Physiol Mol Biol Plants
https://doi.org/10.1007/s12298-021-00986-4

SHORT COMMUNICATION

De novo transcriptome sequencing assisted identification of terpene synthases from black pepper (*Piper nigrum*) berry

Johnson K. George[1] · Sreekumar Shelvy[1] · Abdulkabeer Muhammed Fayad[1] · Palaniyandi Umadevi[1,3] · U. B. Angadi[2] · Mir Asif Iquebal[2] · Sarika Jaiswal[2] · Anil Rai[2] · Dinesh Kumar[2]

Received: 21 December 2020 / Revised: 26 February 2021 / Accepted: 1 April 2021

Abstract Though the volatile profiles of black pepper have been reported already, the information on terpene synthase family genes is not known. In this study, using a combinatorial approach, the berry hybrid transcriptome assembly of llumina and nanopore sequencing, the entire terpene synthase family responsible for the biosynthesis of the flavor-imparting volatiles in black pepper berries was profiled. The profile shows 98 terpene synthases from various terpene synthesis pathways. Three important monoterpene synthases were also validated by targeted amplification, sequencing and homology modeling. This study provides the first of its kind information on the terpene synthase family profile in *Piper nigrum,* which is potentially a major step for further characterization of the functional terpene synthase genes in black pepper.

Keywords *Pn*TPS · *Denovo* transcriptome assembly · Berry transcriptome · Terpenoid pathway · Terpenes · *Piper nigrum*

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12298-021-00986-4.

✉ Palaniyandi Umadevi
Umadevi.P@icar.gov.in

✉ Dinesh Kumar
Dinesh.Kumar@icar.gov.in

[1] ICAR – Indian Institute of Spices Research, Kozhikode, Kerala, India

[2] ICAR – Indian Agricultural Statistics Research Institute, New Delhi, India

[3] Rice Breeding and Genetics Research Center, ICAR–IARI, Aduthurai, Tamil Nadu, India

Introduction

Terpenes constitute the largest class of structurally diverse metabolites in plants. Based on the number of 5-carbon units, terpenoids are classified into monoterpene (C10), sesquiterpenes (C15), diterpenes (C20), triterpenes (C30), tetraterpenes (C40) and polyterpenes which have more than eight 5C units. Apart from being as aroma compounds, these volatiles are the ecologically important info chemicals. They can act as chemical messengers for pollinators, deterrents of herbivores and also mediates plant to plant signals in an ecological system. Each particular terpene synthesis is decided by the presence of unique TPS genes in plants. Black pepper (*Piper nigrum* L.) is a tropical perennial plant native to India. The economic part of black pepper is the berry (matured dry fruits), used in food and pharmaceutical industries. The flavor or aroma of black pepper is attributed by the volatile oil and the pungency is by piperine (Ravindran and Kallupurackal 2001) The de-novo transcriptome assembly has been applied to identify the genes involved in the biosynthesis of secondary metabolites in spice plants viz., *Cinnamomum camphora* (Chen et al. 2018) and *Piper nigrum* (Hu et al. 2015; Jin et al. 2018). Till now there are only two reports on berry transcriptome in black pepper. Hu et al. (2015) used the transcriptome of black pepper fruit to investigate piperine biosynthesis while Jin et al. (2018) used the transcriptome of unripe berry for characterizing sesquiterpene synthases. The terpene synthase family responsible for the biosynthesis of the flavor-imparting volatiles in black pepper berries is still not determined. The identification of terpene synthases would pave the way to understanding the evolutionary and environmental significance of the terpenes in *Piper* species and further help to characterize the important functional terpene synthase genes in black pepper. Hence,

Published online: 04 May 2021

in the present study, using a combinatorial approach viz., hybrid transcriptome assembly, the entire terpene synthase family responsible for the biosynthesis of the flavor-imparting volatiles in black pepper berries was profiled. Three important monoterpene synthases were also validated by targeted amplification and sequencing.

Materials and methods

Genome annotation from berry transcriptome

The developed hybrid transcriptome assembly (Bio Sample accession: SAMN13981803) was used (Kokkat et al. 2021) for the global annotation of TPS present in the berry transcriptome. The unigenes were annotated by BLAST+ executables with protein sequences having an e value of 10^{-10} and KAAS – KEGG automatic annotation server (Moriya et al. 2007) was used for the functional annotation of the transcripts. HMMER 3.0 hmm motif-based search (Wheeler and Eddy 2013) was also attempted against three HMM profiles for terpene synthase family (PF03936, PF01397 and PF06330) to expand the identification of TPS from the hybrid assembly.

Isolation of 3 monoterpene synthase genes

The three unigenes from monoterpene synthases from the KAAS–KEGG identification were used to isolate the gene from the berry cDNA. In order to reconfirm the identity, the unigene sequences of linalool synthase (LS), neomenthol dehydrogenase (NMD) and 1–8 cineol synthase (18CS) from the transcriptome were retrieved from Uniprot (Accession numbers: AAC49395.1—*Clarkia breweri,* NM_001324565.1—*Capsicum annum,* and NM_001281287.1—*Vitis vinefera* respectively). Gene specific primers were designed (Table 1) using the IDT PrimerQuest tool (https://eu.idtdna.com/pages/tools/primerquest). cDNA library was prepared using RevertAid First Strand cDNA Synthesis Kit (Thermofisher, USA). 25 pmol OligoDT primer was used for the reverse transcription.

Table 1 Gene specific primers for Linalool synthase, Neomenthol dehydrogenase and 1,8 Cineol synthase

Sl. No	Oligo Name	Sequence 5′–3′
1	PNLS_F	GAAGAGAGTTCTGGACCAAAGG
2	PNLS_R	CCCTAGTGGAATTGGTTCAACA
3	PNNMD_F	TGGAACAAGCTTTGGAGTAGG
4	PNNMD_R	GATGGGCATTGGAAACAAACA
5	PN1,8CS_F	GCAACCGAAATGTTGACAGTAG
6	PN1,8CS_R	CCCAGCGGAATTGGATTAACTA

PCR amplification of the target genes (*Pn*LS, *Pn*NMD and *Pn*18CS) was carried out using EmeraldAmp® GT PCR Master Mix (Takara bio, India) with the following protocol 94 °C for 5 min, 94 °C for 1 min, 61 °C, 62 °C and 64 °C (respectively for *Pn*LS, *Pn*NMD and *Pn*18CS) for 1 min, 72 °C for 1 min for 35 cycles and 72 °C for 5 min PCR products were purified by using GenElute Gel Extraction Kit (Sigma-Aldrich, USA) and sequenced.

Sequence analysis and phylogenetic tree

The amplification products were sequenced using the Sanger sequencing. The domains present in the sequences were identified using NCBI Conserved Domain Search Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), ORFs were identified with NCBI ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/) and the homologous protein families were identified using InterProScan—EMBL—EBI (http://www.ebi.ac.uk/interpro/search/sequence/). MEGA X software (Kumar et al. 2018) was used for the construction of phylogenetic tree. The amino acid sequences were aligned using MUSCLE (Edgar 2004). The evolutionary history was inferred using the UPGMA method (Sneath and Sokal 1973). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965) and are in the units of the number of amino acid substitutions per site. Plant species and accession numbers used for the construction of tree, the sequence alignment to compare the sequence similarity with other species were given in Supplementary files 1 and 2.

Homology modelling

The 3D model of *Pn*LS was constructed by considering the sequence identity of 45.83% to the template structure Isoprene synthase (3n0f.pdb) of *Populus canescens. Pn*NMD was modelled by considering the sequence identity of 64.13% to the template structure Menthone neomenthol reductase (5l51.pdb) of *Mentha piperita. Pn*18CS was modelled by the sequence identity of 42.86% with the template structure Limonene synthase (2onh.pdb) of *Mentha spicata* using SWISS MODEL workspace (Waterhouse et al. 2018). The active site of the modelled proteins and the docking of the substrate into the active sites were analysed using Schrodinger suit (Schrödinger Release 2018).

Results and discussion

Berry transcriptome

Hu et al. (2015) used the transcriptome of black pepper fruit to investigate piperine biosynthesis while Jin et al. (2018) used the transcriptome of unripe berry for characterizing sesquiterpene synthases. The terpene synthase family responsible for the biosynthesis of the flavor-imparting volatiles in black pepper berries is still not reported. To identify genes involved in terpenoid biosynthesis in black pepper, the transcriptome sequencing with 2 sequencing platforms and hybrid transcriptome assembly was developed from the berries of variety IISR -Thevam. RNA was extracted from the samples and sequenced (Illumina and Nanopore) data was used for the analysis. The illumina sequencing generated a total of 38,153,296 (38.1 million) raw reads. The nanopore produced 4,27,224 raw reads with an average read length of 859 bp and maximum read length of 6039 bp. The *de-novo* hybrid transcriptome assembly produced 308,369 scaffolds with an average scaffold length of 1,119 bp and a maximum length of 27,769 bp (Table 2). The reads generated using illumina platform from this study were more than the reported berry transcriptomes (Jin et al. 2018; Hu et al. 2019).

Prediction of genes for terpene synthesis pathways

The KAAS-KEGG analysis yielded 17 important secondary metabolite pathways from the berry transcriptome (Fig. 1). The analysis revealed that more than 50 percent of the unigenes were mapped to flavonoid, brassinosteroid, carotenoid, phenylpropanoid and terpenoid biosynthesis pathways. This correlates with the nature of essential oil profile in black pepper. The black pepper essential oils are normally rich in piperine and terpenes. The percentage accumulation of the unigenes in phenyl propanoid (60.53%) and lysine biosynthesis (33.33%) suggested the presence of the piperine synthesis pathway as these two pathways are the backbone pathways for piperine synthesis. The other major accumulation of unigenes was towards the terpene synthesis pathways [terpenoid backbone (61%), zeatin, flavonoid, brassinosteroid, carotenoid (above 50%) and the terpenoids mono, sesqui, di terpenoid (more than 20%]. A total of 30 terpenoid back bone pathway genes, 3 monoterpenoid, 9 diterpenoid, 2 sesquiterpenoid, 12 flavonoid, 5 zeatin, 20 carotenoid and 8 brassinosteroid pathway genes were predicted (Table 3). The representative KEGG pathway figure for monoterpenoid and flavonoid is presented as Supplementary File. 3. Though the volatile profiles of black pepper have been reported already (Dosoky et al. 2019) the information on terpene synthase family genes is not known. In this study, we have deduced the profile of TPS family unigenes from this crop for the first time., we have also identified 98 putative functional terpene synthase genes in black pepper through transcriptome mining and by motif-based search. The transcriptome mining alone yielded 89 unigenes with the associated terpenoid synthesis pathway by KEGG analysis.

Table 2 De-novo hybrid transcriptome assembly statistics of *P. nigrum*

Scaffolds generated	308,369
Maximum Scaffold length (bp)	27,769
Minimum Scaffold length (bp)	89
Average Scaffold length (bp)	1119
Median Scaffold length (bp)	173
Total Scaffolds length (bp)	345,060,046
Total number of non-ATGC characters	214,451
Percentage of Non-ATGC characters	0.062
Scaffolds $\geq$ 100 bp	308,365
Scaffolds $\geq$ 200 bp	271,562
Scaffolds $\geq$ 500 bp	197,948
Scaffolds $\geq$ 1 Kbp	119,118
Scaffolds $\geq$ 10 Kbp	119,118
N50 value	1831

Identification of *Pn*TPS by HMMER and motif-based search

In order to expand the identification of TPS family genes in black pepper, the transcriptome was searched against the Pfam-A database locally using the HMMER 3.0 with an E value of $1e^{-5}$. The analysis of two HMM profiles viz., Terpene_synth_C (PF03936) and Terpene Synthase N terminal domain (PF01397) from Pfam database, the putative terpene synthases yielded a total 105 TPS similar genes of different plant species. The retrieved longest peptides from the transcriptome data were analyzed for the presence of terpene synthase motifs. From the terpene synthase motifs DDXXD, DXDD, RR(× 8)W, EDXXD and DTE based search 9 unigenes of terpene synthases were identified. These include 8 monoterpenoid synthase (α-terpineol synthase and α-phellandrene synthase), sesquiterpenoid synthases (Cadinene synthase, Copaene synthase, Caryophyllene synthase, α-guaiene synthase, α-guaiene 2 oxidase) and 1 diterpene synthase (ent—Copalyl diphosphate synthase).

The present study identified the *Pn*TPS from 8 different terpene synthesis pathways from motif based search. Jin et al. (2018), identified only 3 sesquiterpene synthases from unripe berry transcriptome and did not report the genes for guaiene skeleton, a direct precursor to the peppery flavor

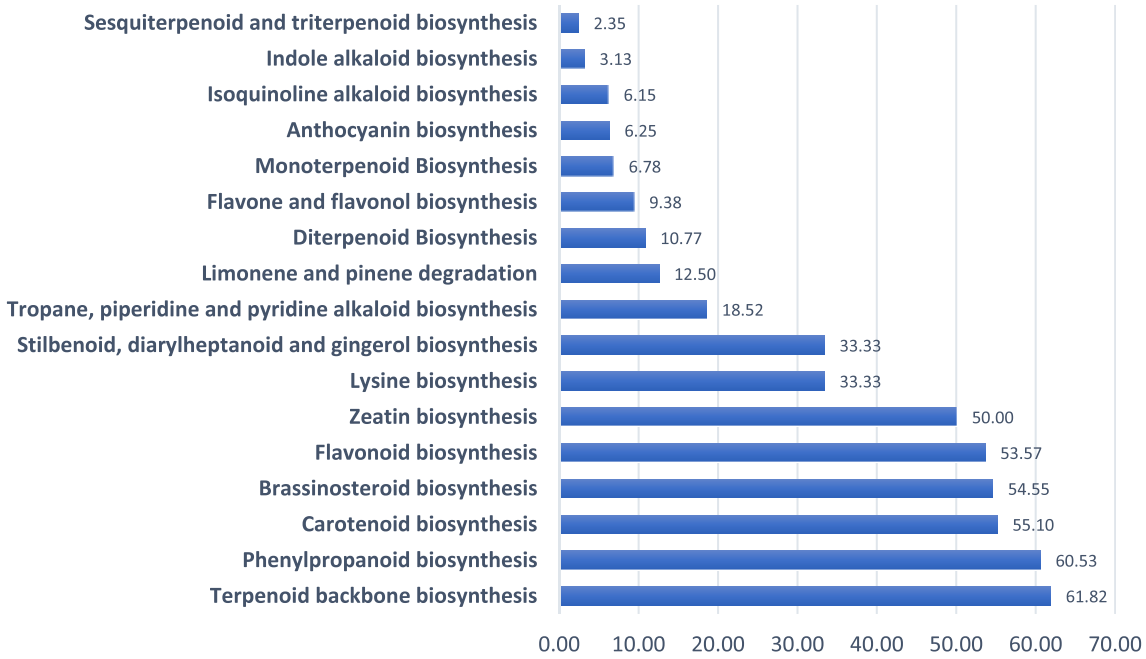

Fig. 1 Percentage enrichment of enzymes annotated in *P. nigrum* transcriptome

compound rotundone. But in this study, the unigenes for the important backbone genes of rotundone synthesis viz., guaiene synthase and oxidase were identified from the hybrid transcriptome through motif-based search.

Isolation and characterization of 3 monoterpene synthase genes

By the targeted amplification, 1 kb of Neomenthol dehydrogenase, above 1 kb of Linalool synthase and 1 kb 1,8 Cineol synthase genes were isolated (Fig. 2). The Blast analysis of the sequenced products (733 bp of NMD, 1269 bp of LS, 1107 bp of 1,8 CS) showed 73.3% similarity for Neomenthol dehydrogenase with *Hevea brasilensis*, 69.9% identity for Linalool synthase with *Coffea arabica* and 72.18% identity for 1,8 cineole with *Arabidopsis thaliana*. The InterProScan search revealed the homologues protein families and signal peptides. The conserved motifs viz., 'DTE', TGXXXGXG (SDR Family) and DDXXD were also identified in LNLS, NMD and 1,8 cineole gene respectively. These TPS from black pepper were named as *Pn*LS, *Pn*NMD and *Pn*18CS. From the phylogenetic analysis (Fig. 3), *Pn*LS sequence in the main clade of the tree was related to *Mentha aquatica*, *Artemisa annua*, *Solanum lycopersicum*, *Coffea arabica* and *Ocimum basilicum*. The *Pn*NMD sequence is related only to *Mucuna pruriens* and differed from the main two clades of neomenthol dehydrogenase. In the case of *Pn*18CS, it was clustered with the major 13 1,8-cineol synthase homologues.

Homology modelling of the PnLS protein by using Isoprene synthase from *Populus canescens* (3n0f) as the template (Fig. 4a) showed 15 amino acids in the active site and the DTE motif of terpene synthase family near to the active site. *Pn*LS were docked with the substrate geranyl pyrophosphate (Fig. 4b). Three amino acid molecules (ASP 34, ASP 90 and LEU 73) were found to be actively participating in the ligand binding.

*Pn*1,8CS were modelled with Isoprene synthase (Fig. 4c) protein derived from *Populus canescens* showed the binding of 14 amino acids in the active site and the terpene synthase motif DDXXD near to the active site. The docking studies with the substrate α-terpineol (Fig. 4d) revealed the active participation of THR 89, LEU 68, TYR 71 and ASP 103 amino acids for the substrate binding. *Pn*NMD modelled with Menthone neomenthol reductase (5L51) from *Mentha piperita* (Fig. 4e) revealed that 12 amino acids present in the active site and the SDR motif TGxxxGxG was located in the lower side of the active site. The docking studies with the substrate menthofuran (Fig. 4f) showed ARG 79 and the GLY 54 from TGxxxGxG motif forming H-bond with ligand molecule.

The targeted isolation of 3 monoterpene genes viz., the putative *Pn*LS, *Pn*NMD and *Pn*18CS proved the presence of the gene in black pepper. Around 30 plant families emit 1,8 cineole (Knudsen et al. 2006) and so far cineole synthases were isolated from *Arabidopsis*, *Citrus, Nicotiana*

Table 3 Identified unigenes in *P. nigrum* transcriptome

Sl. No	Name	Kegg Orthology	Definition
Terpenoid backbone biosynthesis			
1	E2.3.1.9	K00626	acetyl-CoA C-acetyltransferase
2	:2.2.1.7	K01662	1-deoxy-D-xylulose-5-phosphate synthase
3	E2.3.3.10	K01641	Hydroxymethyl glutaryl-CoA synthase
4	1.1.1.267	K00099	1-deoxy-D-xylulose-5-phosphate reductoisomerase
5	1.1.1.34	K00021	hydroxymethylglutaryl-CoA reductase (NADPH)
6	2.7.7.60	K00991	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
7	2.7.1.36	K00869	mevalonate kinase
8	2.7.1.148	K00919	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
9	2.7.4.2	K00938	phosphomevalonate kinase
10	4.6.1.12	K01770	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
11	4.1.1.33	K01597	diphosphomevalonate decarboxylase
12	1.17.7.1	K03526	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase
13	1.17.7.4	K03527	4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase
14	5.3.3.2	K01823	isopentenyl-diphosphate Delta-isomerase
15	2.5.1.1	K00787	farnesyl diphosphate synthase
16	1.1.1.216	K15890	NADP + -dependent farnesol dehydrogenase
17	2.7.1.216	K15892	farnesol kinase
18	2.5.1.10	K00787	farnesyl diphosphate synthase
19	1.8.3.6	K05906	Prenyl cysteine oxidase / farnesyl cysteine lyase
20	3.1.1.-	K15889	Prenyl cysteine alpha-carboxyl methyl esterase
21	2.5.1.82	K05355	hexaprenyl-diphosphate synthase
22	1.3.1.83	K10960	geranylgeranyl diphosphate/geranylgeranyl-bacteriochlorophyllide a reductase
23	2.5.1.85	K05356	all-trans-nonaprenyl-diphosphate synthase
24	:2.5.1.84	K05356	all-trans-nonaprenyl-diphosphate synthase
25	2.1.1.100	K00587	protein-S-isoprenylcysteine O-methyltransferase
26	3.4.24.84	K06013	STE24 endopeptidase
27	3.4.22	K08658	prenyl protein peptidase
28	2.5.1.58	K05955	protein farnesyltransferase/geranylgeranyl transferase type-1 subunit alpha
29	:2.5.1.87	K11778	ditrans,polycis-polyprenyl diphosphate synthase
30	2.5.1.31	K00806	undecaprenyl diphosphate synthase
Monoterpenoid biosynthesis			
1	4.2.3.25	K15086	(3S)-linalool synthase
2	4.2.3.108	K07385	1,8-cineole synthase
3	1.1.1.208	K15095	(+)-neomenthol dehydrogenase
Diterpenoid biosynthesis			
1	5.5.1.13	K04120	ent-copalyl diphosphate synthase
2	5.5.1.13	K20657	ent-copalyl diphosphate/ent-kaurene synthase
3	4.2.3.19	K04121	ent-kaurene synthase
4	1.14.14.86	K04122	ent-kaurene oxidase
5	1.14.14.86	K21292	ent-kaurene oxidase
6	1.14.14.107	K04123	ent-kaurenoic acid monooxygenase
7	1.14.11.12	K05282	gibberellin-44 dioxygenase
8	1.14.11.15	K04124	gibberellin 3beta-dioxygenase
9	1.14.11.13	K04125	gibberellin 2beta-dioxygenase
Sesquiterpenoid and Triterpenoid biosynthesis			
1	2.5.1.21	K00801	farnesyl-diphosphate farnesyltransferase
2	1.14.14.17	K00511	squalene monooxygenase
Carotenoid biosynthesis			
1	2.5.1.32	K02291	15-cis-phytoene synthase
2	1.3.5.5	K02293	15-cis-phytoene desaturase

Table 3 continued

Sl. No	Name	Kegg Orthology	Definition
3	5.2.1.12	K15744	zeta-carotene isomerase
4	1.3.5.6	K00514	zeta-carotene desaturase
5	5.2.1.13	K09835	prolycopene isomerase
6	1.3.99.26 1.3.99.28 1.3.99.29 1.3.99.31	K10027	phytoene desaturase
7	1.3.5.5	K02293	15-cis-phytoene desaturase
8	1.3.5.6	K00514	zeta-carotene desaturase
9	CrtL-b 5.5.1.19	K06443	lycopene beta-cyclase
10	CruA/P 5.5.1.19	K14605	lycopene cyclase CruA
11	CrtL-e 5.5.1.18	K06444	lycopene epsilon-cyclase
12	LUT5 1.14.–	K15747	beta-ring hydroxylase
13	CrtR-b 1.14.15.24	K15746	beta-carotene 3-hydroxylase
14	1.14.14.158	K09837	carotenoid epsilon hydroxylase
15	1.14.15.21	K09838	zeaxanthin epoxidase
16	1.23.5.1	K09839	violaxanthin de-epoxidase
17	1.13.11.51	K09840	9-cis-epoxycarotenoid dioxygenase
18	1.2.3.14	K09842	abscisic-aldehyde oxidase
19	1.1.1.288	K09841	xanthoxin dehydrogenase
20	1.14.14.137	K09843	(+)-abscisic acid 8′-hydroxylase
Zeatin biosynthesis			
1	2.5.1.112	K10760	adenylate dimethylallyl transferase (cytokinin synthase)
3	CYP735A	K10717	cytokinin trans-hydroxylase
4	2.5.1.75	K00791	tRNA dimethylallyl transferase
5	2.5.1.27	K10760	adenylate dimethylallyl transferase (cytokinin synthase)
6	1.5.99.12	K00279	cytokinin dehydrogenase
Flavanoid biosynthesis			
1	2.3.1.133	K13065	shikimate O-hydroxycinnamoyl transferase
2	1.14.14.91	K00487	trans-cinnamate 4-monooxygenase
3	2.3.1.74	K00660	chalcone synthase
4	5.5.1.6	K01859	chalcone isomerase
5	1.14.14.82	K05280	flavonoid 3′-monooxygenase
6	1.14.11.9	K00475	naringenin 3-dioxygenase
7	1.14.20.6	K05278	flavanol synthase
8	1.14.20.4	K05277	anthocyanidin synthase
9	1.3.1.77	K08695	anthocyanidin reductase
10	1.14.14.96	K09754	5-O-(4-coumaroyl)-D-quinate 3′-monooxygenase
11	2.1.1.104	K00588	caffeoyl-CoA O-methyltransferase
12	2.3.1.74, 2.3.1.170	K00660	chalcone synthase
Brassinosteroid biosynthesis pathway			
1	1.14.13.-	K09587	steroid 22-alpha-hydroxylase
2	1.14.13.-	K12639	cytochrome P450 family 724 subfamily B polypeptide 1
3	1.14.14.147	K12637	3-epi-6-deoxocathasterone 23-monooxygenase
4	1.14.14.147	K12638	3-epi-6-deoxocathasterone 23-monooxygenase
5	1.14.-.-	K09588	cytochrome P450 family 90 subfamily A polypeptide 1
6	1.3.1.22	K09591	steroid 5-alpha-reductase
7	1.14.-.-	K09590	brassinosteroid-6-oxidase 1
8	1.14.-.-	K12640	brassinosteroid-6-oxidase 2

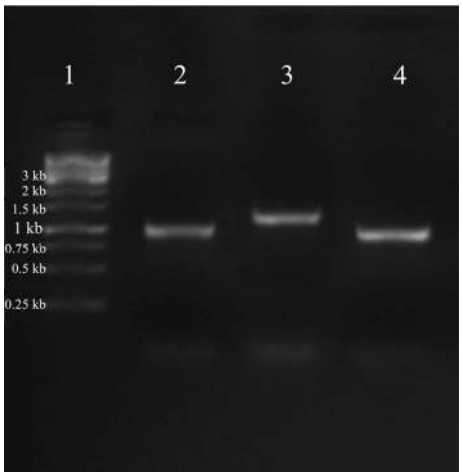

Fig. 2 Targeted gene amplification: Lane 1: 1 kb Gene ruler, Lane 2: Neomenthol dehydrogenase, Lane 3: Linalool synthase Lane 4: 1,8 Cineol synthase

and *Salvia*. Using the leaf transcriptome, a putative linalool synthase gene was isolated from *C. camphora* (Chen et al. 2018). Using degenerate primer approach linalool synthase genes was cloned and sequenced from Mentha (Crowell et al. 2002), Apricot fruit (Wang and Fan 2009), lemon myrtle (Sugiura et al. 2011), and citrus (Shimada et al. 2014). Neomenthol reductase is a member of the short-chain dehydrogenase/reductase (SDR) superfamily. Davis et al. (2005) had demonstrated the cloning, expression, and characterization of menthone reductases from Peppermint. Choi et al. (2008), isolated and functionally characterized a neomenthol reductase (CaMNR1), from pepper (*Capsicum annuum*) leaves infected by *Xanthomonas campestris pv vesicatoria* (Xcv) and also its ortholog SDR gene of Arabidopsis AtSDR1. They suggested the possible role of this monoterpene synthase gene in the regulation plant defenses against a broad spectrum of pathogens. In the present study. from the transcriptome data of black pepper berry three putative genes encoding important monoterpenes were isolated.

In conclusion, this study provides the first of its kind information on the terpene synthase family profile in *Piper nigrum*. Nine sesquiterpene synthases using motif-based search were added which are otherwise not identified by the KASS–KEGG analysis. This is potentially a major step

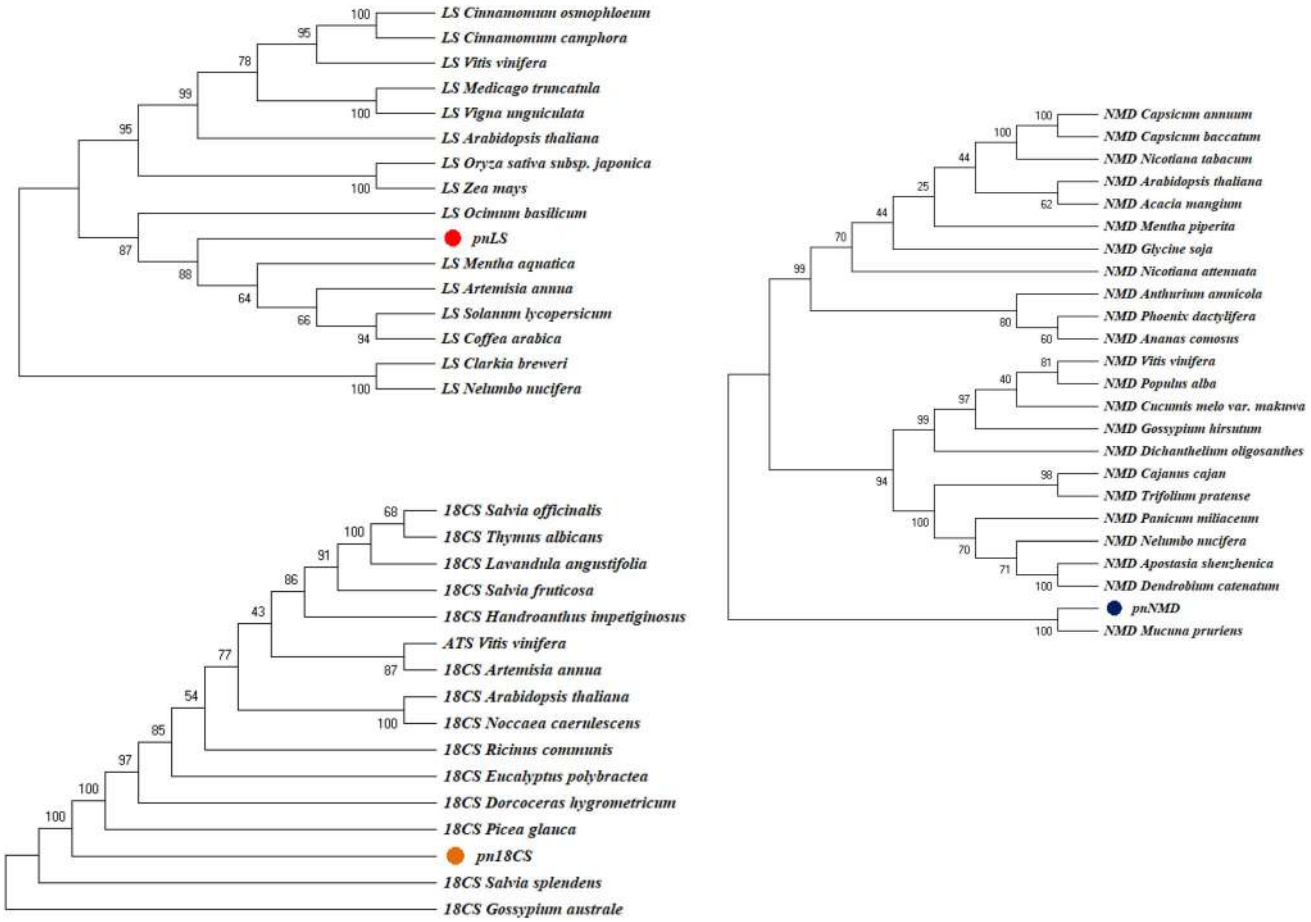

Fig. 3 Phylogeny of *Pn*LS, *Pn*18CS and *Pn*NMD

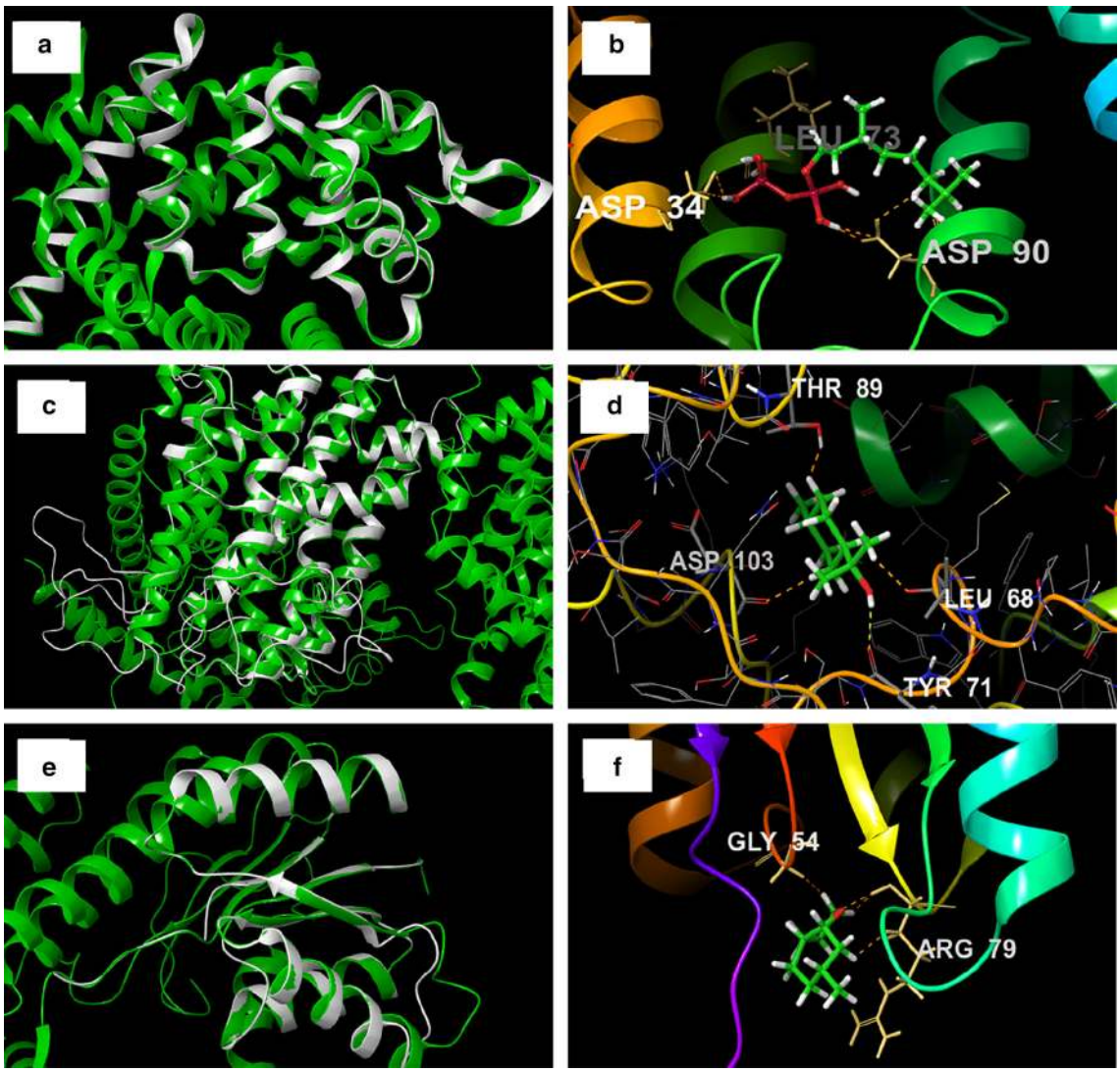

Fig. 4 **a** Template structure Isoprene synthase (3n0f) of *Populus canescens* is superimposed into *Pn*LS. **b** Geranyl pyrophosphate in the active site of *Pn*LS, **c** Template structure Isoprene synthase of *Populus canescens* is superimposed into *Pn*18CS, **d** α-terpineol in the active site of *Pn*18CS, **e** Template structure Menthofuran neomenthol reductase (5L51) of *Mentha piperita* is superimposed into *Pn*NMD, **f** Mentho furan in the active site of *Pn*NMD

for further characterizing the functional terpene synthase genes in the black pepper. Furthermore, it will also pave the way to understanding the evolutionary and environmental significance of the terpenes in *Piper* species.

Acknowledgements Authors are thankful to Indian Council of Agricultural Research, Ministry of Agriculture and Farmers' Welfare, Govt. of India for providing financial assistance in the form of CABin grant (F. no. Agril. Edn.4-1/2013-A&P) as well as Advanced Super Computing Hub for *Omics* Knowledge in Agriculture (ASHOKA) facility at Indian Council of Agricultural Research-Indian Agricultural Statistics Research Institute, New Delhi, India. Authors are thankful to Director, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India and Director, ICAR-Indian Institute of Spice Research, Calicut, India for providing research facility.

Funding This research was funded by Indian Council of Agricultural Research, Ministry of Agriculture and Farmers' Welfare, Govt. of India for providing financial assistance in the form of CABin grant (F. no. Agril. Edn.4–1/2013-A&P).

Declarations

Competing interests On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Chen C, Zheng Y, Zhong Y et al (2018) Transcriptome analysis and identification of genes related to terpenoid biosynthesis in

Cinnamomum camphora. BMC Genomics. https://doi.org/10.1186/s12864-018-4941-1

Choi HW, Lee BG, Kim NH et al (2008) A role for a menthone reductase in resistance against microbial pathogens in plants. Plant Physiol 148(1):383–401

Crowell AL, Williams DC, Davis EM et al (2002) Molecular cloning and characterization of a new linalool synthase. Arch Biochem Biophys 405(1):112–121. https://doi.org/10.1016/S0003-9861(02)00348-X

Davis EM, Ringer KL, McConkey ME, Croteau R (2005) Monoterpene metabolism. Cloning, expression, and characterization of menthone reductases from peppermint. Plant Physiol. https://doi.org/10.1104/pp.104.053306

Dosoky NS, Satyal P, Barata LM et al (2019) Volatiles of black pepper fruits (*Piper nigrum* L.). Molecules. https://doi.org/10.3390/molecules24234244

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340

Hu L, Hao C, Fan R et al (2015) De novo assembly and characterization of fruit transcriptome in black pepper (*Piper nigrum*). PLoS ONE. https://doi.org/10.1371/journal.pone.0129822

Hu L, Xu Z, Wang M et al (2019) The chromosome-scale reference genome of black pepper provides insight into piperine biosynthesis. Nat Commun 10:4702. https://doi.org/10.1038/s41467-019-12607-6

Jin Z, Kwon M, Lee AR et al (2018) Molecular cloning and functional characterization of three terpene synthases from unripe fruit of black pepper (*Piper nigrum*). Arch Biochem Biophys. https://doi.org/10.1016/j.abb.2017.12.011

Knudsen JT, Eriksson R, Gershenzon J, Ståhl B (2006) Diversity and distribution of floral scent. Bot Rev 72:1. https://doi.org/10.1663/0006-8101(2006)72[1:DADOFS]2.0.CO;2

Kokkat JG, Shelvy S, Fayad AM et al (2021) In silico assisted identification of peppery aroma compound 'rotundone' backbone genes from black pepper. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2021.1883113

Kumar S, Stecher G, Li M et al (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. https://doi.org/10.1093/molbev/msy096

Moriya Y, Itoh M, Okuda S et al (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:W182–W185. https://doi.org/10.1093/nar/gkm321

Ravindran PN, Kallupurackal JA (2001) Black pepper. In: Handbook of herbs and spices, pp 62–110

Schrödinger (2018) Schrödinger Release 2018-1. BioLuminate, Schrödinger, LLC, New York

Shimada T, Endo T, Fujii H et al (2014) Characterization of three linalool synthase genes from *Citrus unshiu* Marc. and analysis of linalool-mediated resistance against *Xanthomonas citri* subsp. citri and *Penicilium italicum* in citrus leaves and fruits. Plant Sci. https://doi.org/10.1016/j.plantsci.2014.09.008

Sneath PHA, Sokal RR (1973) Numerical taxonomy. The principles and practice of numerical classification. W.H. Freeman and Company, San Francisco

Sugiura M, Ito S, Saito Y et al (2011) Molecular cloning and characterization of a linalool synthase from Lemon Myrtle. Biosci Biotechnol Biochem. https://doi.org/10.1271/bbb.100922

Wang QR, Fan W (2009) Cloning, expression analysis and sequence characterization of a monoterpene synthase gene, PaLis, from apricot (Prunus armeniaca). In: 3rd international conference on bioinformatics and biomedical engineering, iCBBE 2009

Waterhouse A, Bertoni M, Bienert S et al (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296–W303. https://doi.org/10.1093/nar/gky427

Wheeler TJ, Eddy SR (2013) Nhmmer: DNA homology search with profile HMMs. Bioinformatics. https://doi.org/10.1093/bioinformatics/btt403

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Evolving genes and proteins