Characterization of microbial type terpene synthases from two liverworts Lophocolea bidentata and Nardia scalaris

Conart, Corentin; Bejder, Benjamin Svejdal; Schardelmann, Mikkel; Batth, Rituraj; Kodiripaka, Satish Kumar; Petersen, Bent; Simonsen, Henrik Toft

Published in:
Plant Physiology and Biochemistry

DOI:
10.1016/j.plaphy.2025.110031

Publication date:
2025

Document version
Publisher's PDF, also known as Version of record

Plant Physiology and Biochemistry 226 (2025) 110031

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Characterization of microbial type terpene synthases from two liverworts *Lophocolea bidentata* and *Nardia scalaris*

Corentin Conart [a], Benjamin Svejdal Bejder [b], Mikkel Schardelmann [b], Rituraj Batth [b], Satish Kumar Kodiripaka [b], Bent Petersen [c], Henrik Toft Simonsen [a,b,*]

[a] Université Jean Monnet Saint-Étienne, CNRS, LBVpam UMR 5079, 42023, Saint-Étienne, France
[b] Department of Biotechnology and Biomedicine, The Technical University of Denmark, 2800, Kongens Lyngby, Denmark
[c] Globe Institute, University of Copenhagen, Øster Voldgade 5-7, 1350, Copenhagen K, Denmark

A B S T R A C T

Microbial terpene synthase-like (MTPSL) enzymes from the two liverworts *Nardia scalaris* (Gymnomitriaceae) and *Lopcholea bidentata* (Lophocoleaceae) were characterized in yeast. The hexane extract of the fresh plant material was also analyzed for its terpenoid content. Following the expression of the enzymes in yeast (WAT11), the volatile terpenoid products were analyzed using GC-MS, and their scent was qualitatively evaluated. Four full lengths MTPSLs from *L. bidentata* and four from *N. scalaris* were identified in the transcriptome data and expressed in yeast. Three MTPSLs from *L. bidentata* and two from *N. scalaris* were found to be biochemically active in yeast, and all of these were found to be sesquiterpene synthases without any monoterpenoid products. These represent the first terpene synthases characterized from these two plants. *Lb*MTPSL1 was found to be a murolene synthase, *Lb*MTPSL3 was a cubenene synthase, and *Lb*MTPSL5 had α-selinene and β-elemene as the major products. For *N. scalaris* it was found that *Ns*MTPSL3 was a viridiflorol synthase and *Ns*MTPSL4 was a α/β-chamigrene synthase with at least five other minor products. The 3D structure of the enzymes was modelled using CPH-models, in order to enhance the understanding of the biochemical differences of the enzymes. The models clearly show that these MTPSL only contain one α-domain, which is characteristic of MTPSLs and is previously described for similar enzymes. The bioactivities of the enzymes align with compounds identified in the two liverworts, and this study represents the first onset of further investigations into terpenoid biochemistry of *Lopcholea* and *Nardia* species.

1. Introduction

Liverworts are the second largest group of bryophytes, and are characterized by a significant presence of lipid droplets in the gametophytes. Likewise, liverworts are known to produce a wide range of natural products, including terpenoids, which have various biological activities. The study of terpenoids in liverworts has become increasingly important due to the discovery of new bioactive compounds with potential pharmaceutical and industrial applications (Asakawa et al., 2013; Horn et al., 2021; Ludwiczuk and Asakawa, 2019; Marques et al., 2023; Sabovljević et al., 2016).

The biosynthesis of terpenes in liverworts and bryophytes involve classic plant terpene synthases like the diterpene synthase *ent*-kaurene synthase (Hayashi et al., 2006; Zhan et al., 2015) that have both βγ- and α-domains, and mono- and sesquiterpene synthases that only have β- and α-domains (Karunanithi and Zerbe, 2019). Additionally, bryophytes and other non-seed plants also contain a second class of terpene synthases that only harbor the α-domain, namely the microbial terpene synthase-like enzymes (MTPSL). These are widely distributed among liverworts, mosses, hornworts, lycophytes, and monilophytes, and are divided into four groups (group I-IV), where group I and group II are more similar to the terpene synthases from bacteria and group III and group IV are more similar to terpene synthases from fungi (Jia et al., 2016, 2018). MTPSL enzymes have the same two important motifs for the catalytic activity as other plant terpene synthases, the NSD/DTE and DDxxD domains. The NSD/DTE domain is conserved between MTPSLs and other terpene synthases, but the aspartate-rich DDxxD motif varies between the groups as follows Group I: DDxxD, Group II: DDxxxD, Group III: DDxxx, and Group IV: DDxxD (Jia et al., 2016). It has been observed that 86 % of liverwort and 34 % of moss MTPSLs are in group I. The remaining 66 % of the MTPSLs in mosses and 50 % from hornworts are found in group II, whereas group III are found in low numbers in hornworts and liverworts. Group IV is primarily composed of MTPSLs from monilophytes and lycophytes, with no presence in bryophytes. (Jia et al., 2016). Thus, for this study the liverwort MTPSLs enzymes are most likely to belong to Group I, with a little possibility of being in Group III as well.

The genus *Lophocolea* have several species that are fragrant. In L.

* Corresponding author. Université Jean Monnet Saint-Étienne, CNRS, LBVpam UMR 5079, 42023, Saint-Étienne, France.
E-mail address: henrik.toft.simonsen@univ-st-etienne.fr (H.T. Simonsen).

https://doi.org/10.1016/j.plaphy.2025.110031
Received 3 September 2024; Received in revised form 4 May 2025; Accepted 15 May 2025
Available online 17 May 2025

bidentata, the fragrance has been identified as a mixture of 2-methylisoborneol, geosmin, and the geosmin epoxide ((+)-(4*S*,4a*S*,5*R*,8a*S*)-*trans*-4,8a-dimethyl-4a,5-epoxy-decalin), mono- and sesquiterpenes also found in few other liverwort species and many bacteria (Asakawa, 2007; Rieck et al., 1995; Spörle et al., 1991). The biosynthesis of geosmin is shown to begin with farnesyl diphosphate, going through a germacradienol intermediate, and then losing the 2-propanol side chain to form geosmin (Jiang et al., 2007). The biosynthesis of 2-methylisoborneol includes the elongation of dimethylallyl diphosphate (DMAPP) by 2-methylisopentenyl diphosphate, followed by cyclization to produce 2-methylisoborneol (Gu and Dickschat, 2022). Both terpenoids have unique biosynthetic pathways, mainly studied in bacteria, while their production in plants remains unexplored. Geosmin is not widely spread among liverworts, thus identification of the liverwort biosynthesis could reveal how these fragrances are made in plants, and be compared with bacteria.

Nardia scalaris is less known for its contents of sesquiterpenoids, and so far, no monoterpenids have been reported from *N. scalaris*. The sesquiterpenoid β-barbatene is the main compound in the raw extract of the plant (Asakawa et al., 1981) along with other barbatane-, cuparane- and bisabolane-type sesquiterpenes like *ent*-3,6-peroxycupar-1-ene, (S)-(−)-δ-cuparenol, isoafricanol, α-barbatene, cuparene, and 8-bisabolene (Asakawa, 1995; Asakawa et al., 1981; Langenbahn et al., 1993). These are unique structures that are so far only found in liverworts.

Overall, the study of microbial-like terpene synthases in liverworts provides new insights into the biosynthesis of terpenoids, especially terpenoids only found in this part of the plant kingdom. Here we have identified four MTPSL form *Nardia scalaris* and five from *Lophocolea bidentate*, and we have biochemically characterized five of these, with this being the first studies of terpene synthases from these two plants.

2. Materials and methods

2.1. Plant material collection

The plant material was collected in May 2019, Horserød Hegn, Denmark. 56.020977, 12.471669. The two species were identified by bryologist Irina Goldberg as being *Nardia scalaris* Grey (Gymnomitriaceae) and *Lophocolea bidentata* (L.) Dumort (Lophocoleaceae). The species are stored in the living collection of bryophytes at UJM.

2.2. RNA extraction and sequencing

The plant material was stored at 4 °C just after the collection, and shortly after the plant material was washed and cleaned from soil and other foreign material prior to RNA extraction.

The total RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma, STRN250). RNA integrity was initially confirmed by agarose gel electrophoresis and the visualization of intact ribosomal RNA bands. Subsequent RNA quality control was carried out on a 2100 Bioanalyzer (Agilent Technologies, Hørsholm, Denmark) and each sample received an RNA integrity number (RIN) of greater than 8.5.

RNA samples were submitted to Macrogen (Seoul, Korea) for stranded mRNA library preparation using Illumina Truseq Stranded mRNA library prep kit. The sequencing library was prepared by random fragmentation of the DNA or cDNA sample, followed by 5′ and 3′ adapter ligation. The sequencing was performed by Novaseq 150 bp paired-end sequencing providing 30–40 million reads per sample. The RNA seq data of *L. bidentata* and *N. scalaris* are available on NCBI SRA, with the BioSample numbers SAMN40967014 (*Lb*) and SAMN40967017 (*Ns*).

2.3. Transcriptome assembly and annotation

A *de novo* transcriptome was assembled as a reference for read mapping and Differential Expression (DE) analysis using Trinity v2.4.0. (Haas et al., 2013). As recommended in the Trinity protocol, one single Trinity assembly was generated by combining all reads across samples as input to ease following downstream analysis. Quality trimming and adapter removal was performed using trimmomatic with default parameters (Bolger et al., 2014). Transcript abundance was estimated using the alignment-based quantification method RSEM that uses Bowtie2 (Langmead and Salzberg, 2012), as an alignment method. Transcript and gene expression matrices were generated, and the numbers of expressed genes were calculated. Finally, DE analysis was performed at the gene level using edgeR with a dispersion rate of 0.1 (Robinson et al., 2009). Extractions and clustering of differentially expressed genes were performed with combinations of P-value cutoff for FDR of 10^{-3} and fold-change values of 2, 4, 16, 64 and 256. Functional annotation of the transcriptome was performed on homology searches towards SwissProt, PFAM and the Gene Onthology (GO) resource.

2.4. Validation of transcriptome assembly

Identification of primary metabolism genes was used as an assessment of the *N. scalaris* and *L. bidentata* transcriptome assembly quality. Selected primary metabolism genes from *Marchantia polymorpha v3.1* genome (Bowman et al., 2017) were used to BLAST search the transcriptome. Coding sequences of the following proteins were downloaded from Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html): Hexokinase (glycolysis), pyruvate kinase (glycolysis), Isocitrate dehydrogenase (TCA cycle), Malate dehydrogenase (TCA cycle) (Goodstein et al., 2012).

The *N. scalaris* and *L. bidentata* transcriptomes were searched with each coding sequence, using BLAST + software locally, to check if a homolog was present. Best hits were identified for each gene and the longest contig was used in a tBLASTx search of all six reading frames against the translated NR-nucleotide database, to check if best hits were the same primary metabolism genes from different species. Identification of four primary metabolism genes, expected to be expressed constitutively, serves as a qualitative validation of the success of transcriptome assembly. A BLAST search of the two transcriptomes with the *M. polymorpha* genes encoding hexokinase, isocitrate dehydrogenase, malate dehydrogenase and pyruvate kinase, all returned from 3 to 15 contigs with an E-value close to 0. The longest contig for each gene search was used in a BLASTx search that revealed gene orthologs from other plant species. Thus, the transcriptome assembly was deemed reliable.

2.5. Identification of MTPSL genes

A FASTA file containing the amino acid sequences representing the terpene synthases originating from bryophytes was compiled to use in identifying all possible terpene synthases (TPS) from the two transcriptome sets, this includes fourteen proteins from *M. polymorpha* (two hypothetical proteins, three diterpene synthases and nine MTPSLs), one from *Jungermannia subulate* (*ent*-kaurene synthases), two from *Scapania nemorea* (MTPSLs), three from *Anthoceros punctatus* (MTPSLs), two from *Anthoceros agrestis* (MTPSLs), one from *Physcomitrium patens* (*ent*-kaurene synthase), one from *Pseudotaxiphyllum elegans* (MTPSL), one from *Anomodon rostratus* (MTPSL), one from *Hypnum plumaeforme* (diterpene synthase), and one from *Sphagnum lescurii* (MTPSL).

The gene list was used to do a translated BLAST (tBLASTx) search of the *N. scalaris* and *L. bidentata* transcriptomes. The returned whole contig hits were in turn used to do a tBLASTn search of the NCBI non-redundant nucleotide database to inspect if the top hits would be terpene synthases. Following this, the protein Basic Local Alignment Search Tool (BLASTp) was used for protein sequences of the MTPSLs (Altschul et al., 1997; Shiryev et al., 2007). All searches were conducted from the NCBI BLAST web server (https://blast.ncbi.nlm.nih.gov/Blast.cgi), using default parameters. The top ten hits returned by each search were investigated if they had an E-value $\leq 10^{-5}$. Entry pages for each protein hit were inspected at the UniProt KnowledgeBase webpage

Table 1
The table show the identification of different compounds in the hexane extract of *Lophocolea bidentata*. The compounds were identified by comparison of the obtained MS and RI data with literature data as listed in the references (Adams, 2007; Andriamaharavo, 2014; Juliani et al., 2004; Karioti et al., 2004).

S.no.	RI (Ref)	Name	RI	Content Percentage	References for RI	Identified previously in the species
Lophocolea bidentata						
1.	978	1-octen-3-ol	978	0,87 %	(Adams, 2007)	(Ludwiczuk and Asakawa, 2015)
2.	1095	1-octen-3-yl-acetate	1109	1,44 %	(Adams, 2007)	(Ludwiczuk and Asakawa, 2015)
3.	1180	2-methylisoborneol	1197	1,18 %	(Adams, 2007)	(Asakawa, 2007)
4.	1389	β-elemene	1407	1,00 %	(Adams, 2007)	(Ludwiczuk and Asakawa, 2015)
5.	1457	β-barbatene	1475	4,51 %	(Yousefzadi et al., 2011)	(Ludwiczuk and Asakawa, 2015)
6.	1498	α-selinene	1521	3,25 %	(Adams, 2007)	(Ludwiczuk and Asakawa, 2015)
7.	1508	cuparene	1532	0,45 %	(Zhao et al., 2009)	(Ludwiczuk and Asakawa, 2015)
8.	1627	cubenol	1654	0,34 %	(Juliani et al., 2004)	(Ludwiczuk and Asakawa, 2015)
9.	981	(-)-β-pinene	984	0,56 %	(Karioti et al., 2004)	-
10.	1339	δ-elemene	1352	0,41 %	(Adams, 2007)	-
11.	1466	4-epi-α-acoradiene	1489	0,28 %	(Skaltsa et al., 2003)	-
12.	1478	γ-muurolene	1496	0,40 %	(Adams, 2007)	-
13.	1513	α-cuprenene	1527	0,76 %	(Andriamaharavo, 2014)	-
14.	1550	δ-cuprenene	1574	0,60 %	(Andriamaharavo, 2014)	-
15.	1649	β-eudesmol	1679	0,30 %	(Adams, 2007)	-
16.	-	unidentified	1252	2,33 %		
17.	-	unidentified	1717	0,46 %		
18.	-	unidentified	2326	19,3 %		
19.	-	unidentified	2389	0,32 %		
20.	-	unidentified	1505	0,65 %		
21.	-	unidentified	1384	16,2 %	} naptha-ketone	
22.	-	unidentified	2006	24,9 %		
23.	-	unidentified	2071	1,31 %	} sesquiterpene lactone	
24.	-	unidentified	2091	11,9 %		
25.	-	unidentified	2113	6,08 %		

Table 2
The table show the identification of different compounds in the hexane extract of *Nardia scalaris*. The compounds were identified by comparison of the obtained MS and RI data with literature data as listed in the references (Hnawia et al., 2008; Skaltsa et al., 2003; Zhao et al., 2009).

S.no.	RI (Ref)	Name	RI	Content Percentage	References for RI	Identified previously in the species
Nardia scalaris						
1.	1457	-barbatene	1440	3,00 %	(Yousefzadi et al., 2011)	(Rieck et al., 1995)
2.	1465	-barbatene	1476	10,0 %	(Hnawia et al., 2008)	(Asakawa et al., 1981)
3.		(-)- -cuparenol	1808	1,00 %	(Langenbahn et al., 1993)	(Langenbahn et al., 1993)
4	1466	4-epi- -acoradiene	1490	1,00 %	(Skaltsa et al., 2003)	-
5.	1491	2-isopropenyl-4 ,8-dimethyl-1,2,3,4,4 ,5,6,7-octahydronaphthalene	1497	1,00 %	(Andriamaharavo, 2014)	-
6.	1513	-cuprenene	1528	2,00 %	(Andriamaharavo, 2014)	-
7.	1508	cuparene	1533	3,00 %	(Zhao et al., 2009)	-
8.	1600	viridiflorol	1607	1,00 %	(Asuming et al., 2005)	-
9.	-	unidentified	1391	3,00 %		
10.	-	unidentified	1522	3,00 %		
11.	-	unidentified	1559	1,00 %		
12.	-	unidentified	1973	5,00 %		
13.	-	unidentified	2343	37,0 %		
14.	-	unidentified	2363	8,00 %		
15.	-	unidentified	1990	15,0 %		
16.	-	unidentified	2005	4,00 %		
17.	-	unidentified	2256	1,00 %	sesquiterpene lactone	

(https://www.uniprot.org/), using their respective accession numbers. Accession number, protein name, function and origin species and percent sequence identity were recorded for each protein hit. The InterProScan web server (https://www.ebi.ac.uk/interpro/search/sequence/) was used to verify the MTPSLs, based on InterPro signatures (Mitchell et al., 2019). 5 full length potential MTPSLs were identified from the transcriptome data set of *L. bidentata* (*Lb*MTPSL1-5) and 4 from *N. scalaris* (*Ns*MTPSL1-4). *Lb*MTPSL2 was found to not be full length. The sequences were confirmed by a PCR analysis on the original cDNA, and the genes are available from GenBank with the accessions PQ094571 - PQ094578, respectively. To confirm the localization of the proteins all eight sequences was analyzed using the TargetP tool (https://services.healthtech.dtu.dk/services/TargetP-2.0/) (Almagro Armenteros et al., 2019). This confirmed for all sequences that these were localized to the cytoplasm, and thus most likely would use farnesyl diphosphate as a substrate (supplementary data).

Additionally, one TPS gene, which aligned with *ent*-kaurene synthase from *Machantia polymorpha* was found in each transcriptome set (data not show as this was not studied further).

2.6. Multiple sequence alignment of MTPSLs

Amino acid sequences of the isolated MTPSLs were aligned using the Clustal Omega tool build into Geneious Prime software (Geneious Prime® 2024.0.5 Build, 2024-04-12 19:00), which also include a logo builder. MTPSLs from other liverworts were obtained from NCBI, and their protein accession numbers are used in the alignments (supplementary data). A three was constructed based on the alignment using the in-build tools in Geneious, which uses a Jukes-Cantor distance model and a Neighbor-Joining tree model, and provides the bootstrap values (see supplementary data).

2.7. Expression of MTPSL in S. cerevisiae

Gene sequences were synthetized from Twist Bioscience and subsequently amplified by PCR (Phusion U polymerase thermosfisher (reference F533S) with uracil extended primers and compatible with USER cloning system (New England Biolab). 10 μl of PCR product was directly cloned into 75 ng of pYEDP60u vector (Hamann and Møller, 2007), compatible with USER cloning system with 1 μl of USER enzyme. Successful cloning was verified by sanger sequencing (Eurofins genomics).

Plasmids were transferred to *S. cerevisiae* WAT11 strain (Urban et al., 1997) by electro transformation with 1 μg of plasmid and 50 μl of electro competent cells with biorad MicroPulser Electroporator (1.5 kV, 200 mA, 25 μF). After 3 h at 30 °C of regeneration in 1 M sorbitol yeast were plated on selection media (SGAI: yeast nitrogen base 6.7 g/L, glucose 20 g/L, tryptophan 20 μg/L, Adenine 0.2 mg/L agar 15 g/L). After 3 days at 30 °C, a single colony was cultured in 10 ml of YPGAE media (Yeast extract 10 g/L, bacto peptone 10 g/L, glucose 10 g/L, ethanol 3 %, adenine 0.2 mg/L) and grown for 60 h with agitation at 30 °C. Integration of the plasmids was confirmed by PCR. Three cell lines was obtained for further analysis in biological triplicates.

For biochemical analysis, a single colony was again grown in YPGAE media for 60 h with agitation at 30 °C, in 100 ml in a 1 L bottle. Induction was started by adding 10 ml of fresh buffered YPLAE media (Yeast extract 10 g/L, bacto peptone 10 g/L, galactose 10 g/L, ethanol 3 %, adenine 0.2 mg/L, Tris-HCL 20 mM pH 7). Volatiles emitted were trapped on 100 mg porapak Q 80/100 mesh from a 1 L bottle with stirring with a flow of about 400 ml/min for 72 h. For additional analysis of the *N. scalaris* 1 mg/L of farnesyl diphosphate and geranyl diphosphate was added to the YPGAE media in two different tests for the four enzymes. This did not provide additional peaks in the GC-MS analysis (see supplementary data).

Every 24 h, VOCs trapped on porapak are eluted with 1 ml of n-hexane before GC-MS analysis was performed following the same protocol as described below.

2.8. Extraction and chemical analysis of plant material

Fresh plant material (1g) was extracted with 10 ml hexane, this was concentrated to 100 μl to yield a raw extract, performed in triplicates. The raw extract and the eluted hexane samples from the different VOC traps of enzyme expressions (see above) were analyzed using an Agilent 9000 Intuvo gas chromatography system connected to an Agilent 5977B mass detector, Agilent Technologies. Temperature of guard chip was set at 65 °C at the beginning of the run before following the temperature of the oven in track oven mode and the bus temperature was set at 220 °C. The run time was 15.5 min, with an initial temperature of 50 °C, hold initial time was 1 min, and the temperature increase was set to 20 °C/min, ending at 300 °C with a final hold time of 2 min. The injection volume was 1 μL. The injection port was set to 250 °C, with a split ratio of 25:1. The column was an Agilent 19091S-433UI-INT: US20040208, HP-5MS UI, and flow rate of helium was 1.5 mL/min. The MSD transfer line temperature was set to 320 °C. The MS was set to 230 °C, and the MS Quad was set to 150 °C. The acquisition mode was in Scan mode from 30 to 300 in mass range.

Products were identified based on their retention times and electron ionization mass spectra compared to previously published data as given in the results part.

2.9. Homology-based structural modelling and docking

The three-dimensional structure of the MTPSLs was modelled using the modelling tool, CPHmodels 3.2 (Nielsen et al., 2010). The CPHmodels 3.2 code is available at https://services.healthtech.dtu.dk/service.php?CPHmodels-3.2. The MTPSLs amino acid sequences were submitted in FASTA format at the website under the tab “Submission”, returning a structural model for each of them. All models generated were submitted to “ProQ – Protein Quality Predictor” web server (Wallner and Elofsson, 2003). Returned results from the server were used to decide on models to be used for further analysis. For the models chosen through quality assessment, a manual search of the modelling templates was performed and visualized using the RCSB protein data bank webserver (Berman et al., 2000; Burley et al., 2023). A model of selinadiene synthase from *Streptomyces pristinaespiralis* in complex with dihydrofarnesyl pyrophosphate (https://doi.org/10.2210/pdb4OKZ/pdb) and epi-isozizaene synthase *Streptomyces coelicolor* A3(2) in complex with Mg, inorganic pyrophosphate and benzyl triethyl ammonium cation (https://doi.org/10.2210/pdb3KB9/pdb) were identified as common templates for the MTPSLs (Aaron et al., 2010; Baer et al., 2014). The models were downloaded from the PDB website (https://www.rcsb.org/). Using the PyMOL software (The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC), model structures were visualized and superposed. Distances between Mg^{2+} and coordinating residues were measured.

3. Results

3.1. Identification of terpenoids from raw extracts

In order to compare the enzymatic products with what is found in the plants, the hexane extracts of *Lophocolea bidentata* and *Nardia scalaris* were examined by GC-MS, and several compounds were identified. In L. *bidentata* 25 compounds could be observed, of these 15 was identified, and 8 of these have not been reported before in the species, but are found in sister species and other liverworts. For *N. scalaris* a total of 17 compounds were found, and 8 was identified, where only 3 have previously been reported from the plant, since there are only few studies on the mono- and sesquiterpenoids in *N. scalaris*. Tables 1 and 2 show the identified compounds for the GC-MS analysis for the two species.

Table 3
The lengths of coding sequences and of the translated protein sequences for *Lb*MTPSL1, 3,4,5 and *Ns*MTPSL1-4 are given. Here '*Lb*' is an abbreviation for *Lophocolea bidentata and 'Ns' for Nardia scalaris.*

	*Lb*MTPSL1	*Lb*MTPSL3	*Lb*MTPSL4	*Lb*MTPSL5	*Ns*MTPSL1	NsMTPSL2	NsMTPSL3	NsMTPSL4
Coding sequence length (bp)	1128	1287	1155	1164	1200	1152	1602	1050
Protein length (aa)	375	428	384	387	399	383	533	349

Fig. 1. Sequence alignment of *Lb*MTPSL1,3,4,5 and *Ns*MTPSL1-4 using Clustal omega alignment tool in Geneious software (Geneious Prime® 2024.0.5) with the inbuild logo design. The figure highlights the conserved DDxxD/E/x domains of the 8 MTPSLs, which confirm that these are indeed terpene synthases.

3.2. Identification of MTPSL genes from Lophocolea bidentata and Nardia scalaris

Transcriptome data was collected from both plants and assembled using the Trinity assembly pipeline. BLAST searches in the transcriptomes found five contig hits of putative *LbMTPSLs* for *Lophocolea bidentata* and four for *Nardia scalaris NsMTPSLs*. One of the five sequences from *L. bidentata* was incomplete, so only four were analyzed in this study. The full genome of both species may contain additional MTPSLs and TPSs. A re-BLAST search of the identified MTPSL sequences against NCBI primarily returned matches with sesquiterpene synthase activity, though with a few monoterpene synthases. All eight full length genes were analyzed for target peptides, and was found to be localized to the cytosol in the plant cells indicating a sesquiterpene synthase activity (supplementary data). The lengths of the MTPSL coding sequences and their translated proteins are listed in Table 3.

The substrate binding in terpene synthases is facilitated by two highly conserved motifs: the DDxxD/E and NSD/DTE motifs. While the NSD/DTE motif is highly conserved in MTPSLs, including those identified in this work (see full alignment in supplementary data), the aspartate-rich DDxxD/E motif tends to exist in different noncanonical forms (Jia et al., 2016). As seen in Fig. 1 it is observed that in *L. bidentata* there are two variations among the aspartate-rich motifs, *Lb*MTPSL1 and 4 display a DDxxE motif and *Lb*MTPSL3 and 5 display the canonical DDxxD motif. This places all four of them in group I of MTPSLs. For *N. scalaris* it is observed that *Ns*MTPSL1, 2 and 3 have the canonical DDxxD motif therefore placed in group I, whereas *Ns*MTPSL4 have an unusual DDxxx motif, and hence are part of the group III of MTPSLs (Jia et al., 2016). Additionally, *Ns*MTPSL3 is unusually long for MTPSLs, but a similar length is observed from other MTPSLs, like a MTPSL from *Jungermannia exsertifolia* (GenBank: UKS51569.1) and *Mp*MTPSL8 from *M. polymorpha* subsp. *ruderalis* (GenBank: BFI32343.1). This can be observed in the alignment (see supplementary data), that also together with the Neighbor-Joining tree show that *Ns*MTPSL 4 have an unusual sequence compared to other characterized MTPSLs.

3.3. Catalytic functions of liverwort MTPSLs

After identifying the MTPSLs, they were expressed in *Saccharomyces cerevisiae* WAT11 to investigate their functionality. GC-MS analysis of trapped volatiles revealed the following activities: *LbMTPSL1* was a γ-muurolene synthase, *LbMTPSL3* a cubenene synthase, *LbMTPSL5* a β-elemene/α-selinene synthase, *NsMTPSL3* a viridiflorol synthase, and *NsMTPSL4* primarily produced α- and β-chamigrene along with at least five minor products (Fig. 2, supplementary data). *LbMTPSL4*, *NsMTPSL1*, and *NsMTPSL2* were inactive in yeast, though further studies are needed. Feeding mutant yeast expressing the *N. scalaris* enzymes with farnesyl diphosphate or geranyl diphosphate did not yield additional products, confirming these enzymes function as sesquiterpene synthases and that *Ns*MTPSL2 is inactive in yeast (see supplementary data).

3.4. Three-dimensional structure models and docking of MTPSLs

Tertiary structure models of the MTPSLs were investigated by model visualization using PyMOL and the models build in CPHmodels. The purpose was to visualize the positioning of the Mg^{2+} and FPP binding sites. MTPSLs, including the structures examined in this study, exhibit a predominantly α-helical conformation, with some disordered coil regions connecting the α-helices. Superposition of MTPSLs on each other shows high structural conservation. All the models of the MTPSLs shared structural similarity and are structural analogous with the *Streptomyces* selinadiene synthase. The comparison with selinadiene synthase was used to confirm the position of the DDxxD/E motif in all the enzymes, as demonstrated by the overlay of selinadiene synthase with *LbMTPSL3* (Fig. 3). Along with the sequence alignment (Fig. 1), this confirmed both the presence and position of the DDxxD/E motif in all active enzymes. This information was then used to classify the enzymes into their respective MTPSL groups.

4. Discussion

In this work, eight full length putative MTPSL genes from the liverworts *L. bidentata* and *N. scalaris,* four from each, were examined in terms of their functionality as terpene synthases. Among these, five were successfully characterized by expressing them in yeast and analyzing the volatile compounds emitted by the yeast. The raw extracts of the two species were also analyzed, and in *L. bidentata* it was observed that the three characterized enzymes were responsible for the biosynthesis of some of the extracted compounds. γ-muurolene was produced by *Lb*MTPSL1 in yeast and identified from the *L. bidentata* extracts. This is also true for the products β-elemene and α-selinene from *Lb*MTPSL5 that was found in the extracts here, and are also previously isolated from the plant (Ludwiczuk and Asakawa, 2015). It is important to note that β-elemene products may result from Cope rearrangements of germacrene sesquiterpenoids (Andersen et al., 2017), and that MTPSL

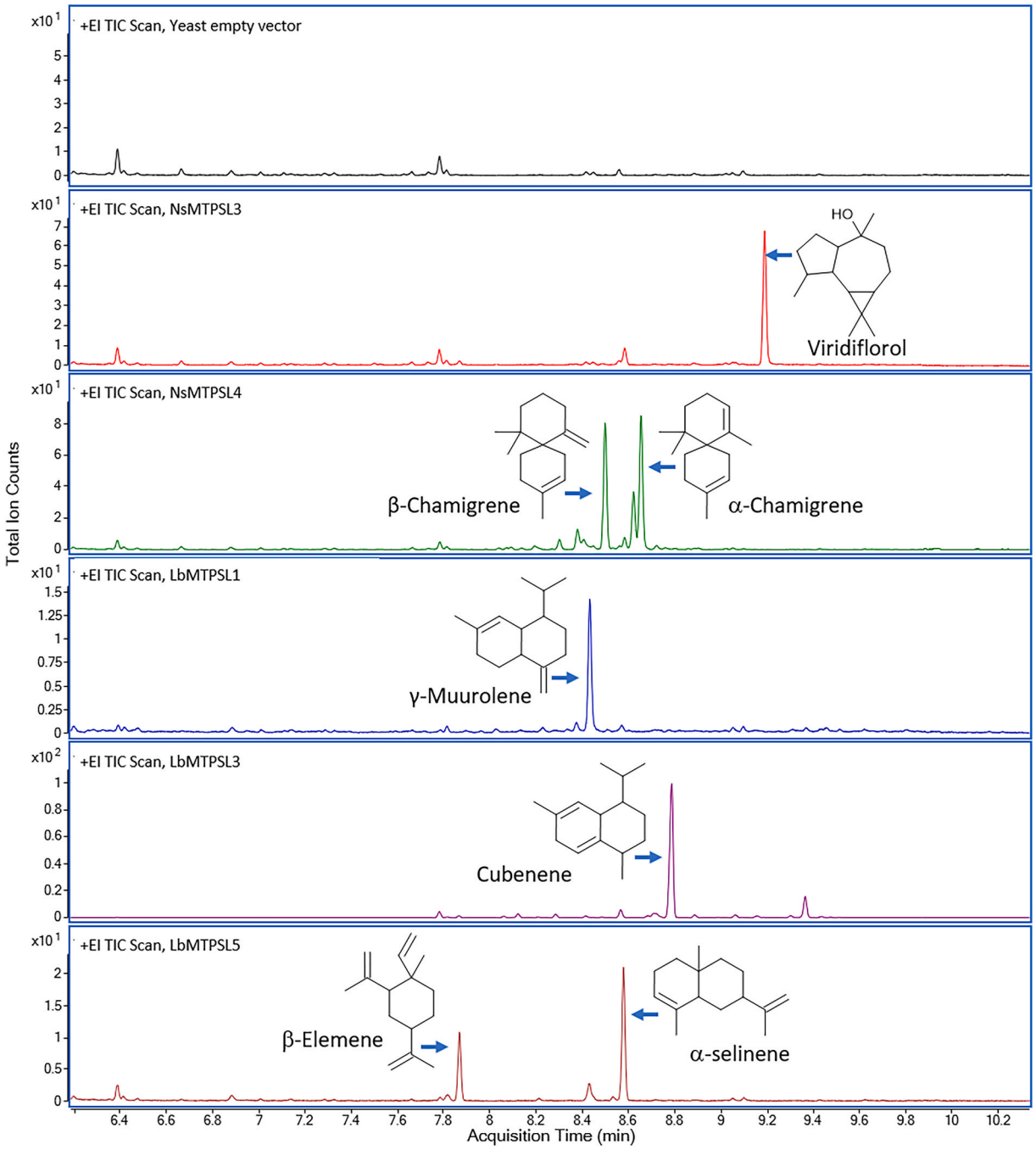

Fig. 2. GC-MS chromatograms of the hexane extraxt of the VOC collection that was obtained during cultivation of yeast expressing the 8 different MTPSLs. The structures of the identified molecules were determined based on comparison of the MS spectral data and the obtained RI values with literature data.

germacrene synthases have been characterized from other liverworts (Fan et al., 2021). However, we were unable to determine whether this occurred in our investigation, necessitating further studies. The product of *Lb*MTPSL3, cubenene, was not seen in the plant extracts properly due to low abundance. Likewise cubenene is a likely precursor for other terpenoids found in *L. bidentata* like cubenol through a hydroxylation of C7 in the molecule. This reaction could very likely be performed by an yet unidentified cytochrome P450. A comprehensive characterization of all MTPSLs in *L. bidentate* requires not only the isolation of the remaining fragments of *LbMTPSL2 and expression of the yet inactive enzymes in a different host (e.g. plants),* but also the identification and analysis of currently uncharacterized genes of which we have preliminary evidence that there are at least two more MTPSLs. There is still the question on how geosmin is produced in plants, and if this resemble the biosynthesis found in bacteria. Of the obtained molecules, α-selinene from *Lb*MTPSL5 could serve as a precursor for geosmin, through a reduction of the double bond, a hydroxylation and a los of isopropanyl side chain to yield the final product. It also remains to be investigated what is the ecological function of these compounds or their derivatives. Several terpenoids, especially sesquiterpene lactones various ecological and pharmacological functions (Simonsen et al., 2013), but this is not very well studied in bryophytes.

In *N. scalaris Ns*MTPSL3 produced viridiflorol, a compound previously isolated from the plant and also detected in the plant extracts in this study (Asuming et al., 2005). Meanwhile, *Ns*MTPSL4 predominantly produced α- and β-chamigrene, which were not observed in the extracts.

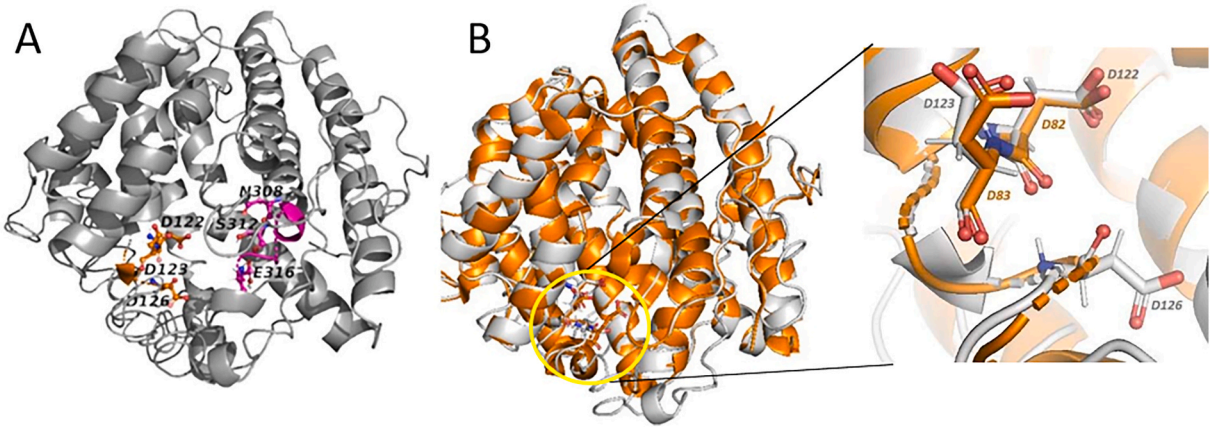

Fig. 3. Modeling of *Lb*MTPSL3. A) Structural model of *Lb*MTPSL3 obtained by using CPHmodels showing metal binding motifs D122D123xxD126 (orange); N308S312E316 (magenta). B) Superposition of *Lb*MTPSL3 (grey) model with X-ray structure of top hit template selinadiene synthase from *Streptomyces pristinaespiralis* (orange) (PDB ID: 4okm); The D122D123xxD126 region is in the circle and the inset shows the closeup view of DDxxD superposition. The figure shows how the plant MTPSL have structural relationship with the bacterial TPS. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

However, these compounds might serve as precursors for other metabolites in the plant not yet identified. *Ns*MTPSL4 also produced a range of minor products, where one of them could be barbatene, which is seen in the plant and have been shown to be the major compound among mono- and sesquiterpenoids in the plant extracts previously (Yousefzadi et al., 2011). The identification of these compounds requires further work, including characterization of all MTPSLs and isolation of the compounds. As barbatene is rather unique it is of interest to clarify the biosynthesis of this compound in plants.

One of the characteristics of MTPSL genes is that they encode terpene synthases with a simple domain architecture of a single α-domain (Jia et al., 2016, 2018; Li et al., 2012). They lack the β- and γ-domains of typical plant TPS and are shorter proteins, in generally around 350 amino acids in length. The *Ns*MTPSL1, 2 and 4 and the *Lb*MTPSLs appear to conform to these observed features. They range from 349 to 428 amino acids in length and only have the α-domain. *Ns*MTPSL3 differ from this generality, with a length of 533 amino acids. To emphasize that this is a generalization, it is worth mentioning that a review the data from other MTPSL studies reveals that many of the proteins range near 500 amino acids. Other functionally characterized MTPSLs from e.g. *M. polymorpha* have a protein lengths of above 450 amino acids (Jia et al., 2018; Kumar et al., 2016). Hence, the lengths of the MTPSLs identified in this work are in line with what other studies have observed.

MTPSLs belong to class I terpene synthases and use the DDxxD and NSE/DTE motifs for catalytic activity, which were also found in this study and several other studies of MTPSLs from bryophytes and other non-seed plants (Jia et al., 2018; Li et al., 2012, 2012, Xu et al., 2024a, 2024b; Yan et al., 2023; Zhao et al., 2021). While the NSE/DTE motif is highly conserved in MTPSLs, the DDxxD motif exhibits variations. This pattern is also observed in the MTPSLs of *L. bidentata* and *N. scalaris*. Among the eight full-length proteins analyzed, only NsMTPSL4 did not fall into the group of enzymes with a conserved DDxxD/E motif, which is the predominant group among the MTPSLs functionally characterized from other bryophytes, whether they are mono- and sesquiterpene synthases (Bowman et al., 2017; Conart and Simonsen, 2025; Kumar et al., 2016; Takizawa et al., 2021). However, the motif seen in *Ns*MTPL4 is not exclusive to this enzyme but rather common. Though the glutamine in the *Ns*MTPSL4 motif is not charged like the regular aspartate residue contained in these motifs, glutamine might still be able to function as part of the Mg^{2+} ion coordination in the enzyme. This hypothesis is based on the solved crystal structures of TPSs, where it is evident that the asparagine of the NSE motif, clearly functions as ion coordinator (Baer et al., 2014). Because asparagine and glutamine share similar physio-chemical properties, as they both have an amide side chain, the glutamine residue might be functional in the DDxxQ motif of *Ns*MTPSL4. Alanine substitutions of conserved motif residues have shown that availability of electrons in the DDxxX motifs are crucial for the functionality. Substitutions of aspartate to alanine in this region in MTPSL enzymes from *Anthoceros punctatus* and *Anthoceros agrestis* led to non-active enzymes (Xiong et al., 2018).

A recent study that compare microbial and plant terpene synthases found that the position of the asymmetric diphosphate- $(Mg^{2+})_3$ cluster in the enzymes and the subsequent positioning of the hydrocarbon tail in the enzyme display enantiomeric stereochemistry (Schwartz et al., 2024), likewise a few have been investigated for the mechanistic and showing some interesting stereochemical properties going through a germacrene intermediate towards more complex sesquiterpene skeletons (Schwartz et al., 2024; Xu et al., 2024a), which potentially could explain the biosynthesis of some of the compounds found here like viridiflorol and chamigrene, this though require further studies. Our enzymes align well with the microbial template from *Streptomyces*, however further modelling is needed to confirm that the MTPSLs here have the same enantiomeric chemistry as microbial terpene synthases or if they are like regular TPSs from plants, these studies are ongoing. This observed difference, could explain the enantiomeric chemistry of terpenoids seen between bryophytes and angiosperms (Ludwiczuk and Asakawa, 2019). Studies to explain this difference is ongoing along with further characterization of the remaining enzymes.

5. Conclusion

Five MTPSL enzymes from the two liverworts *Lophocolea bidentata* and *Nardia scalaris*, *Lb*MTPSL1, 3, and 5 and *Ns*MTPSL3 and 4 were characterized as murolene synthase (*Lb*MTPSL1), cubenene synthase (*Lb*MTPSL3), α-selinene/β-elemene synthase (*Lb*MTPSL5), viridiflorol synthase (*Ns*MTPSL3), and chamigrene synthase (*Ns*MTPSL4). The characterization was based on heterologous expression in yeast, and provides insight into the biosynthesis of terpenoids in these two liverworts. Several other MTPSL genes were also found, but did not show activity under the present settings. Thus, further work is needed to determine their activity and also to provide data on the *in-planta* activity of the characterized enzymes. The enzymes described might provide the first step towards a description of sesquiterpene lactone biosynthesis in liverworts though several more enzymes need to be tested to obtain such knowledge.

Ethics approval and consent to participate

Not applicable.

Author contributions

CC, RB and HTS conceived and designed the experiments; CC, BSB, SKK and HTS collected the plant material; CC, BSB, CG, MS, SSR, SKK and RB performed the experiments and analyzed data. CC, BB and HTS wrote the manuscript. All authors have read and approved the final manuscript.

Availability of data

In this study, transcriptomes of the liverworts *Nardia scalaris* and *Lophocolea bidentata* are made available on NCBI SRI with the BioSample numbers SAMN40967014 (*Lb*) and SAMN40967017 (*Ns*).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Corentin Conartt reports financial support was provided by Fondation de le Universite Jean Monnet ELAN Recherche grant. Rituraj Batth reports financial support was provided by Villum Experiment Villum Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Irina Goldberg for her valuable help with the preliminary identifications of the liverworts. This research was funded by Fondation de l'Université Jean Monnet, ELAN Recherche grant for CC and Villum Experiment grant for RB.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphy.2025.110031.

Data availability

Data will be made available on request.

References

Aaron, J.A., Lin, X., Cane, D.E., Christianson, D.W., 2010. Structure of epi-isozizaene synthase from *Streptomyces coelicolor* A3(2), a platform for new terpenoid cyclization templates. Biochemistry 49, 1787–1797. https://doi.org/10.1021/bi902088z.

Adams, R.P., 2007. Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry. Allured Pub Corp, Carol Stream, IL, USA.

Almagro Armenteros, J.J., Salvatore, M., Emanuelsson, O., Winther, O., Von Heijne, G., Elofsson, A., Nielsen, H., 2019. Detecting sequence signals in targeting peptides using deep learning. Life Sci. Alliance 2, e201900429. https://doi.org/10.26508/lsa.201900429.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Andersen, T.B., Martinez-Swatson, K.A., Rasmussen, S.A., Boughton, B.A., Jørgensen, K., Andersen-Ranberg, J., Nyberg, N., Christensen, S.B., Simonsen, H.T., 2017. Localization and in-Vivo characterization of *Thapsia garganica* CYP76AE2 indicates a role in thapsigargin biosynthesis. Plant Physiol. 174, 56–72. https://doi.org/10.1104/pp.16.00055.

Andriamaharavo, N.R., 2014. Retention Data.

Asakawa, Y., 2007. Biologically active compounds from bryophytes. Pure Appl. Chem. 79, 557–580. https://doi.org/10.1351/pac200779040557.

Asakawa, Y., 1995. Chemical constituents of the bryophytes. In: Progress in the Chemistry of Organic Natural Products. Springer, pp. 1–562.

Asakawa, Y., Ludwiczuk, A., Nagashima, F., 2013. Chemical constituents of Bryophyta. In: Progress in the Chemistry of Organic Natural Products, Progress in the Chemistry of Organic Natural Products. Springer, Vienna, Vienna, pp. 563–605. https://doi.org/10.1007/978-3-7091-1084-3_5.

Asakawa, Y., Matsuda, R., Toyota, M., Suire, C., Takemoto, T., Inoue, H., Hattori, S., Mizutani, M., 1981. Chemosystematics of bryophytes VIII. The distribution of terpenoids and aromatic compounds in Japanese hepaticae and anthocerotae. J. Hattori Bot. Lab. 50, 165–177. https://doi.org/10.18968/jhbl.50.0_165.

Asuming, W.A., Beauchamp, P.S., Descalzo, J.T., Dev, B.C., Dev, V., Frost, S., Ma, C.W., 2005. Essential oil composition of four Lomatium Raf. species and their chemotaxonomy. Biochem. Systemat. Ecol. 33, 17–26. https://doi.org/10.1016/j.bse.2004.06.005.

Baer, P., Rabe, P., Fischer, K., Citron, C.A., Klapschinski, T.A., Groll, M., Dickschat, J.S., 2014. Induced-fit mechanism in class I terpene cyclases. Angew. Chem. Int. Ed. 53, 7652–7656. https://doi.org/10.1002/anie.201403648.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242. https://doi.org/10.1093/nar/28.1.235.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S.S., Althoff, F., Araki, T., Arteaga-Vazquez, M.A., Balasubrmanian, S., Barry, K., Bauer, D., Boehm, C.R., Briginshaw, L., Caballero-Perez, J., Catarino, B., Chen, F., Chiyoda, S., Chovatia, M., Davies, K.M., Delmans, M., Demura, T., Dierschke, T., Dolan, L., Dorantes-Acosta, A. E., Eklund, D.M., Florent, S.N., Flores-Sandoval, E., Fujiyama, A., Fukuzawa, H., Galik, B., Grimanelli, D., Grimwood, J., Grossniklaus, U., Hamada, T., Haseloff, J., Hetherington, A.J., Higo, A., Hirakawa, Y., Hundley, H.N., Ikeda, Y., Inoue, K., Inoue, S.-I., Ishida, S., Jia, Q., Kakita, M., Kanazawa, T., Kawai, Y., Kawashima, T., Kennedy, M., Kinose, K., Kinoshita, T., Kohara, Y., Koide, E., Komatsu, K., Kopischke, S., Kubo, M., Kyozuka, J., Lagercrantz, U., Lin, S.-S., Lindquist, E., Lipzen, A.M., Lu, C.-W., De Luna, E., Martienssen, R.A., Minamino, N., Mizutani, Masaharu, Mizutani, Miya, Mochizuki, N., Monte, I., Mosher, R., Nagasaki, H., Nakagami, H., Naramoto, S., Nishitani, K., Ohtani, M., Okamoto, T., Okumura, M., Phillips, J., Pollak, B., Reinders, A., Rövekamp, M., Sano, R., Sawa, S., Schmid, M.W., Shirakawa, M., Solano, R., Spunde, A., Suetsugu, N., Sugano, S., Sugiyama, A., Sun, R., Suzuki, Y., Takenaka, M., Takezawa, D., Tomogane, H., Tsuzuki, M., Ueda, T., Umeda, M., Ward, J.M., Watanabe, Y., Yazaki, K., Yokoyama, R., Yoshitake, Y., Yotsui, I., Zachgo, S., Schmutz, J., 2017. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. Cell 171, 287–304.e15. https://doi.org/10.1016/j.cell.2017.09.030.

Burley, S.K., Bhikadiya, C., Bi, C., Bittrich, S., Chao, H., Chen, L., Craig, P.A., Crichlow, G. V., Dalenberg, K., Duarte, J.M., Dutta, S., Fayazi, M., Feng, Z., Flatt, J.W., Ganesan, S., Ghosh, S., Goodsell, D.S., Green, R.K., Guranovic, V., Henry, J., Hudson, B.P., Khokhriakov, I., Lawson, C.L., Liang, Y., Lowe, R., Peisach, E., Persikova, I., Piehl, D.W., Rose, Y., Sali, A., Segura, J., Sekharan, M., Shao, C., Vallat, B., Voigt, M., Webb, B., Westbrook, J.D., Whetstone, S., Young, J.Y., Zalevsky, A., Zardecki, C., 2023. RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. Nucleic Acids Res. 51, D488–D508. https://doi.org/10.1093/nar/gkac1077.

Conart, C., Simonsen, H.T., 2025. Tamariscol biosynthesis in *Frullania tamarisci*. Phytochemistry 229, 114301. https://doi.org/10.1016/j.phytochem.2024.114301.

Fan, H., Wei, G., Chen, X., Guo, H., Crandall-Stotler, B., Köllner, T.G., Chen, F., 2021. Sesquiterpene biosynthesis in a leafy liverwort Radula lindenbergiana Gottsche ex C. Hartm. Phytochemistry 190, 112847. https://doi.org/10.1016/j.phytochem.2021.112847.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40, D1178–D1186. https://doi.org/10.1093/nar/gkr944.

Gu, B., Dickschat, J.S., 2022. A non-natural biosynthesis pathway toward 2-methylisoborneol. Chem. Commun. 58, 4316–4319. https://doi.org/10.1039/D2CC00636G.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512. https://doi.org/10.1038/nprot.2013.084.

Hamann, T., Møller, B.L., 2007. Improved cloning and expression of cytochrome P450s and cytochrome P450 reductase in yeast. Protein Expr. Purif. 56, 121–127. https://doi.org/10.1016/j.pep.2007.06.007.

Hayashi, K., Kawaide, H., Notomi, M., Sakigi, Y., Matsuo, A., Nozaki, H., 2006. Identification and functional analysis of bifunctional ent-kaurene synthase from the moss *Physcomitrella patens*. FEBS (Fed. Eur. Biochem. Soc.) Lett. 580, 6175–6181. https://doi.org/10.1016/j.febslet.2006.10.018.

Hnawia, E., Menut, C., Agrebi, A., Cabalion, P., 2008. Wood essential oils of two endemic trees from New Caledonia: *callitris sulcata* (Parl.) Schltr. and *Callitris neocaledonica* Dummer. Biochem. Systemat. Ecol. 36, 859–866. https://doi.org/10.1016/j.bse.2008.08.007.

Horn, A., Pascal, A., Lončarević, I., Volpatto Marques, R., Lu, Y., Miguel, S., Bourgaud, F., Thorsteinsdóttir, M., Cronberg, N., Becker, J.D., Reski, R., Simonsen, H.T., 2021. Natural products from bryophytes: from basic biology to biotechnological applications. Crit. Rev. Plant Sci. 40, 191–217. https://doi.org/10.1080/07352689.2021.1911034.

Jia, Q., Köllner, T.G., Gershenzon, J., Chen, F., 2018. MTPSLs: new terpene synthases in nonseed plants. Trends Plant Sci. 23, 121–128. https://doi.org/10.1016/j.tplants.2017.09.014.

Jia, Q., Li, G., Köllner, T.G., Fu, J., Chen, X., Xiong, W., Crandall-Stotler, B.J., Bowman, J. L., Weston, D.J., Zhang, Y., Chen, L., Xie, Y., Li, F.W., Rothfels, C.J., Larsson, A., Graham, S.W., Stevenson, D.W., Wong, G.K.S., Gershenzon, J., Chen, F., 2016. Microbial-type terpene synthase genes occur widely in nonseed land plants, but not in seed plants. Proc. Natl. Acad. Sci. U. S. A 113, 12328–12333. https://doi.org/10.1073/pnas.1607973113.

Jiang, J., He, X., Cane, D.E., 2007. Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme. Nat. Chem. Biol. 3, 711–715. https://doi.org/10.1038/nchembio.2007.29.

Juliani, H.R., Zygadlo, J.A., Scrivanti, R., De La Sota, E., Simon, J.E., 2004. The essential oil of Anemia tomentosa (Savigny) Sw. var. anthriscifolia (Schrad.) Mickel. Flavour & Fragrance J 19, 541–543. https://doi.org/10.1002/ffj.1341.

Karioti, A., Hadjipavlou-Litina, D., Mensah, M.L.K., Fleischer, T.C., Skaltsa, H., 2004. Composition and antioxidant activity of the essential oils of *Xylopia aethiopica* (Dun) A. Rich. (Annonaceae) leaves, stem bark, root bark, and fresh and dried fruits, growing in Ghana. J. Agric. Food Chem. 52, 8094–8098. https://doi.org/10.1021/jf040150j.

Karunanithi, P.S., Zerbe, P., 2019. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. Front. Plant Sci. 10.

Kumar, S., Kempinski, C., Zhuang, X., Norris, A., Mafu, S., Zi, J., Bell, S.A., Nybo, S.E., Kinison, S.E., Jiang, Z., Goklany, S., Linscott, K.B., Chen, X., Jia, Q., Brown, S.D., Bowman, J.L., Babbitt, P.C., Peters, R.J., Chen, F., Chappell, J., 2016. Molecular diversity of terpene synthases in the liverwort *Marchantia polymorpha*. Plant Cell 28, 2632–2650. https://doi.org/10.1105/tpc.16.00062.

Langenbahn, U., Burkhardt, G., Becker, H., 1993. Diterpene malonates and other terpenes from *Nardia succulenta* and *N. scalaris*. Phytochemistry 33, 1173–1179. https://doi.org/10.1016/0031-9422(93)85044-R.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.

Li, G., Köllner, T.G., Yin, Y., Jiang, Y., Chen, H., Xu, Y., Gershenzon, J., Pichersky, E., Chen, F., 2012. Nonseed plant *Selaginella moellendorffii* has both seed plant and microbial types of terpene synthases. Proc. Natl. Acad. Sci. USA 109, 14711–14715. https://doi.org/10.1073/pnas.1204300109.

Ludwiczuk, A., Asakawa, Y., 2019. Bryophytes as a source of bioactive volatile terpenoids – a review. Food Chem. Toxicol. 132, 110649. https://doi.org/10.1016/j.fct.2019.110649.

Ludwiczuk, A., Asakawa, Y., 2015. Chemotaxonomic value of essential oil components in liverwort species. Rev. Flavour & Fragrance J. 30, 189–196. https://doi.org/10.1002/ffj.3236.

Marques, R.V., Salwinski, A., Enemark-Rasmussen, K., Gotfredsen, C.H., Lu, Y., Hocquigny, N., Risler, A., Duval, R.E., Miguel, S., Bourgaud, F., Simonsen, H.T., 2023. Extracts from the liverwort *Bazzania trilobata* with potential dermo-cosmetic properties. In: Murthy, H.N. (Ed.), Bioactive Compounds in Bryophytes and Pteridophytes, Reference Series in Phytochemistry. Springer International Publishing, Cham, pp. 147–164. https://doi.org/10.1007/978-3-031-23243-5_9.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.-Y., El-Gebali, S., Fraser, M.I., Gough, J., Haft, D.R., Huang, H., Letunic, I., Lopez, R., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Nuka, G., Orengo, C., Pandurangan, A.P., Paysan-Lafosse, T., Pesseat, S., Potter, S.C., Qureshi, M.A., Rawlings, N.D., Redaschi, N., Richardson, L.J., Rivoire, C., Salazar, G.A., Sangrador-Vegas, A., Sigrist, C.J.A., Sillitoe, I., Sutton, G. G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Yong, S.-Y., Finn, R.D., 2019. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. 47, D351–D360. https://doi.org/10.1093/nar/gky1100.

Nielsen, M., Lundegaard, C., Lund, O., Petersen, T.N., 2010. CPHmodels-3.0-remote homology modeling using structure-guided sequence profiles. Nucleic Acids Res. 38, W576–W581. https://doi.org/10.1093/nar/gkq535.

Rieck, A., Bülow, N., König, W.A., 1995. An epoxy-trinoreudesmane sesquiterpene from the liverwort *Lophocolea bidentata*. Phytochemistry 40, 847–851. https://doi.org/10.1016/0031-9422(95)00355-B.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2009. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

Sabovljević, M.S., Sabovljević, A.D., Ikram, N.K.K., Peramuna, A., Bae, H., Simonsen, H. T., 2016. Bryophytes – an emerging source for herbal remedies and chemical production. Plant Gen. Resour. 14, 314–327. https://doi.org/10.1017/S1479262116000320.

Schwartz, R., Zev, S., Major, D.T., 2024. Differential substrate sensing in terpene synthases from plants and microorganisms: insight from structural, bioinformatic, and EnzyDock analyses. Angew. Chem. Int. Ed. 63, e202400743. https://doi.org/10.1002/anie.202400743.

Shiryev, S.A., Papadopoulos, J.S., Schäffer, A.A., Agarwala, R., 2007. Improved BLAST searches using longer words for protein seeding. Bioinformatics 23, 2949–2951. https://doi.org/10.1093/bioinformatics/btm479.

Simonsen, H.T., Weitzel, C., Christensen, S.B., 2013. Guaianolide sesquiterpenoids: pharmacology and biosynthesis. In: Ramawat, K.G., Merillon, J.M. (Eds.), Natural Products: Phytochemistry, Botany and Metabolism of Alkaloids, Phenolics and Terpenes, Handbook of Natural Products. Springer-Verlag, Berlin, Germany, pp. 3069–3098. https://doi.org/10.1007/978-3-642-22144-6_134.

Skaltsa, H.D., Demetzos, C., Lazari, D., Sokovic, M., 2003. Essential oil analysis and antimicrobial activity of eight Stachys species from Greece. Phytochemistry 64, 743–752. https://doi.org/10.1016/S0031-9422(03)00386-8.

Spörle, J., Becker, H., Salazar Allen, N., ahabir G upta, M.P., 1991. Occurrence of (-)-Geosmin and other terpenoids in an axenic culture of the liverwort *Symphyogyna brongniartii*. Z. Naturforsch. 46c (1), 3–188.

Takizawa, R., Hatada, M., Moriwaki, Y., Abe, S., Yamashita, Y., Arimitsu, R., Yamato, K. T., Nishihama, R., Kohchi, T., Koeduka, T., Chen, F., Matsui, K., 2021. Fungal-type terpene synthases in *Marchantia polymorpha* are involved in sesquiterpene biosynthesis in oil body cells. Plant Cell Physiol. 62, 528–537. https://doi.org/10.1093/pcp/pcaa175.

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., Pompon, D., 1997. Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. J. Biol. Chem. 272, 19176–19186.

Wallner, B., Elofsson, A., 2003. Can correct protein models be identified? Protein Sci. 12, 1073–1086. https://doi.org/10.1110/ps.0236803.

Xiong, W., Fu, J., Köllner, T.G., Chen, X., Jia, Q., Guo, Haobo, Qian, P., Guo, Hong, Wu, G., Chen, F., 2018. Biochemical characterization of microbial type terpene synthases in two closely related species of hornworts, *Anthoceros punctatus* and *Anthoceros agrestis*. Phytochemistry 149, 116–122. https://doi.org/10.1016/j.phytochem.2018.02.011.

Xu, H., Köllner, T.G., Chen, F., Dickschat, J.S., 2024a. Mechanistic characterisation of a sesquiterpene synthase for asterisca-1,6-diene from the liverwort *Radula lindenbergiana* and implications for pentalenene biosynthesis. Org. Biomol. Chem. 22, 1360–1364. https://doi.org/10.1039/D3OB02088F.

Xu, H., Köllner, T.G., Chen, F., Dickschat, J.S., 2024b. Functional and mechanistic characterization of the 4,5-diepi-Isoishwarane synthase from the liverwort *Radula lindenbergiana*. Chembiochem 25, e202400104. https://doi.org/10.1002/cbic.202400104.

Yan, X., Li, Y., Li, W., Liang, D., Nie, S., Chen, R., Qiao, J., Wen, M., Caiyin, Q., 2023. Transcriptome analysis and identification of sesquiterpene synthases in liverwort *Jungermannia exsertifolia*. Bioengineering 10. https://doi.org/10.3390/bioengineering10050569.

Yousefzadi, M., Heidari, M., Akbarpour, M., Mirjalili, M.H., Zeinali, A., Parsa, M., 2011. *In vitro* cytotoxic activity of the essential oil of *Dorema ammoniacum* D. Don. Middle East J. Sci. Res. 7, 511–514.

Zhan, X., Bach, S.S., Hansen, N.L., Lunde, C., Simonsen, H.T., 2015. Additional diterpenes from *Physcomitrella patens* synthesized by copalyl diphosphate/kaurene synthase (PpCPS/KS). Plant Physiol. Biochem. 96, 110–114. https://doi.org/10.1016/j.plaphy.2015.07.011.

Zhao, C., Zeng, Y., Wan, M., Li, R., Liang, Y., Li, C., Zeng, Z., Chau, F., 2009. Comparative analysis of essential oils from eight herbal medicines with pungent flavor and cool nature by GC–MS and chemometric resolution methods. J. Separ. Sci. 32, 660–670. https://doi.org/10.1002/jssc.200800484.

Zhao, Y., Hu, T., Liu, R., Hao, Z., Liang, G., Li, G., 2021. Biochemical characterization and function of eight microbial type terpene synthases from lycophyte *Selaginella moellendorffii*. Int. J. Mol. Sci. 22. https://doi.org/10.3390/ijms22020605.