family*. Plant Physiol, 2011. **157**(2): p. 770-89.
32. F., T.A., et al., *10-epijunenol, a new cis-eudesmane sesquiterpenoid*. Tetrahedron, 1976. **32**: p. 2261-2264.
33. Lago, J.H., C.B. Brochini, and N.F. Roque, *Terpenes from leaves of Guarea macrophylla (Meliaceae)*. Phytochemistry, 2000. **55**(7): p. 727-31.
34. Vattekkatte, A., S. Garms, and W. Boland, *Alternate Cyclization Cascade Initiated by Substrate Isomer in Multiproduct Terpene Synthase from Medicago truncatula*. J Org Chem, 2017. **82**(6): p. 2855-2861.
35. Benedito, V.A., et al., *A gene expression atlas of the model legume Medicago truncatula*. Plant J, 2008. **55**(3): p. 504-13.
36. Carrere, S., J. Verdier, and P. Gamas, *MtExpress, a Comprehensive and Curated RNAseq-based Gene Expression Atlas for the Model Legume Medicago truncatula*. Plant Cell Physiol, 2021. **62**(9): p. 1494-1500.
37. Schnee, C., et al., *The maize gene terpene synthase 1 encodes a sesquiterpene synthase catalyzing the formation of (E)-beta-farnesene, (E)-nerolidol, and (E,E)-farnesol after herbivore damage*. Plant Physiol, 2002. **130**(4): p. 2049-60.
38. Martin, D.M., J. Fäldt, and J. Bohlmann, *Functional characterization of nine Norway Spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily*. Plant Physiol, 2004. **135**(4): p. 1908-27.
39. Anand, A., et al., *Salicylic acid and systemic acquired resistance play a role in attenuating crown gall disease caused by Agrobacterium tumefaciens*. Plant Physiol, 2008. **146**(2): p. 703-15.
40. Nguyen, D.T., et al., *Biochemical conservation and evolution of germacrene A oxidase in asteraceae*. J Biol Chem, 2010. **285**(22): p. 16588-98.
41. Ding, Y., et al., *Multiple genes recruited from hormone pathways partition maize diterpenoid defences*. Nat Plants, 2019. **5**(10): p. 1043-1056.
42. Ding, Y., et al., *Selinene Volatiles Are Essential Precursors for Maize Defense Promoting Fungal Pathogen Resistance*. Plant Physiol, 2017. **175**(3): p. 1455-1468.

In review

Figure 1.TIFF

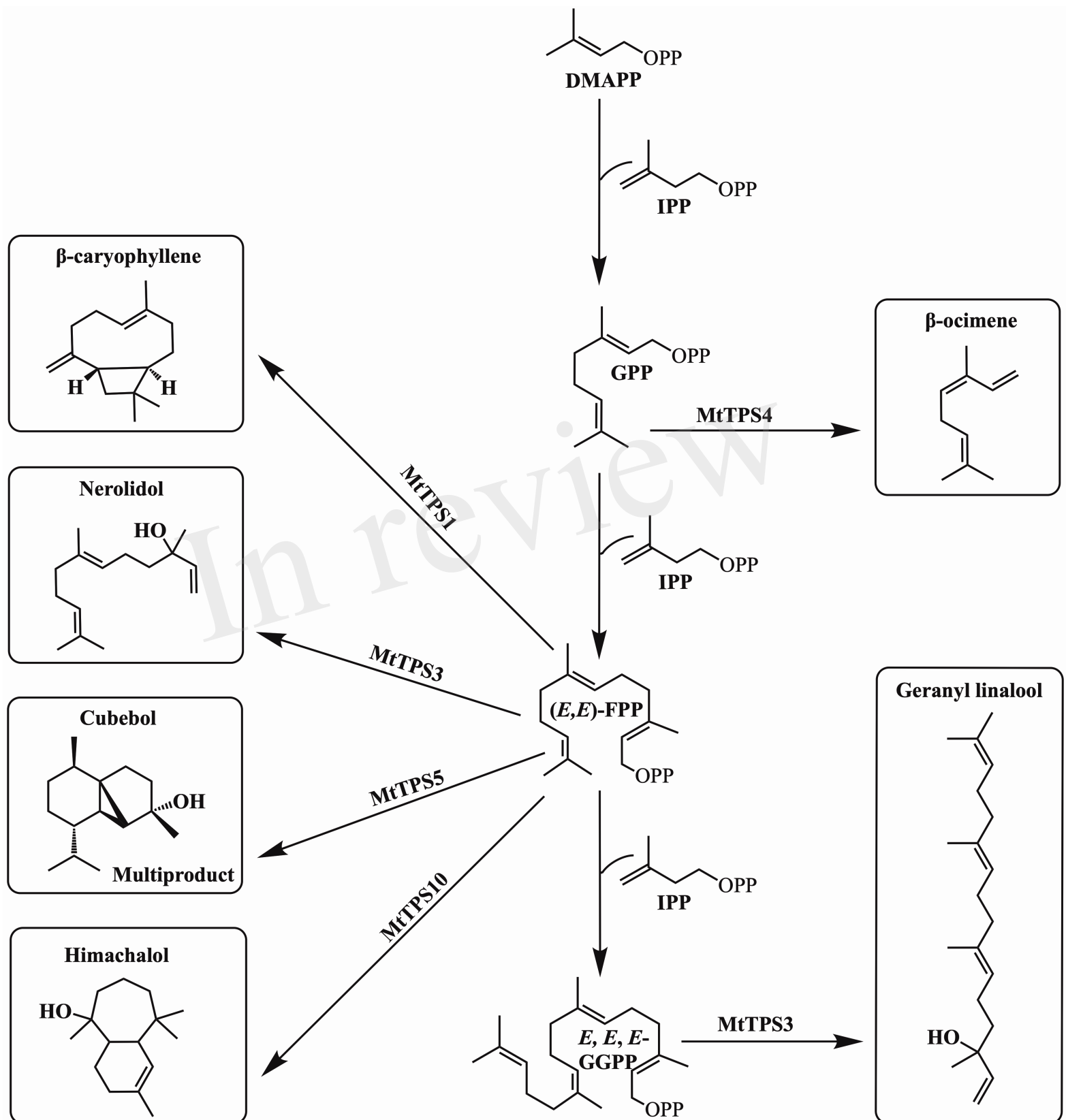
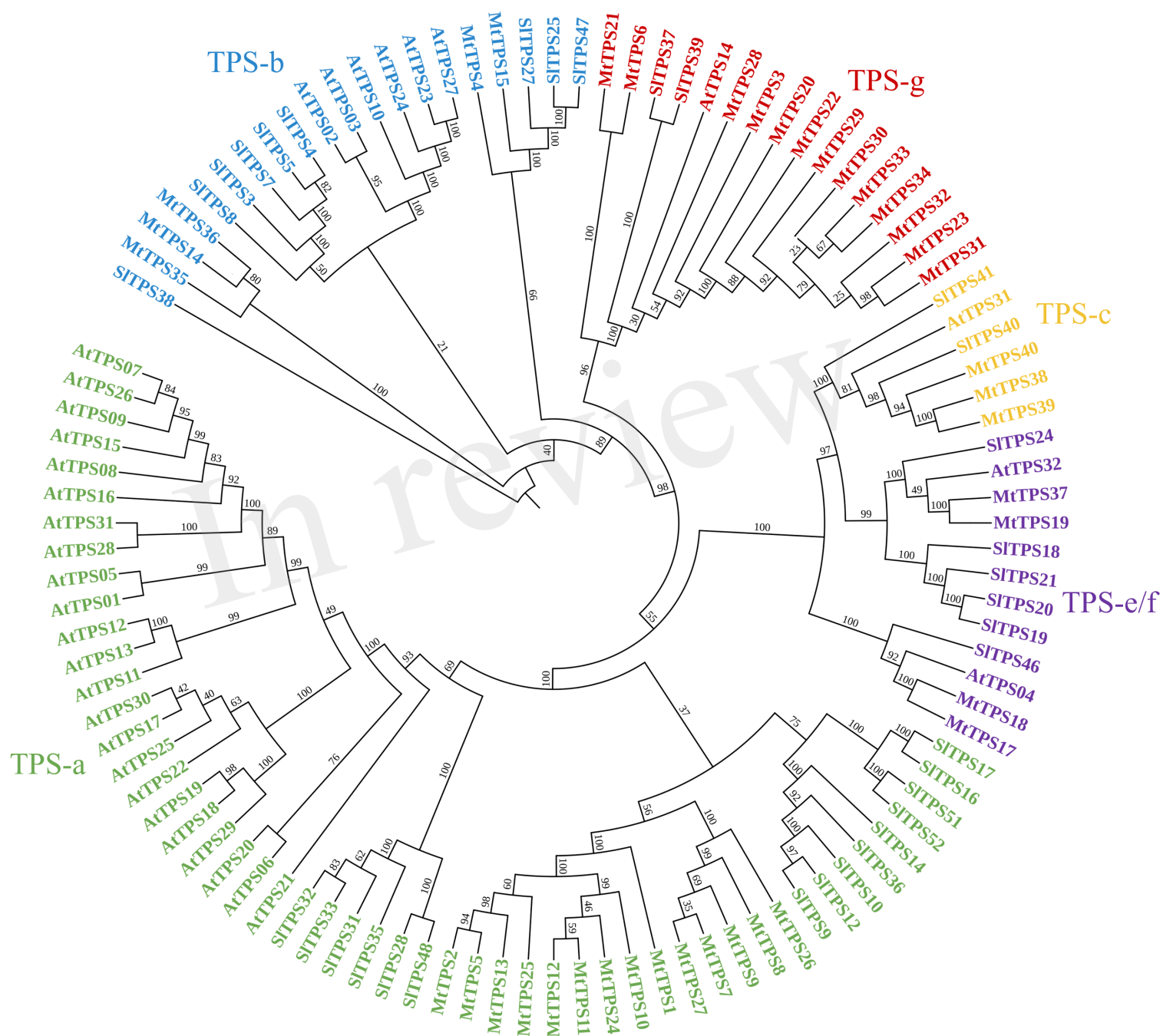


Figure 2.TIF



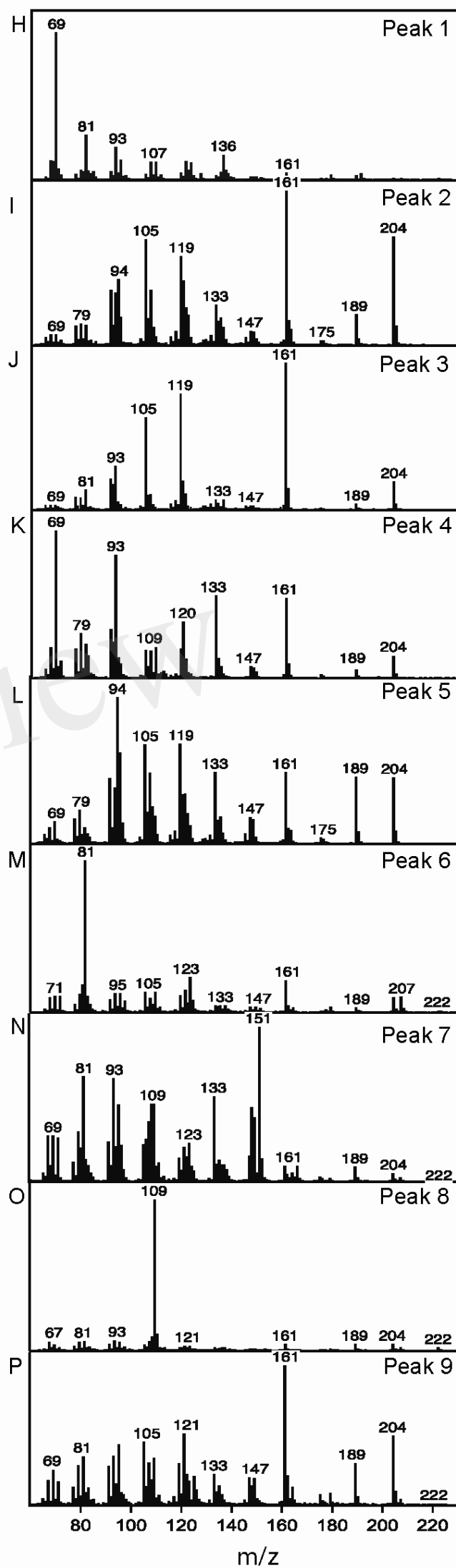
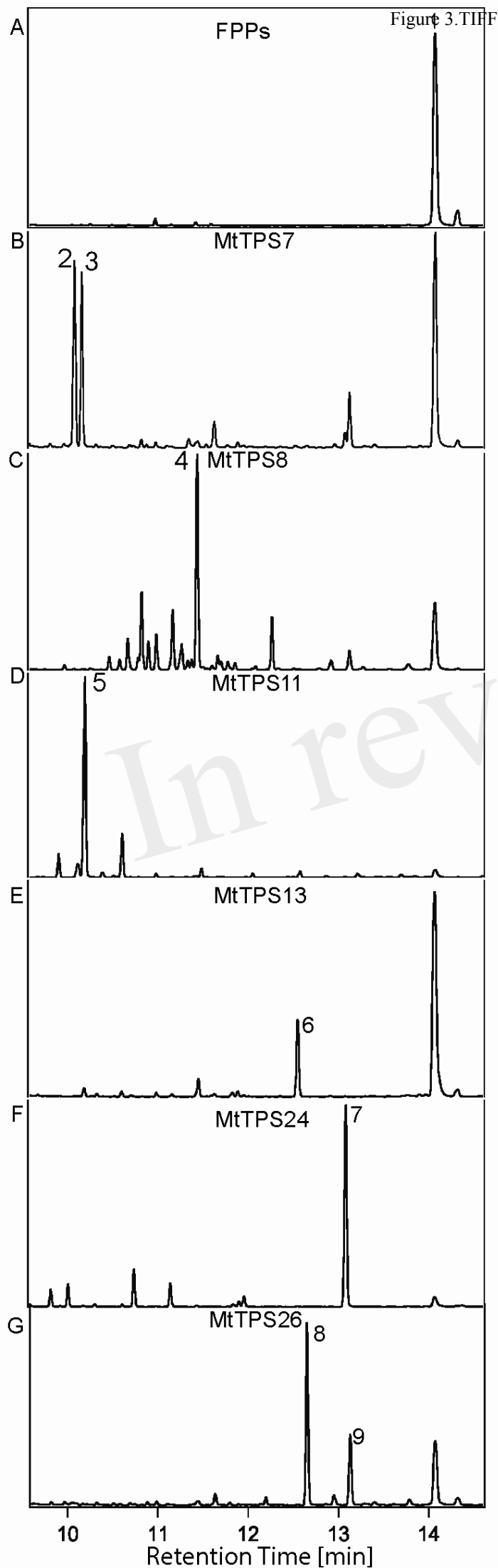


Figure 4.TIFF

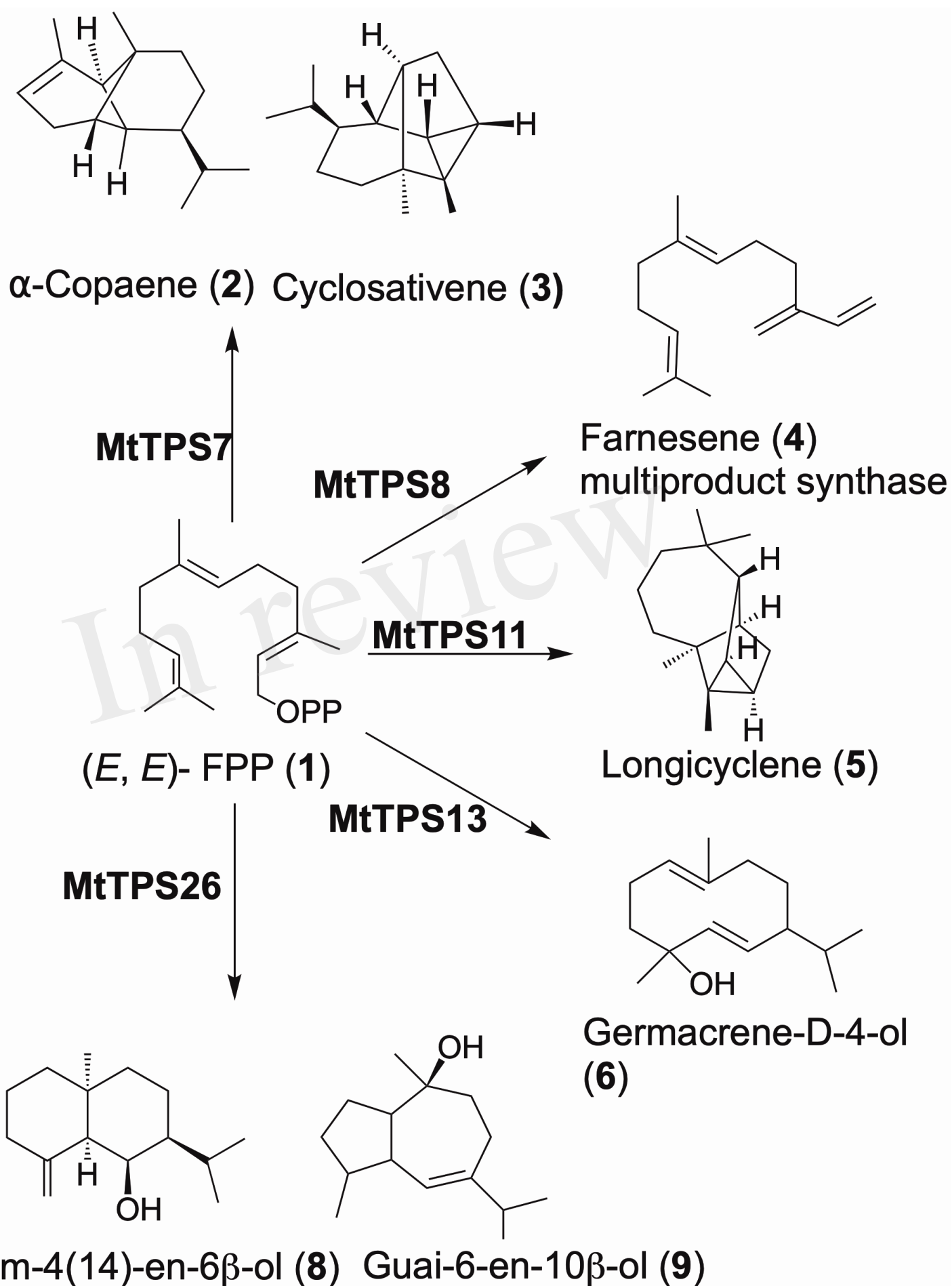


Figure 5.TIF

