South African Journal of Botany 159 (2023) 617–626

Contents lists available at ScienceDirect

South African Journal of Botany

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

Oxygenated sesquiterpenes, molecular docking, and the trait-linked occurrence of essential oil in *Knema angustifolia* (Roxb.) Warb. (Myristicaceae).

Rubi Barman[1,2], Jadumoni Saikia[1,2], Prasanna Sarmah[1,2], Parthapratim Konwar[1,2], Manoj Kumar[1,2], Pranjit Kumar Bora[1,2], Prastuti Bhattacharyya[1], Siddhartha Proteem Saikia[1,2], Saikat Haldar[1,2,*], Dipanwita Banik[1,2,*]

[1] *Agrotechnology and Rural Development Division, CSIR-North East Institute of Science and Technology, Jorhat 785006, Assam, India*
[2] *Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India*

ARTICLE INFO

Article History:
Received 2 December 2022
Revised 6 June 2023
Accepted 14 June 2023
Available online xxx

Edited by: Dr S.C. Pendota

Key words:
docking
essential oil
Knema
neurodegenerative
oxygenated sesquiterpenes
trait

ABSTRACT

The folk medicinal plant *Knema angustifolia* (Roxb.) Warb. of *Myristicaceae* is reported with the essential oil (0.02±0.001 - 0.04±0.01%) in leaves and twigs respectively, for the first time. The GC-MS analysis of essential oil found 15 oxygenated sesquiterpenes viz., Globulol (30.83-35.46%), Spathulenol (21.46-23.98%), Viridiflorol (6.24-12.40%) and others. Several species under the family *Myristicaceae* had been used in neurodegenerative diseases. The site-specific docking of the markers against *tau* protein, the major causal factor for neurodegenerative diseases were carried out which found biomarkers viz., α-Cadinol, δ-Cadinol, *epi*-Cubenol, Neointermedeol, Shyobunol, and τ-Cadinol with moderate to high binding affinity (-5.1 to -6.2 Kcal/mol). The markers followed ADMET parameters including Human Intestinal Absorption, Blood Brain Barrier, CNS permeability, and Lipinski rule of five, and were found as moderately bioactive. τ-Cadinol and δ-Cadinol showed high binding affinity (-6.2 Kcal/mol) like Galantamine, an FDA-approved drug. Moreover, τ-Cadinol and δ-Cadinol showed moderate activity as nuclear receptors, enzyme inhibitors, and ion channel modulators and exhibited potential as a base structure to develop candidate oral drugs. The potential of trait-linked occurrence of essential oil among *K. angustifolia* and allied species were assessed through ancestral reconstruction using Mesquite ver. 3.61 upon Bayesian consensus tree of plastid genes viz., *psbA-trnH, rbcL,* and *matK*. The essential oil was found as a shared ancestral trait in *K. angustifolia*. For the long-term conservation of *K. angustifolia,* the MaxEnt modeling was carried out which predicted climate-suitable habitat in NE India and Western Himalayas.

1. Introduction

Plant essential oils are volatile and aromatic secondary metabolites in nature (Husnu and Demirci 2007). More than 17000 plant species of 92 family biosynthesize essential oils which play a vital role in the defense and reproduction (de Wilde, 2000; Prins et al., 2010; Kumari et al., 2014; Prakash et al., 2015). Plant essential oils are widely used in flavour and fragrance, pharmaceuticals, food and beverages, and other industries (Koul et al., 2008; Daviet and Schalk, 2010; Prakash et al., 2015). During field surveys in North East India, *Knema angustifolia* (Roxb.) Warb., a tree species of *Myristicaceae* was collected with aromatic leaves and twigs. The genus *Knema* comprises approximately 96 tree species distributed from tropics to subtropics. Initially the essential oil was reported from *Knema kunstleri* (King) Warb. Salleh et al., 2021). The commonly distributed species of *Knema* in North East India are viz., *Knema erratica* (Hook.f. & Thomson) J. Sinclair, *K. linifolia* (Roxb.) Warb., *K. lenta* Warb., *K. tenuinervia* W.J. de Wilde, *K. globularia* (Lam.) Warb. and *K. angustifolia* (Roxb.) Warb. of which a few have been used in folk medicine (Banik and Bora, 2016; Salleh and Farediah, 2017; POWO, 2021; Salleh et al., 2021). *K. angustifolia* had been used as ethnomedicinal species in Thailand and previously reported with anti-oxidant activity (Phadungkit et al., 2010). Many species of the family *Myristicaceae* contain essential oil and had been used as spices and hallucinogenic medicines to treat neurodegenerative ailments (Barman et al., 2021). Recent studies showed the potential of psychedelic substances to treat neurodegenerative disorders (Morales-Garcia et al., 2020;

* AcSIR - Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India
E-mail addresses: saikatchembiol@gmail.com (S. Haldar), dipanwitabanik@neist.res.in, banikdipanwita@yahoo.com (D. Banik).

https://doi.org/10.1016/j.sajb.2023.06.030
0254-6299/

Kozlowska et al., 2021; Saeger and Olson, 2022). Further, the neurodegenerative diseases viz., dementia, Alzheimer's disease, and Parkinson's disease exhibited disease prognosis through *tau* protein mediated neurofibrillary tangles (Goedert et al., 1989; Shin et al., 1991; Lei et al., 2010). The present study aimed to estimate the yield of essential oil of *Knema angustifolia,* identify the chemical markers, *in silico tau* binding affinity of the biomarkers, and analyze the trait-linked occurrence potential of essential oil among closely related species.

2. Materials and methods

2.1. Collection and identification of plant species

Fresh leaves and young twigs of *Knema angustifolia* were collected from adjoining locality of Gibbon wildlife sanctuary (26°43′00″N and 94°23′00″E), Holongapar, Jorhat and *Knema erratica* from Deoparbat, Golaghat (26°36′03″N and 93°43′52″E), Assam, India. The other available species, selected for the study were also collected from different parts of NE India. The voucher specimens (*D. Banik* 1851, 1852, 1853, 1856, 1857; *D. Banik, R. Barman & P. Konwar* 1311; *R. Barman & J. Saikia* 1320; *R. Barman & D. Banik* 1613; *D. Banik & J. Saikia* 0447, 0452; *D. Banik, P. Konwar & B. Das* 0876, 0877) were deposited at the Herbarium (CSIRNEIST). The general morphology of the live species were studied under a stereo-zoom binocular dissecting microscope (Magnus-TZ). The species were identified based on comparing morphological characters of the live specimens with the taxonomic literature, type and authentic specimens (de Wilde, 2000; 2002; Li and Wilson, 2008; Banik and Bora, 2016) (**MS1**; Fig. S1). Freshly collected plant materials were stored in the refrigerator at −20°C for further analysis.

2.2. Isolation of Essential oil

Hydro-distillation was performed with 200 g of each plant part viz., fresh leaves and young twigs in 250 − 300 ml of water for 6 hours through Clevenger type apparatus. Pale yellow oil lighter than water was carefully collected using *n*-hexane and the yield percentage was measured in *w/w* with respect to the fresh weight of the samples. The extracted essential oil (EO) was passed through anhydrous sodium sulfate to remove the remaining traces of water. The extraction was carried out in triplicates and the mean value was used to calculate the yield percentage of EO. The extracted EO was stored at 4°C under dark till further analysis.

2.3. GC-MS analysis

The compositional analysis of the EO was carried out in an Agilent 8890 GC system attached with an HP-5MS UI capillary column (30 m × 250 μm × 0.25 μm) and triple quadrupole mass detector (Agilent 7010B Triple Quadrupole MS). EO samples were dissolved in analytical grade ethyl acetate at ~50-100 ppm of which 0.3 μL was injected with a split ratio of 1:10 (Gogoi et al., 2021). Total runtime of 28.75 min was programmed as follows: (a) initially at 50°C for 1 min, (b) followed by increasing up to 200°C with a rate of 8°C/min and hold for 2 min, c) subsequently increasing up to 260°C with a ramp of 15°C/min and hold for 3 min. The flow rate of carrier gas (He) was maintained at 1.0 mL/min. The inlet temperature was kept at 250°C. Data were analyzed in Agilent MassHunter Qualitative Analysis 10.0 software (Gogoi et al., 2021). The relative percentage of the individual constituent was determined by the area under the peak. The library-based structural hits were obtained by comparing 70 eV EI-MS spectra in positive ion mode with NIST Mass Spectral Database (NIST MS Search v.2.3). The quality of structural prediction was assessed by the match and R match values of the individual hits. The identity of the constituents was also validated with their linear retention indices (RI) relative to C_{10}-C_{40} *n*-alkanes (Mahanta et al., 2020; Bora et al., 2021; Gogoi et al., 2021).

2.4. Molecular docking

Among the markers of the essential oil of *K. angustifolia* 12 dominant marker compounds (Area % > 1.1 as per GC-MS analysis) were selected for the site specific molecular docking with human *tau* protein. The three-dimensional crystallographic structure of the human *tau* protein (PDB ID: 2V17) was retrieved from the RCSB PDB database (https://www.rcsb.org/) and the information on the active sites of protein 2V17, namely, THR 386, ASP 387, HIS 388, GLY 389, ALA 390, and GLU 391 was retrieved from PDBsum (Sevcik et al., 2007; Laskowski et al., 2018). The protein molecule was refined by removing the undesirable water molecules, heteroatoms, and ions. The canonical SMILES format of 16 compounds was obtained from the NCBI PubChem database (Kim et al., 2021) and then converted to PDB format using Corina classic demo (2022) (https://mn-am.com/demos-services/) and Autodoc 4.2 of PMV ver. 1.5.7 was used to estimate binding affinity (Sanner, 1999). FDA-approved drugs viz., Donepezil, Galantamine, Memantine, and Rivastigmine were used as control (Alzheimer's Association, 2023). Autodock 1.5.7 (Sanner, 1999; Trott and Olson, 2010) was employed to modify the ligands to attach polar hydrogen atoms, Kollman's charges, gasteiger charges, rotatable bonds, and torsional degree of freedom to ensure optimum binding with the receptors. The grid box was constructed in the active site of the protein (PDB ID: 2V17) with grid size value (X=-25.708, Y=-28.357, Z=13.969, dimensions 50×50×50 Å2). Grid values were saved as "config.txt" and the protein and ligands in pdbqt format. Vina program was run through the windows command prompt for each protein-ligand configuration file to estimate the binding affinity.

Docking scores were autogenerated as an output file in .txt format. The highest binding affinity of the ligand docking conformation was recorded. The Discovery studio visualizer (Biovia Dassault, 2021) and PyMOL (Lilkova et al., 2015) were used to visualize ligand interaction with amino acid residues of the protein target and bond types.

ADMET and drug-likeness prediction for the pharmacokinetic profile of the drug molecule was performed viz., lipophilicity (XlogP3), log S (solubility), TPSA (polarity), size (molecular weight), insaturation (fraction Csp3), and flexibility (number of rotatable bonds) (Daina et al., 2017). pkCSM online server was used in forecasting the response of candidate drugs (Pires et al., 2015). Bioactivity score of all compounds for drug target were evaluated using online server Molinspiration Cheminformatics free web services, 2022 (https://www.molinspiration.com).

2.5. Trait tracing

Knema angustifolia and its allied species from North East India viz., *Knema erratica, K. globularia, K. tenuinervia, K. lenta, K. linifolia,* and outside NE India viz., *K. cinerea* (Poir.) Warb., *K. kunstleri* (King) Warb. were used in the study (Sinclair, 1958; 1961; de Wilde, 1979; 2000; Banik and Bora, 2016). *Myristica fragrans* Houtt. was used as an outgroup (**MS1**; Fig. S1).

The morphological matrix was prepared with 56 qualitative and quantitative characters of the 9 studied species. The quantitative characters were measured on a metric scale. For the natural range of value, the mean value was used to enumerate the character states using primary and secondary data from live species and taxonomic literature respectively where the non-availability of the data was referred to as missing trait (Siga et al., 2022). The descriptive binary and multistate conditions or traits of 56 characters of the selected 9 species i.e., altogether 228 traits were enumerated as 0 − 9 within parenthesis and considered unweighted (Puce et al., 2016; Table S1-S2).

The plant DNA barcode candidate of chloroplast regions viz., *rbcL, matK* and *psbA-trnH* were selected. The nucleotide sequences of the selected species with more than 300 base pairs were downloaded from NCBI and the GenBank accession numbers were provided (Table S3). As the sequences of *K. angustifolia, K. erratica* and *K.*

kunstleri were not available in GenBank, genomic DNA was isolated from freshly collected young leaves of *K. angustifolia* and *K. erratica* using the DNeasy Plant Mini Kit (QIAGEN) following manufacturer's protocol and *K. kunstleri* was treated as missing data. The three barcode loci *rbcL, matK* and *psbA-trnH* were amplified in PCR (**MS2**). The amplified PCR products were sequenced bidirectionally at M/S Agrigenome Pvt. Ltd, India. The amplified sequences were edited manually using CLC Genomic workbench 22.0.1 and BLAST analysis was executed to check the quality of the sequences (Madden, 2002; **MS2**).

Both the downloaded and amplified sequences of similar loci were assembled individually and trimmed manually in MEGA 10.0 (Kumar et al., 2018). The assembled sequences were then aligned with MUSCLE and the inter and intra-generic distances for each barcode locus were calculated using pairwise distance matrix in MEGA 10.0. The assessment of the barcode loci was done using PAUP ver. 4.0a169 and Mega 10.0 (Swofford, 1998). (**MS3**; Table S4).

Bayesian interference (BI) tree was reconstructed in MrBayes v3.2.7 using each single and multi-loci barcode to get the tree topology (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). The JmodelTest v2.1.9 was used to assess the best-fit substitution model as per Bayesian Information Criterion (BIC) which indicated F81 model for *matK* and *psbA-trnH* and K80 model for *rbcL* (Guindon and Gascuel, 2003; Darriba et al., 2012). BI of the combination of three loci was assessed by running Markov Chain Monte Carlo (MCMC) with four replicates for 1M generations and sampling individual trees after 1000th generations with one cold and three hot chains. A 50% majority-rule consensus BI tree of three combined loci was reconstructed with a posterior probability value as branch support by discarding the initial 25% trees (Ruchisansakun et al., 2015; Puce et al., 2016; Siga et al., 2022; **MS4**). FigTree ver (2018) v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) was used to visualize all the possible tree topologies.

Parsimonious ancestral traits of all 56 characters including essential oil and morphology were reconstructed by tracing trait distribution onto the best topology of the 50% majority-rule consensus BI tree of combined loci (*rbcL* + *matK* + *psbA-trnH*) using Mesquite version 3.61 (build 927) to visualize shared, derived, and independently originated traits where the tree topology was not considered as true phylogeny of the genus *Knema* (Knapp, 2002; Maddison and Maddison, 2019).

2.6. Predictive Habitat Modelling

Nearly 22 occurrence records of *K. angustifolia* from different localities of NE India were compiled from systematic field survey and herbarium records by converting into the decimal format of geo-coordinates using Google Earth Pro (v7.3.4.) and saved as Microsoft excel CSV file compatible with MaxEnt model. Nearly 19 bioclimatic variables from WorldClim along with elevation were used at a spatial resolution of 30 arc-second (~1 km) in the model (Hijmans et al., 2005; Kumar et al., 2021). The parameters used in MaxEnt (ver. 3.4.4e) were viz., iterations 5000, convergence threshold 0.00001, prevalence 0.5, background point 10000 with 10 replicates. The model was executed at 25% training threshold (Pearson et al., 2004; Kumar et al., 2021). Model reliability was assessed using bootstrap and performance was evaluated based on the AUC value by means of ROC. AUC value indicated prediction e.g., very good (0.95<AUC<1.0), good (0.9<AUC<0.95) fair (0.8<AUC<0.9) and poor (AUC<0.8) (Swets, 1988; Thuiller et al., 2005; Phillips et al., 2006; Wang et al., 2007; Phillips and Dudik, 2008; Phillips and Elith, 2010; Kumar et al., 2021).

3. Results

3.1. Essential oil analysis

Hydro-distillation of leaves and young twigs of *K. angustifolia* yielded 0.02±0.001% and 0.04±0.01% volatile oil respectively with pungent-woody smell whereas the yield of essential oil from leaves of *Knema kunstleri* was reported as 0.12% in Malaysia (Salleh et al., 2021). The GC-MS analysis of both the leaf and twig essential oils revealed an exclusive richness of oxygenated sesquiterpenes. Nearly 15 constituents were identified comprising of 90.67% and 99.92% of the oil of leaf and twigs respectively. The identified constituents, their retention time (R_t), relative %, match and R match values, RI values (exp. and lit.) and structural class have been represented in tables (Table 1; Table S5; Fig. S2). Among a series of identified oxygenated sesquiterpenes (OS), Spathulenol (**4**) and Globulol (**5**) were the major ones, contributing 59.44% and 52.29% in the leaf and twig oils respectively. The identity of these major constituents was confirmed on the basis of library hits with high match and R match values (>900) and their RI values. These two sesquiterpene alcohols are closely structurally related to the identical tricyclic skeleton. Another two sesquiterpene alcohols with the same tricyclic skeleton namely Viridiflorol (**6**) and Ledol (**7**) were also identified in both the oil samples with relatively minor abundance (4.50-12.40%). A bunch of bicyclic sesquiterpene alcohols *viz.* τ, δ and α-Cadinol (**10, 11, 12**), *epi*-Cubenol (**9**), Maaliol (**3**) and Neointermedeol (**13**) were the next dominating class of oil constituents contributing in total 10.61% and 25.44% of the leaf

Table 1
GC-MS analysis of essential oil of *Knema angustifolia* (Retention time, percentage area and library-based identification)

Compounds	RT (min)	Area %		match	R match	RI_{exp}	RI_{lit}	Class
		Leaves	Twig					
1. Cubebol	16.09	0.51	0.70	797	808	1517	1515	OS
2. *epi*-Globulol	16.76	0.94	0.59	876	931	1557	1554	OS
3. Maaliol	16.89	1.04	1.54	848	883	1576	1574	OS
4. Spathulenol	17.03	23.98	21.46	940	942	1579	1576	OS
5. Globulol	17.13	35.46	30.83	906	931	1586	1583	OS
6. Viridiflorol	17.26	6.24	12.40	886	918	1592	1591	OS
7. Ledol	17.42	4.56	7.24	885	937	1605	1602	OS
8. Humulene 1,2-epoxide	17.51	2.59	1.19	872	906	1610	1606	OS
9. *epi*-Cubenol	17.74	4.21	5.10	818	888	1630	1627	OS
10. τ-Cadinol	17.93	1.20	4.09	882	892	1645	1640	OS
11. δ-Cadinol	17.98	0.63	1.76	836	934	1647	1645	OS
12. α-Cadinol	18.10	0.34	8.41	892	917	1654	1653	OS
13. Neointermedeol	18.15	2.25	3.95	866	899	1663	1660	OS
14. Mustakone	18.47	4.59	0.32	842	904	1690	1687	OS
15. Shyobunol	18.54	2.13	0.34	787	806	1692	1687	OS
Oxygeneted Sesquiterpene		90.67	99.92					
Total % of compounds identified		**90.67**	**99.92**					
Unidentified		**9.33**	**0.08**					

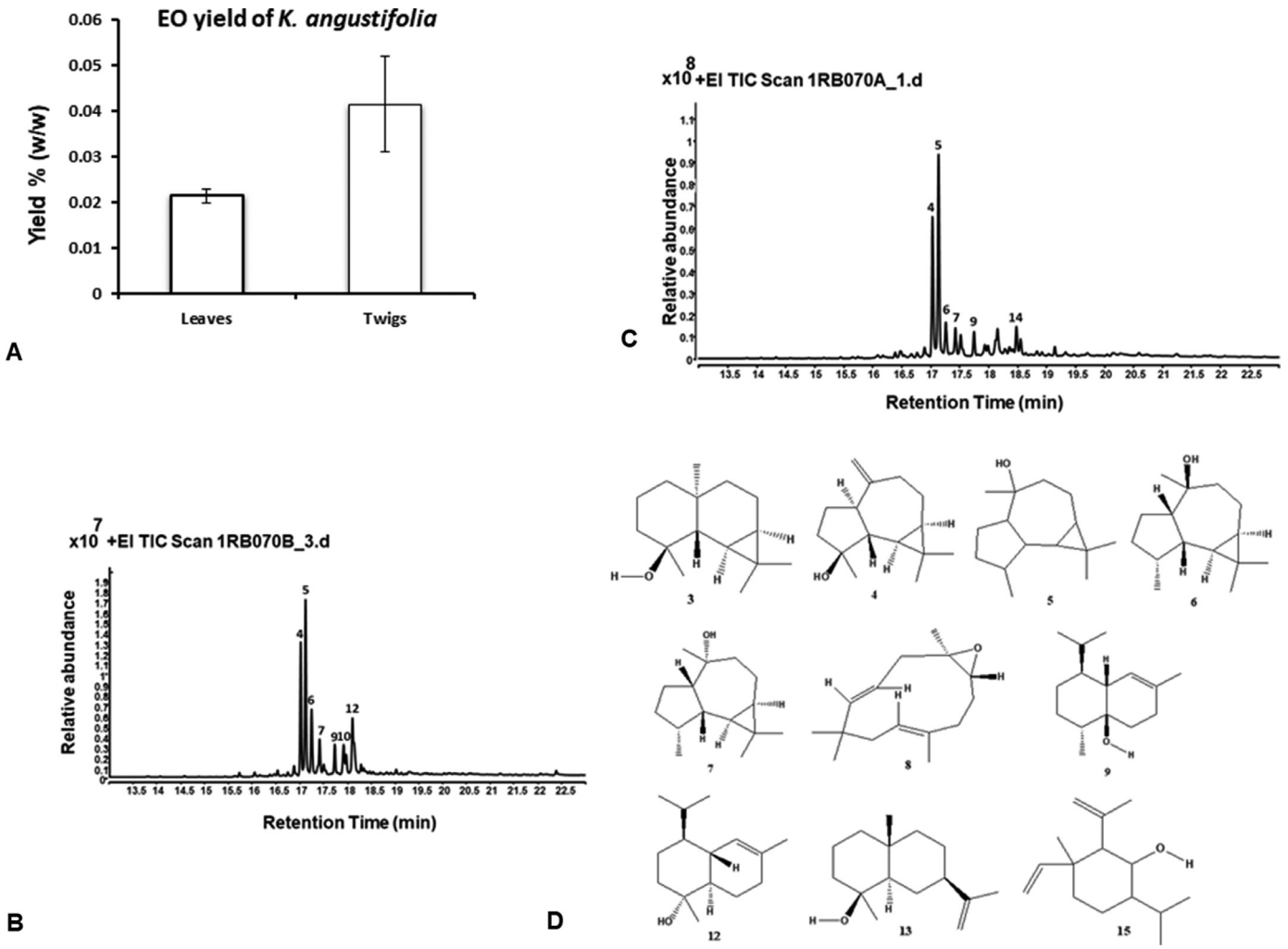

Fig. 1. Essential oil in *Knema angustifolia*: A. Yield from leaves and twigs, B-C. GC-MS chromatograms; B. leaves, C. twigs, D. marker compounds detected by GC-MS (3 = Maaliol, 4 = Spathulenol, 5 = Globulol, 6 = Viridiflorol, 7 = Ledol, 8 = Humulene 1,2-epoxide, 9 = *epi*-Cubenol, 10 = τ-Cadinol, 12 = α-Cadinol, 13 = Neointermedeol, 14 = Mustakone, 15 = Shyobunol).

and twig samples respectively. Mustakone (1**4**), an α,β-unsaturated sesquiterpene ketone with a complex bridged tricyclic skeleton was also detected with 4.59 and 0.32% respectively. A few minor (<3.0%) oxygenated sesquiterpenes were found viz., Cubebol (**1**), Humulene 1,2-epoxide (**8**) and Shyobunol (**15**). A previous study on essential oil of *K. kunstleri* also reported a high abundance of sesquiterpenes viz., Sesquiterpene Hydrocarbons (SH) 77.3% and Oxygenated Sesquiterpene (OS) 14.4% comprising majorly of β-Caryophyllene (23.2%), Bicyclogermacrene (9.6%), δ-Cadinene (7.3%), α-Humulene (5.7%) and Germacrene D (4%). Common compounds detected in both *Knema kunstleri* and *Knema angustifolia* were viz., Globulol, Spathulenol, Humulene, *epi*-Cubenol, α-Cadinol (Salleh et al., 2021). The study found only 26.67% similarity of the chemical markers of essential oil of *Knema angustifolia* with *Knema kunstleri* (Table S6). However, the essential oil from *K. angustifolia* was unique in terms of its richness and diversity in oxygenated sesquiterpenes (Fig. 1).

3.2. Evaluating docking results

The binding affinities of 12 marker compounds and 4 standard drugs of neurodegenerative disease against the target protein human *tau* (2V17) was calculated using site specific docking experiment (Table S7; Fig. 2). Among the markers, δ-Cadinol and τ-Cadinol showed the highest docking score of -6.2 Kcal/mol followed by α-Cadinol, Globulol with -6.1 Kcal/mol, Spathulenol, Ledol, Viridiflorol with -6.0 Kcal/mol and Mustakone with -5.9 Kcal/mol. Whereas, the binding affinity of the standard drugs was viz., Donepezil (-6.3 Kcal/mol), Galantamine (-6.2 Kcal/mol), Memantine (-5.1 Kcal/mol), and Rivastigmine (-5.1 Kcal/mol). The molecular interaction of ligands with protein targets was equilibrated by the hydrogen and hydrophobic bonds. Both conventional hydrogen and hydrophobic interactions were observed in the markers and control drugs for bond types and amino acid residues (Fig. 3; S3 –S4). Both τ-Cadinol and δ-Cadinol showed hydrogen and hydrophobic bonds with amino acid residue ALA 44 (4.33 Å), PRO 95 (4.79 Å), PHE 98 (3.24 Å), LEU45 (2.47 Å). Similarly, α-Cadinol, Globulol showed hydrogen and alkyl bond with PRO 95 (4.86 Å and 5.17 Å), PHE 98 (2.03 Å and 2.69 Å), ALA 44 (4.27 Å and 3.63 Å) and ILE 4 (4.82 Å). Spathulenol and Ledol showed hydrogen, alkyl, and donor-donor interaction with amino acid residues including PHE 98 (2.41 Å, 2.71 Å, and 2.69 Å), ALA 44 (4.85 Å and 3.64 Å) and ILE 4 (4.85 Å). Virdiflorol showed alkyl interaction with amino acid residues PRO 95 (5.18 Å), ALA 44 (3.63 Å), and ILE 4 (4.85 Å). Mustakone showed alkyl bond interaction with ALA 44 (3.60 and 4.44 Å). Furthermore, standard drugs Donepezil showed carbon hydrogen and alkyl bond with amino acid residues including ALA 44 (5.09 Å), LYS 43 (4.92 Å), PRO 95 (4.79 Å), SER 63 (3.38 Å), VAL 2 (3.46 Å). Galantamine showed conventional hydrogen, carbon hydrogen, alkyl and pi alkyl interaction with target amino acid residues containing GLU 46 (2.94 Å), ILE 4 (3.59 Å), ALA 44 (3.69 Å, 3.84 Å, 4.08 Å), LYS 43 (4.24 Å). Memantine exhibited conventional

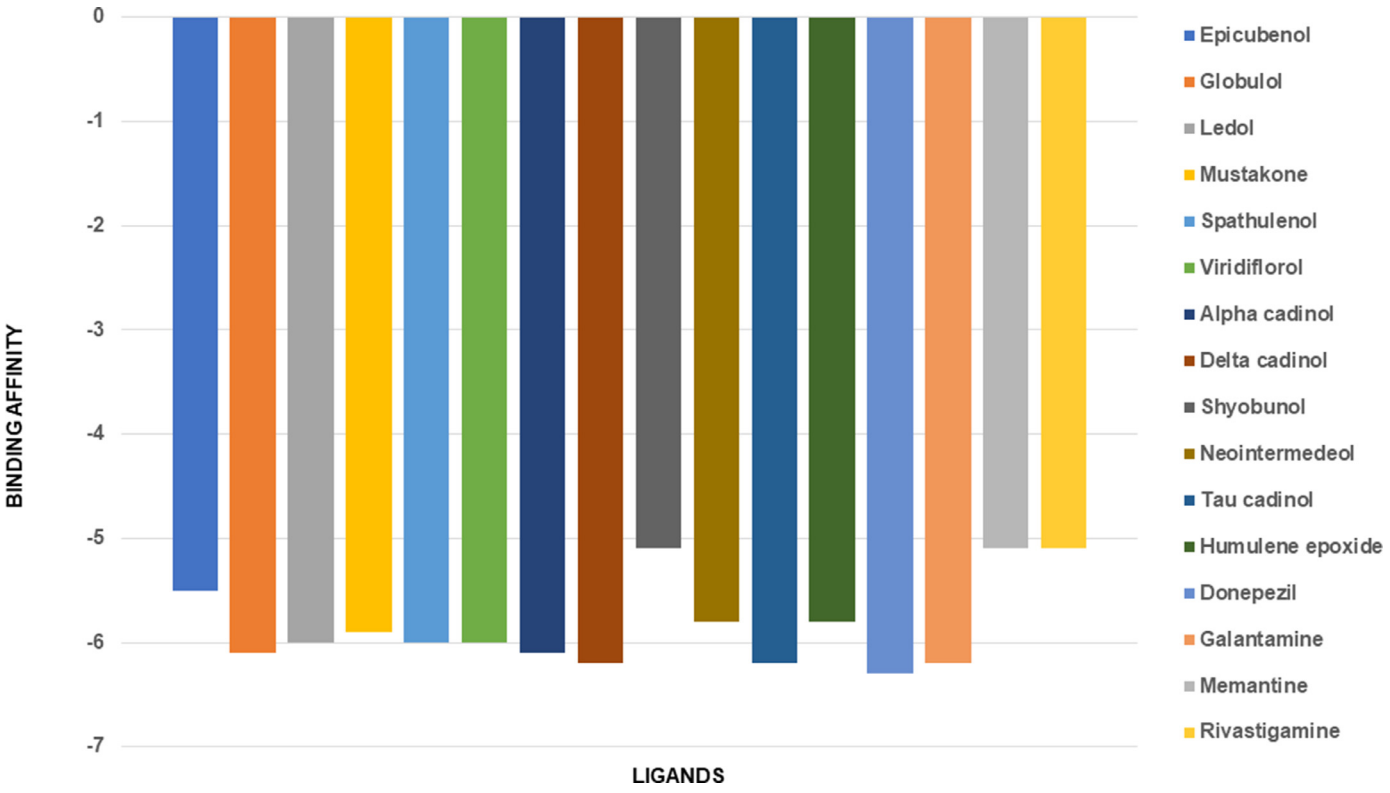

Fig. 2. Clustered columns showing the binding affinity value (Kcal/mol) of 16 ligand compounds with target protein human tau (PDB ID: 2V17)

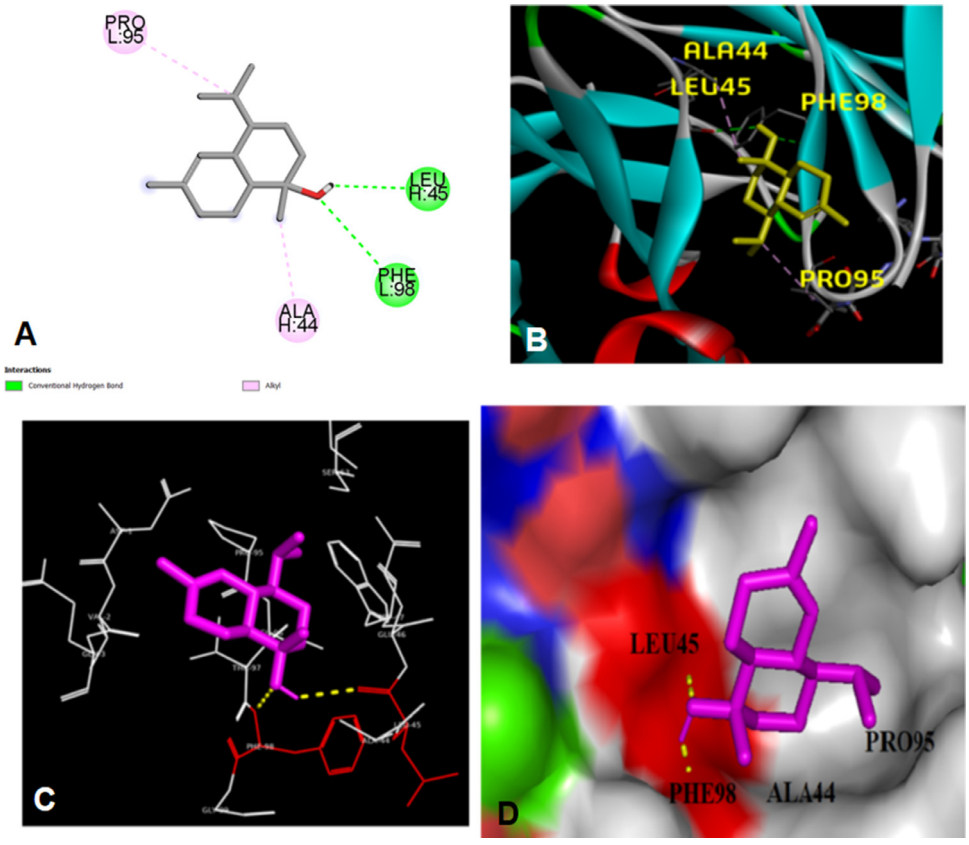

Fig. 3. A-B. 2-D and 3-D interaction of target protein and δ-Cadinol visualized in Discovery studio visualizer, C-D. Two Hydrogen bond (Red colour) and two Hydrophobic bond (Gray colour) were observed in PyMOL.

hydrogen and alkyl bond with PRO 95 (4.67 Å) and GLU 46 (3.72 Å). Lastly, Rivastigmine showed conventional, carbon hydrogen, pi anion, pi sigma and pi alkyl interaction with amino acid residues PHE 98 (2.97 Å and 3.58 Å), GLU 46 (3.72 Å and 4.64 Å) THR 97 (3.88 Å), PRO 95 (5.46 Å) and ASP 1 (3.39 Å and 3.49 Å).

3.3. ADMET, Drug-Likeness and Bioactivity score

ADMET properties of ligands were determined using the pkCSM online server. The ADMET results revealed higher human intestinal absorption (HIA) value of Mustakone (98.25%), and Humulene epoxide (95.52%) suggesting more effective absorption by the intestinal tract upon oral administration than all four standard drugs (88.46 - 94.99%) while *epi*-Cubenol showed similar absorption value like Galantamine (94.99%). An early blood-brain barrier (BBB) disruption had been established in neurodegenerative diseases (Montagne et al., 2017). ADMET predicted the penetrance of a compound through the BBB reflecting the transportation of molecules and ions between the blood and brain. A value of log BB > 0.3 ensured the penetrance of marker compounds through the BBB (Vardhan and Sahoo, 2020). Mustakone showed best BBB penetrance (log BB 0.74) followed by Humulene epoxide (log BB 0.67), Globulol, Ledol, and Viridiflorol (log BB 0.63) compared to all the experimental standard drugs (log BB 0.08 - log BB 0.60). P-glycoprotein (PGP) is responsible for the absorption and disposal of drugs through efflux while weak substrates of PGP may experience lesser extrusion with optimum delivery (Constantinides and Wasan, 2007; Amin, 2013). All the marker compounds except standard drug, Donepezil were not reported as a substrate of PGP suggesting that the efficacy of the rest of the experimental markers and drugs may not be affected by the overexpression of PGP. Further, markers with a log PS > -2 have been considered to penetrate the CNS and were referred as CNS active showing possibility of being fatal and toxic while markers with log PS > -3 were unable to cross CNS and were termed as CNS inactive (Fatima, et al. 2020). Most of the tested markers exhibited as CNS inactive or CNS log PS > -3 except Mustakone which was found CNS active like Donepezil. Oral rat chronic toxicity was recorded in Globulol, Ledol and Viridiflorol (1.19 log mg/kg bw/day) and Humulene epoxide (1.13 log mg/kg bw/day) compared to the standard drugs (0.97 - 1.25 log mg/kg bw/day) while the rest markers showed lesser toxicity than the standard drugs. Minnow toxicity for all the marker compounds were assessed which showed higher log LC_{50} values than -0.3 (Vardhan and Sahoo, 2020) reflecting that most of the markers used in the study were not acutely toxic while Donepezil showed -2.01 log LC_{50} for acute toxicity in Minnow fishes. The markers along with the control drugs were assessed for hepatotoxicity and all the compounds revealed no toxicity except Donepezil and Galantamine (Table 2).

Drug-likeness prediction found that among the experimental 12 markers, only 6 viz., α-Cadinol, δ-Cadinol, τ-Cadinol, *epi*-Cubenol, Neointermedeol and Shyobunol showed drug-likeness properties for oral drug (Table 3; Fig. S5). All the 6 markers exhibited lower molecular weight (222.37 g/mol) while the lowest was found in standard drug, Memantine (179.30 g/mol) reflecting better absorption. However, all the compounds followed lipophilicity while Galantamine showed the lowest inferring lesser toxic than the rest. Mustakone and Humulene epoxide violated the TPSA range (20 - 130 Å2) inferring that except these two, the rest are polar in nature. Lower log S value of Shyobunol exhibited solubility (log S -4.99) (Nisha et al., 2016). Insaturation of the compounds was tested based on carbon fraction in sp3 hybridization which showed positive results among 50% of the experimental markers inferring increase in saturation, molecular solubility or better binding pockets of larger ligands for clinical testing (Lovering et al., 2009; Wei et al., 2020). Shyobunol exhibited flexibility with 3 rotatable bonds reflecting molecular flexibility which may enhance oral bioavailability (Khanna and Ranganathan, 2009).

Table 2
ADMET profile of markers of essential oil of ***K. angustifolia*** and control drugs assessed through pkCSM

Biomarkers/ Standard drugs	Water solubility Numeric (log mol/L)	Caco2 permeability Numeric (log Papp in 10-6 cm/s)	Intestinal absorption (human) Numeric (% Absorbed)	P-glycoprotein substrate Categorical (Yes/No)	P-glycoprotein I Inhibitor Categorical (Yes/No)	P-glycoprotein II inhibitor Categorical (Yes/No)	VDss (human) Numeric (log L/kg	BBB permeability Numeric (log BB)	CNS permeability Numeric (log PS)	Oral Rat Acute Toxicity (LD50) Numeric (mol/kg)	Oral Rat Chronic Toxicity (LOAEL) Numeric (log mg/kgbw/day)	Minnow toxicity Numeric (log mM	Hepatotoxicity
α-Cadinol	-4.07	1.48	94.30	NO	NO	NO	0.42	0.60	-2.15	1.92	1.47	0.74	NO
δ-Cadinol	-4.07	1.48	94.30	NO	NO	NO	0.42	0.60	-2.15	1.92	1.47	0.74	NO
epi- Cubenol	-4.21	1.64	94.99	NO	NO	NO	0.46	0.60	-2.16	2.00	1.31	0.65	NO
Globulol	-4.31	1.48	92.81	NO	NO	NO	0.56	0.63	-2.18	1.62	1.19	1.06	NO
Humulene epoxide	-4.05	1.42	95.52	NO	NO	NO	0.44	0.67	-2.92	1.62	1.13	1.19	NO
Ledol	-4.31	1.48	92.81	NO	NO	NO	0.56	0.63	-2.18	1.62	1.19	1.06	NO
Mustakone	-4.16	0.99	98.25	NO	NO	NO	0.36	0.74	-1.81	1.71	1.21	0.75	NO
Neointermedeol	-4.39	1.51	93.70	NO	NO	NO	0.49	0.60	-2.33	1.72	1.28	0.73	NO
Shyobunol	-4.13	1.32	94.65	NO	NO	NO	0.31	0.60	-2.04	1.69	1.64	0.45	NO
Spathulenol	-3.87	1.39	93.23	NO	NO	NO	0.52	0.60	-2.45	1.69	1.39	1.27	NO
τ-Cadinol	-4.07	1.48	94.30	NO	NO	NO	0.42	0.60	-2.15	0.92	1.47	0.74	NO
Viridiflorol	-4.31	1.48	92.81	NO	NO	NO	0.56	0.63	-2.18	1.62	1.19	1.06	NO
Donepezil	-4.65	1.27	93.71	YES	YES	YES	1.27	0.16	-1.46	2.75	0.99	-2.01	YES
Galantamine	-2.64	1.59	94.99	NO	NO	NO	0.89	-0.08	-2.51	2.73	0.97	1.67	YES
Memantine	-2.32	1.33	91.23	NO	NO	NO	0.99	0.60	-2.48	2.67	1.25	1.43	NO
Rivastigmine	-2.35	1.57	88.46	NO	NO	NO	0.62	0.51	-2.25	3.40	1.16	1.37	NO

Table 3
Drug-likeness prediction of markers of ***Knema angustifolia*** and control drugs of neurodegenerative diseases assessed through Swiss ADME

Compound	Cannonical SMILES	XLogP3	MW	TPSA	Log S	Saturation	Flexibility
α-Cadinol	CC1=CC2C(CCC(C2CC1)(C)O)C(C)C	3.34	222.37	20.23 Å2	-3.44	0.87	1
δ-Cadinol	CC1=CC2C(CCC(C2CC1)(C)O)C(C)C	3.34	222.37	20.23 Å2	-3.44	0.87	1
epi-Cubenol	CC1CCC(C2C1(CCC(=C2)C)O)C(C)C	3.70	222.37	20.23 Å2	-3.82	0.87	1
Globulol	CC1CCC2C1C3C(C3(C)C)CCC2(C)O	3.74	222.37	20.23 Å2	-3.86	1.00	0
Humulene epoxide	CC1=CCCC2(C(O2)CC(C=CC1)(C)C)C	3.91	220.35	12.53 Å2	-3.87	0.73	0
Ledol	CC1CCC2C1C3C(C3(C)C)CCC2(C)O	3.74	222.37	20.23 Å2	-3.86	1.00	0
Mustakone	CC1=CC(=O)C2C3C1C2(CCC3C(C)C)C	3.41	218.33	17.07 Å2	-3.45	0.80	1
Neointermedeol	CC(=C)C1CCC2(CCCC(C2C1)(C)O)C	4.47	222.37	20.23 Å2	-4.61	0.87	1
Shyobunol	CC(C)C1CCC(C(C1O)C(=C)C)(C)C=C	4.83	222.37	20.23 Å2	-4.99	0.73	3
Spathulenol	CC1(C2C1C3C(CCC3(C)O)C(=C)CC2)C	3.11	220.35	20.23 Å2	-3.20	0.87	0
τ-Cadinol	CC1=CC2C(CCC(C2CC1)(C)O)C(C)C	3.34	222.37	20.23 Å2	-3.44	0.87	1
Viridiflorol	CC1CCC2C1C3C(C3(C)C)CCC2(C)O	3.74	222.37	20.23 Å2	-3.86	1.00	0
Donepezil	COC1=C(C=C2C(=C1)CC(C2=O)CC3CCN(CC3)CC4=CC=CC=C4)OC	4.28	379.49	38.77 Å2	-4.81	0.46	6
Galantamine	CN1CCC23C=CC(CC2OC4=C(C=CC(=C34)C1)OC)O	1.84	287.35	41.93 Å2	-2.34	0.53	1
Memantine	CC12CC3CC(C1)(CC(C3)(C2)N)C	3.28	179.30	26.02 Å2	-3.50	1.00	0
Rivastigmine	CCN(C)C(=O)OC1=CC=CC(=C1)C(C)N(C)C	2.29	250.34	32.78 Å2	-2.62	0.50	6

XLOGP3 −0.7 to +5.0, size: **MW** 150 - 500 g/mol, polarity: **TPSA** 20 - 130Å2, solubility: **log S** not higher than - 6, saturation: fraction of carbons in the sp3 hybridization 0.25 - 1, and flexibility: 0-9 rotatable bonds

Bioactivity scores of all the 12 markers were evaluated where 8 markers viz., τ-Cadinol, Spathulenol, Shyobunol, Neointermedeol, Humulene epoxide, *epi*-cubenol, δ-Cadinol and α-Cadinol scored (0.12-0.81) for at least 1 biological activity which inferred most likely to display moderate to good biological activity as values between 0.00 to 0.50 were considered as moderately active, above 0.50 were considered as highly active and values below 0.00 as inactive (Abdelrheem et al., 2021; Table 4). Humulene epoxide scored 0.81 for nuclear receptor ligand and 0.69 for enzyme inhibitor inferring highly bioactive. Neointermedeol scored 0.34 inferring moderately active as ion channel modulator, 0.59 indicating highly active as nuclear receptor ligand and 0.34 as moderately active as enzyme inhibitor. Spathulenol scored 0.28 as nuclear receptor ligand. Among the standard drugs, Galantamine and Donepezil scored 0.93 and 0.22 for GPCR ligand respectively. Galantamine also scored 1.02 for enzyme inhibitor, 0.26 for ion channel modulator and 0.20 for nuclear receptor ligand inferring highly to moderately active.

3.4. Trait-linked occurrence potential of essential oil

The outsourced sequences of *Knema angustifolia* and *Knema erratica* for *rbcL* and *psbA-trnH* were edited with CLC Genomics Workbench 22.0.1 (QIAGEN) and converted to consensus sequences. Identity of the consensus sequences was confirmed by BLASTn analysis by similarity 91.67-99.47% and query coverage 86-95% except for *psbA-trnH* locus of *K. angustifolia* where it was 78% (**MS4**; Table S8). However, sequencing was not successful for locus *matK*.

The detailed statistics of the three assessed loci using PAUP ver. 4.0a169 showed the highest alignment length 750 bp and conserved sites 745 bp in *matK*, maximum variable sites 25bp and parsimony informative sites 10bp in *psbA-trnH*. The consistency index was found as 0.9 for *psbA-trnH*, shortest tree length (5) in *matK*, and highest intergeneric (0.06±0.02) and interspecific divergence (0.03±0.03) in *psbA-trnH* (Table S4). Similar tree topology was observed in the individual dataset of each locus. The partition homogeneity test was also in congruence among the datasets (p = 0.98) and the combined dataset of plastid genes was used to reconstruct the Bayesian Inference (BI) tree for trait tracing.

The 50% majority-rule consensus BI tree exhibited monophyletic clade of the genus *Knema* with *M. fragrans* as distinct outgroup in congruence to previously published majority-rule consensus tree of *Myristicaceae* (Sauquet et al., 2003; Doyle et al., 2004; Doyle and Endress, 2010). The *Knema* clade comprised of a well-supported subclade of 5 species with 97% posterior probability (pp) nesting *K. kunstleri* + *K. cinerea* with 76% pp and *K. tenuinervia* + *K. linifolia* with 98% pp respectively as the most closely related taxa. *K. globularia, K angustifolia,* and *K. lenta* showed well-supported paraphyly within the *Knema* clade (Fig. S6; MS5).

Traced traits onto tree topology exhibited majorly homoplasy (67.85%), followed by synapomorphy (12.50%), plesiomorphy (10.71%), and symplesiomorphy (8.92%) (data not shown). However, both the essential oil-containing species *K. angustifolia* and *K. kunstleri* showed 30.34% traits common with homoplasy (16.07%), synapomorphy (10.71%), 1.78% each of plesiomorphy and

Table 4
Prediction of bioactivity score of tested compounds

Compound	GPCR ligand	Ion channel modulator	Kinase inhibitor	Nuclear receptor ligand	Protease inhibitor	Enzyme inhibitor
α-Cadinol	-0.09	0.05	-0.87	0.39	-0.63	0.40
δ-Cadinol	-0.09	0.05	-0.87	0.39	-0.63	0.40
Globulol	-0.50	-0.29	-0.82	-0.22	-0.48	-0.13
***epi*-Cubenol**	-0.17	0.41	-0.69	0.11	-0.48	0.27
Humulene epoxide	0.14	-0.20	-1.14	0.81	-0.29	0.69
Neointermedeol	-0.12	0.37	-1.01	0.59	-0.41	0.34
Ledol	-0.50	-0.29	-0.82	-0.22	-0.48	-0.13
Shyobunol	-0.32	0.00	-1.02	0.68	-0.22	0.35
Spathulenol	-0.42	-0.28	-0.68	0.28	-0.36	0.06
τ-Cadinol	-0.09	0.05	-0.87	0.39	-0.63	0.40
Viridiflorol	-0.50	-0.29	-0.82	-0.22	-0.48	-0.13
Donepezil	0.22	-0.14	-0.16	0.03	0.03	0.25
Galantamine	0.93	0.26	-0.15	0.20	0.01	1.02
Memantine	-0.28	0.12	-1.09	-1.02	-0.60	-0.47
Rivastigmine	-0.05	0.06	-0.38	-0.35	-0.18	-0.01

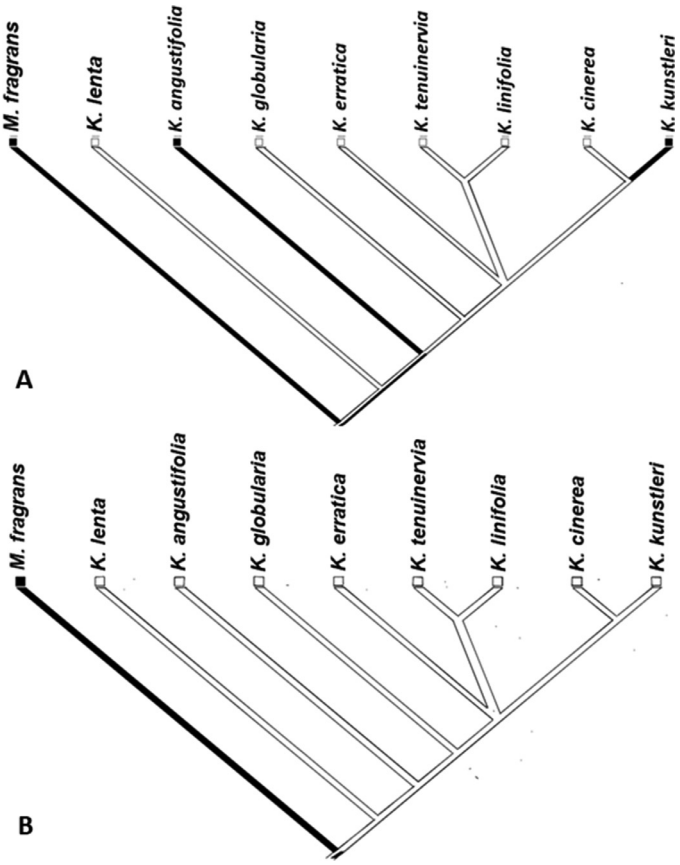

Fig. 4. Trait analysis of *Knema lenta, K. angustifolia, K. globularia, K. erratica, K. cinerea, K. kunstleri, K. tenuinervia, K. linifolia* and *Myristica fragrans* (Outgroup) by Mesquite 3.61 onto 50% majority rule consensus tree reconstructed by MrBayes. A: Essential oil, black - present, white - absent, B: aril, black - lacerated upto base, white - only at apex.

symplesiomorphy (data not shown). The presence of EO was found as a shared ancestral trait whereas, several synapomorphies viz., dioecious trees, apically lacerate aril, brown fruit, and anthers attached to disc were distributed within the clade *Knema* (Fig. 4). The current study found that the occurrence of essential oil is linked to terete young twigs and glabrous lamina upper surface among the studied species whereas, EO is linked to puberulous to glabrous young twigs, puberulous to glabrous mature fruit indumentum having hairs 0.15 - 0.3 mm; stellate or otherwise within the genus *Knema*.

3.5. Climate suitable habitat for Knema angustifolia

The suitable habitat was measured and accuracy was signified by the Area under Curve (AUC) score of training data 0.992 (±0.003) and test data 0.99 (±0.002) (Fielding and Bell, 1997; Elith et al., 2011). The model performance test showed an ideal result for ROC (mean AUC 0.992). Whereas, AUC partial/AUC random showed significantly higher values than predicted, reflecting good consistency of the model (Adhikari et al., 2018; Fig. S7). Among the bioclimatic variables, precipitation of warmest quarter (BIO18) was found as the most influencing bioclimatic variable for the geographic distribution of *Knema angustifolia* (Fig. S7; Table S9). Predictive modelling found suitable climatic habitat of the species in North East India, North Eastern part of West Bengal, Western part of Jammu and Kashmir, and parts of Himachal Pradesh for long-term conservation and re-establishment (Fig. S8).

4. Discussion

Among the commonly distributed species of *Knema* in North East India, essential oil was detected only from *Knema angustifolia* during the current study by the authors which was in congruence with the previously reported essential oil from leaves in *Knema kunstleri*. The essential oil of leaves of *Knema kunstleri* exhibited the dominance of sesquiterpenes (91.70%) i.e., Sesquiterpene Hydrocarbon (77.30%) and Oxygenated Sesquiterpene (14.40%) (Salleh et al., 2021) whereas the leaf essential oil of *Knema angustifolia* found oxygenated sesquiterpenes (90.67-99.92%). The oxygenated sesquiterpenes detected in the essential oil of *Knema angustifolia* were viz., Globulol (30.83-35.46%), Spathulenol (21.46-23.98%), Viridiflorol (6.24-12.40%), Ledol (4.56%), Mustakone (4.59%) and others. The dominant markers viz., Globulol has high commercial potential as natural food preservative (Boukhatem et al., 2013), Spathulenol as adjuvant in chemotherapy (Martins et al., 2010), Viridiflorol in flavour and fragrans (Ntana et al., 2020) as shown in patents (PCT/US2020/050354) and others.

Site specific molecular docking found that τ-Cadinol and δ-Cadinol exhibited nearly equal binding affinity against human *tau* protein (PDB ID 2V17) in comparison to Galantamine. ADMET analysis revealed that all the tested compounds exhibited the better human intestinal absorption (HIA). Most of the tested markers revealed better BBB penetrance compared to Donepezil, Galantamine, and Rivastigmine. Most of the markers exhibited CNS inactive except Mustakone indicating most of the markers may not be fatal and toxic. α-Cadinol and δ-Cadinol showed lower oral rat chronic toxicity (LOEFL) than the tested drugs. The ecotoxicological assessment by Minnow toxicity revealed lesser toxicity of the markers compared to Donepezil. None of the markers showed hepatotoxicity. For drug-likeness prediction all the markers exhibited lesser molecular weight than standard drugs reflecting better absorption of the markers than the standard drugs. All the markers except Neointermedeol and Shyobunol followed lesser lipophilicity and toxicity than Donepezil. However, only 6 markers viz., α-Cadinol, δ-Cadinol, τ-Cadinol, *epi*-Cubenol, Neointermedeol, Shyobunol followed Lipinski rule of five for drug-likeness properties. Though nearly 8 markers viz., τ-Cadinol, Spathulenol, Shyobunol, Neointermedeol, Humulene epoxide, *epi*-Cubenol, δ-Cadinol and α-Cadinol scored (0.12-0.81) for at least 1 biological activity, altogether the 6 markers exhibited site specific docking, ADMET analysis, followed Lipinski rule of five and scored (0.05 - 0.68) for at least 1 biological activity as enzyme inhibitor, ion channel modulator and nuclear receptor ligand. Neointermedeol scored 0.34 - 0.59 as moderate to high biological activity as enzyme inhibitor (0.34), ion channel modulator (0.37) and nuclear receptor ligand (0.59). Among the rest, Shyobunol (0.68; 0.35), δ-Cadinol (0.39; 0.40), and τ-Cadinol (0.39; 0.40) displayed high to moderate activity as nuclear receptor ligand and enzyme inhibitor respectively.

The phylogenetic multi-loci framework of the genus *Knema* was reconstructed in the current study for the first time. Traced traits exhibited abundance of homoplasy reflecting possibly independent evolution (Hall, 2007). The morphological traits like puberulous to glabrous and terete young twigs, glabrous lamina upper surface, puberulous to glabrous mature fruit with 0.15 - 0.3 mm hairs, stellate or otherwise showed positive correlation with presence of essential oil. Thus, the species having such morphological features like *Knema globularia, Knema cinerea,* and *Knema lenta* require thorough investigation as additional or novel source of essential oil under the genus *Knema.*

For future conservation of the species, the precipitation of warmest quarter, precipitation seasonality, min temperature of coldest month and elevation were found as significant factors in the predictive modelling showing the suitable geographical region from North East India to western Himalayas. Further, a previous study also showed that precipitation of the warmest quarter, precipitation seasonality, minimum temperature in the coldest month were inversely related to the yield of essential oil whereas, the solar radiation is directly related to the increased yield of EO in *Lippia thymoides* Mart. & Schauer (Silva et al., 2018).

5. Conclusion

The study reported presence of essential oil in *K. angustifolia* for the first time with comparatively higher yield in twigs (0.04± 0.01%) than the leaves (0.02±0.001). The GC-MS analysis showed abundance of 15 oxygenated sesquiterpenes viz., Globulol (30.83-35.46%), Spathulenol (21.46-23.98%), Viridiflorol (6.24-12.40%), Ledol (4.56%), Mustakone (4.59%) and others. Site specific docking, ADMET analysis, drug-likeness prediction and bioactivity scoring was conducted among the 12 dominant markers. After compilation of the results, the study found that only 6 markers viz., α-Cadinol, δ-Cadinol, *epi*-Cubenol, Neointermedeol, Shyobunol, and τ-Cadinol qualified most of the parameters. Among these 6 markers, τ-Cadinol, δ-Cadinol displayed high docking score -6.2 Kcal/mol similar to Galantamine followed by Neointermedeol as -5.8 Kcal/mol as *tau* binding affinity. All these three markers followed most of the ADMET properties, drug-likeness parameters and scored moderate to high activity as nuclear receptor ligands and enzyme inhibitors. Compared to δ-Cadinol, τ-Cadinol showed more potential as a base structure to develop candidate oral drug as it was present with a greater area coverage (> 1.1%) in the leaf essential oil of *Knema angustifolia.* The trait analysis exhibited occurrence of essential oils as shared ancestral trait revealing possibility of finding essential oils in other species of *Knema* as well like *Knema globularia, Knema cinerea, Knema lenta* and others. Predictive modelling found that precipitation of warmest quarter as the highest influential climatic variable and *Knema angustifolia* can be reintroduced in North East India to Western Himalayas for long term conservation. However, in-depth study would have been useful for essential oil yield and chemical marker-linked prediction of suitable habitat of *Knema angustifolia* for future.

Author's contributions

RB, JS, and PKB: Sample collection, Data curation, Methodology, Investigation, Experiments, Formal analysis, writing – original draft (part). **PS, PK, MK, PB:** In-silico analysis viz., Molecular docking, Maxent modelling, writing – original draft (part). **SH:** Methodology, Formal analysis. **SPS:** Review, Editing. **DB:** Conceptualization, Sample collection, Identification, Methodology, Supervision, Review, Editing, writing – original and final draft

Declaration of Competing Interest

There is no potential conflict of interests among the authors.

Acknowledgements

The study was funded by Council of Scientific and Industrial Research, Ministry of Science & Technology, Govt. of India, New Delhi and Department of Biotechnology (DBT), Govt. of India, New Delhi (File no. BT/PR27725/NER/95/1349/2018 dated 11.01.2019) and CSIR FBR Project MLP0041. First six authors acknowledge Academy of Scientific and Innovative Research (AcSIR), Ghaziabad - 201002, India for the opportunity and platform to carry out the work. Rubi Barman acknowledges Department of Science & Technology, Govt. of India for DST-Inspire Fellowship (DST/INSPIRE Fellowship/2017/IF170731) for pursuing Ph.D. work. All the authors acknowledge Director, CSIR-NEIST for the logistics and support.

Supplementary materials

Supplementary material associated with this article can be found in the online version at 10.1016/j.sajb.2023.06.030.

References

Abdelrheem, D.A., Rahman, A.A., Elsayed, K.N., Abd El-Mageed, H.R., Mohamed, H.S., Ahmed, S.A., 2021. Isolation, characterization, in vitro anticancer activity, dft calculations, molecular docking, bioactivity score, drug-likeness and admet studies of eight phytoconstituents from brown alga *Sargassum platycarpum*. J. Mol. Struct. 1225, 129245. https://doi.org/10.1016/j.molstruc.2020.129245.

Adhikari, D., Mir, A.H., Upadhaya, K., Iralu, V., Roy, D.K., 2018. Abundance and habitat-suitability relationship deteriorate in fragmented forest landscapes: a case of *Adinandra griffithii* Dyer, a threatened endemic tree from Meghalaya in northeast India. Ecol. Process. 7, 3. https://doi.org/10.1186/s13717-018-0114-z.

Alzheimer's Association. 2023. FDA approved treatments for Alzheimer's. https://www.alz.org/media/documents/fda-approved-treatments-alzheimers-ts.pdf. Accessed on 10.11.2022

Amin, M.L., 2013. P-glycoprotein inhibition for optimal drug delivery. Drug target insights 7, 27–34.

Banik, D., Bora, P.P., 2016. A Taxonomic study on the diversity of Indian *Knema* Lour. (*Myristicaceae*). Taiwania 61 (2), 141–158.

Barman, R., Bora, P.K., Saikia, J., Kemprai, P., Saikia, S.P., Haldar, S., Banik, D., 2021. Nutmegs and wild nutmegs: An update on ethnomedicines, phytochemicals, pharmacology, and toxicity of the *Myristicaceae* species. Phytother. Res. 35 (9), 4632–4659 .

BIOVIA Dassault, 2021. Systèmes, Discovery studio visualizer, version v21.1.0.20298. Dassault Systèmes, San Diego.

Bora, P.K., Saikia, J., Kemprai, P., Saikia, S.P., Banik, D., Haldar, S., 2021. Evaluation of Postharvest Drying, Key Odorants, and Phytotoxins in Plai (*Zingiber montanum*) Essential Oil. J. Agric. Food Chem. 69, 5500–5509.

Boukhatem, M.N., Kameli, A., Ferhat, A.M., Saidi, F., Mekarnia, M., 2013. Rose geranium essential oil as a source of new and safe anti-inflammatory drugs. Libyan J. Med. 8 (1), 22520.

CLC Genomics Workbench 22.0.1 (QIAGEN). QIAGEN, Aarhus, Denmark. www.qiagenbioinformatics.com. Accessed on 1.12.2022

Constantinides, P.P., Wasan, K.M., 2007. Lipid formulation strategies for enhancing intestinal transport and absorption of P-glycoprotein (P-gp) substrate drugs: in vitro/in vivo case studies. J. Pharm. Sci. 96 (2), 235–248.

Corina classic demo, 2022. https://mn-am.com/demos-services/. Accessed on 7 November 2022.

Daina, A., Michielin, O., Zoete, V., 2017. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness, and medicinalchemistry friendliness of small molecules. Sci. Rep. 7 (1), 42717.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods. 9 (8), 772.

Daviet, L., Schalk, M., 2010. Biotechnology in plant essential oil production: progress and perspective in metabolic engineering of the terpene pathway. Flavour Fragr. J. 25, 123–127.

de Wilde, W.J.J.O., 1979. New account of the genus *Knema* (Myristicaceae). Blumea 25 (2), 321–478.

de Wilde, W.J.J.O., 2000. Myristicaceae. In: *Flora Malesiana: Series I-Seed Plants*. National Herbarium Nederland, Universiteit Leiden, Netherlands, pp. 1-634.

de Wilde, W.J.J.O., 2002. Myristicaceae. Flora of Thailand, 7. The Forest Herbarium, Royal Forest Department, Bangkok, Thailand, pp. 720–777. part 4.

Doyle, J., Sauquet, H., Scharaschkin, T., Thomas, A.L., 2004. Phylogeny, molecular and fossil dating, and biogeographic history of *Annonaceae* and *Myristicaceae* (*Magnoliales*). Int. J. Plant Sci. 165 (4 Suppl), S55–S67.

Doyle, J.A., Endress, P.K., 2010. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. J. Syst. Evol. 48 (1), 1–35.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Diversity Distrib 17, 43–57.

Fatima, S., Gupta, P., Sharma, S., Sharma, A., Agarwal, S.M., 2020. ADMET profiling of geographically diverse phytochemical using chemoinformatic tools. Future Med. Chem. 12, 69–87.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24 (1), 38–49.

FigTree ver. 2018-11-25-v1.4.4. http://tree.bio.ed.ac.uk/software/figtree/Accessed on 1.12.2022

Goedert, M.G.S.M., Spillantini, M.G., Jakes, R., Rutherford, D., Crowther, R.A., 1989. Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. Neuron 3, 519–526.

Gogoi, M., Hati-Boruah, J.L., Bora, P.K., Das, D.J., Famhawite, V., Biswas, A., Puro, N., Kalita, J., Haldar, S., Baishya, R., 2021. *Citrus macroptera* induces apoptosis via death receptor and mitochondrial mediated pathway as prooxidant in human non-small cell lung cancer cells. Food Biosci 43, 101293. https://doi.org/10.1016/j.fbio.2021.101293.

Guindon, S., Gascuel, O., 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum-likelihood. Syst. Biol. 52 (5), 696–704.

Hall, B.K., 2007. Homology and homoplasy. Philosophy of Biology. North-Holland, pp. 429–453.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high-resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Husnu, C.B.K., Demirci, F., 2007. Chemistry of essential oils. In: Berger, R.G. (Ed.), Flavours and Fragrances: Chemistry, Bioprocessing and Sustainability. Springer, Berlin, Heidelberg, pp. 43–67.

Khanna, V., Ranganathan, S., 2009. Physicochemical property space distribution among human metabolites, drugs, and toxins. BMC Bioinf 10 (Suppl 15), 1–18.

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Bolton, E.E., 2021. PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res 49 (D1), D1388–D1395.
Knapp, S., 2002. Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the *Solanaceae*. J. Expt. Bot. 53 (377), 2001–2022.
Koul, O., Walia, S., Dhaliwal, G.S., 2008. Essential Oils as Green Pesticides: Potential and Constraints. Biopestic Int 4 (1), 63–84.
Kozlowska, U., Nichols, C., Wiatr, K., Figiel, M., 2021. From psychiatry to neurology: Psychedelics as prospective therapeutics for neurodegenerative disorders. J. Neurochem. 162 (1), 89–108.
Kumar, D., Rawat, S., Joshi, R., 2021. Predicting the current and future suitable habitat distribution of the habitat tree *Oroxylum indicum* (L.) Kurz in India. J. Appl. Res. Med. Aromat. Plants 23, 100309. https://doi.org/10.1016/j.jarmap.2021.100309.
Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol. Biol. Evol. 35 (6), 1547–1549.
Kumari, S., Pundhir, S., Priya, P., Jeena, G., Punetha, A., Chawla, K., Jafaree, Z.F., Mondal, S., Yadav, G., 2014. EssOilDB: a database of essential oils reflecting terpene composition and variability in the plant kingdom. Database 1-12, article ID bau120. doi: 10.1093/database/bau120
Laskowski, R.A., Jabłońska, J., Pravda, L., Vařeková, R.S., Thornton, J.M., 2018. PDBsum: Structural summaries of PDB entries. Protein Sci 27 (1), 129–134. http://www.ebi.ac.uk/pdbsum. Accessed on 10.11.2022.
Lei, P., Ayton, S., Finkelstein, D.I., Adlard, P.A., Masters, C.L., Bush, A.I., 2010. Tau protein: relevance to Parkinson's disease. Int. J. Biochem. Cell Biol. 42 (11), 1775–1778.
Li, B., Wilson, T.K., 2008. Myristicaceae. Flora of China, 7. Science Press. Beijing, Missouri Botanical Garden Press, St. Louis, pp. 96–101. Wu, Z. Y., Raven PH, Hong DY (Eds.).
Lilkova, E., Petkov, P., Ilieva, N., Litov, L., 2015. The PyMOL molecular graphics system, version 2.0. LLC, Schrodinger.
Lovering, F., Bikker, J., Humblet, C., 2009. Escape from flatland: increasing saturation as an approach to improving clinical success. J. Med. Chem. 52 (21), 6752–6756.
[Updated 2003] Madden, T., 2002. The BLAST Sequence Analysis Tool. In: McEntyre, J., Ostell, J. (Eds.), The NCBI Handbook [Internet]. National Center for Biotechnology Information, Bethesda (MD)US, Chapter16.
Maddison, W.P., Maddison, D.R., 2019. Mesquite: a modular system for evolutionary analysis. Version 3.61 http://www.mesquiteproject.org. Accessed on 30 July 2022.
Mahanta, B.P., Sarma, N., Kemprai, P., Begum, T., Saikia, L., Lal, M., Haldar, S., 2020. Hydrodistillation based multifaceted value addition to *Kaempferia galanga* L. leaves, an agricultural residue. Ind. Crops Prod. 154, 112642. https://doi.org/10.1016/j.indcrop.2020.112642.
Martins, A., Hajdu, Z., Anndrea, V., Hohmann, J., 2010. Spathulenol inhibit the human ABCB1 efflux pump. Planta. Med. 76 (12), P608.
Molinspiration Cheminformatics free web services, https://www.molinspiration.com, Slovensky Grob, Slovakia. Accessed on 1.12.2022
Montagne, A., Zhao, Z., Zlokovic, B.V., 2017. Alzheimer's disease: a matter of blood–-brain barrier dysfunction? J. Expt. Med. 214 (11), 3151–3169.
Morales-Garcia, J.A., Calleja-Conde, J., Lopez-Moreno, J.A., Alonso-Gil, S., Sanz-SanCristobal, M., Riba, J., Perez-Castillo, A., 2020. N, N-dimethyltryptamine compound found in the hallucinogenic tea ayahuasca, regulates adult neurogenesis in vitro and in vivo. Transl. Psychiatry 10 (1), 1–14.
Nisha, C.M., Kumar, A., Nair, P., Gupta, N., Silakari, C., Tripathi, T., Kumar, A., 2016. Molecular docking and in silico ADMET study reveal acylguanidine 7a as a potential inhibitor of β-secretase. Adv. Bioinf. https://doi.org/10.1155/2016/9258578 Article ID 9258578.
Ntana, F., Collinge, D.B., Hamberger, B., Bhat, WE.W., Johnson, S., Jensen, B., Jorgen, H., Jorgensen, L., 2020. Method for producing the sesquiterpene viridiflorol with a fungal enzyme (PCT/US2020/050354). Michigan State University, MSU Technologies.
Pearson, R.G., Dawson, T.P., Liu, C., 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. Ecography 27 (3), 285–298.
Phadungkit, M., Rattarom, R., Rattana, S., 2010. Phytochemical screening, antioxidant, antibacterial and cytotoxic activities of *Knema angustifolia* extracts. J. Med. Plants Res. 4 (13), 1269–1272.
Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modelling of species geographic distributions. Ecol. Modell. 190, 231–259.
Phillips, S.J., Dudik, M., 2008. Modelling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. Ecography 31 (2), 161–175.
Phillips, S.J., Elith, J., 2010. POC plots: calibrating species distribution models with presence-only data. Ecology 91 (8), 2476–2484.
Pires, D.E., Blundell, T.L., Ascher, D.B., 2015. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J. Med. Chem. 58 (9), 4066–4072. https://biosig.lab.uq.edu.au/pkcsm/. Accessed on 10.11.2022.
POWO, 2021. Plants of the World Online>Facilitated by the Royal Botanic Gardens, Kew. Published on the internet; http://www.plantsoftheworldonline.org/Accessed on 29.11.2022
Prakash, B., Kedia, A., Mishra, P.K., Dubey, N.K., 2015. Plant essential oils as food preservatives to control moulds, mycotoxin contamination and oxidative deterioration of agri-food commodities e Potentials and challenges. Food Control 47, 381–391.
Prins, C.L., Vieira, I.J.C., Freitas, S.P., 2010. Growth regulators and essential oil production. Braz. J. Plant Physiol. 22, 91–102.
Puce, S., Pica, D., Schiaparelli, S., Negrisolo, E., 2016. Integration of Morphological Data into Molecular Phylogenetic Analysis: Toward the Identikit of the Stylasterid Ancestor. PLoS one 11 (8), e0161423. https://doi.org/10.1371/journal.pone.0161423.
Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Ruchisansakun, S., van-der-Niet, T., Janssens, S.B., Triboun, P., Techaprasan, I., Jenjittikul, T., Suksathan, P., 2015. Phylogenetic Analyses of Molecular Data and Reconstruction of Morphological Character Evolution in Asian *Impatiens* section *Semeiocardium* (*Balsaminaceae*). Syst. Bot. 40 (4), 1063–1074.
Saeger, H.N., Olson, D.E., 2022. Psychedelic-inspired approaches for treating neurodegenerative disorders. J. Neurochem. 162 (1), 109–127.
Salleh, W.M.N.H.W., Anuar, M.Z.A., Khamis, S., Nafiah, M.A., Sul'ain, M.D., 2021. Chemical investigation and biological activities of the essential oil of *Knema kunstleri* Warb. from Malaysia. Nat. Prod. Res. 35 (13), 2279–2284.
Salleh, W.M.N.H.W., Farediah, A., 2017. Phytochemistry and Biological Activities of the Genus *Knema* (*Myristicaceae*). Pharm. Sci. 23 (4), 249–255.
Sanner, M.F., 1999. Python: A Programming Language for Software Integration and Development. J. Mol. Graphics Mod. 17 (1), 57–61.
Sauquet, H., Doyle, J.A., Scharaschkin, T., Thomas, B., Hilu, K.W., Chatrou, L.W., Thomas, A.L., 2003. Phylogenetic analysis of Magnoliales and *Myristicaceae* based on multiple data sets: implications for character evolution. Bot. J. Linn. Soc. 142 (2), 125–186.
Sevcik, J., Skrabana, R., Dvorsky, R., Csokova, N., Iqbal, K., Novak, M., 2007. X-ray structure of the PHF core C-terminus: insight into the folding of the intrinsically disordered protein tau in Alzheimer's disease. FEBS Lett 581 (30), 5872–5878.
Shin, R.W., Iwaki, T., Kitamoto, T., Tateishi, J., 1991. Hydrated autoclave pretreatment enhances tau immunoreactivity in formalin-fixed normal and Alzheimer's disease brain tissues. Lab. Invest. 64, 693–702.
Siga, A., Sarkar, A., Konwar, P., Saikia, J., Saikia, S.P., Banik, D., 2022. Trait-linked phylogenetic framework of *Paphiopedilum* distributed in India revealed species passport trait to prevent unethical trade through *in-silico* study. S. Afr. J. Bot. 150, 420–430.
Silva, S.G., Figueiredo, P.L.B., Nascimento, L.D., da Costa, W.A., Maia, J.G.S., Andrade, E.H.A., 2018. Planting and seasonal and circadian evaluation of a thymol-type oil from *Lippia thymoides*. Mart. & Schauer. Chem. Cent. J. 12 (1), 1–11.
Sinclair, J., 1958. A revision of the Malaysian *Myristicaceae*, 16. Gard Bull, Singapore, pp. 205–472.
Sinclair, J., 1961. The Genus *Knema* (*Myristacaceae*) in Malaysia and Outside Malaysia, 18s. Gard Bull, Singapore, pp. 102–327.
Swets, J., 1988. Measuring the accuracy of diagnostic systems. Science 240 (4857), 1285–1293.
Swofford, D.L., 1998. PAUP*. Phylogenetic analysis using parsimony ver. 4.0. Sinauer, Sunderland, Mass, USA.
Thuiller, W., Lavorel, S., Araujo, M.B., Sykes, M.T., Prentice, I.C., 2005. Climate change threats to plant diversity in Europe. Proc. Nat. Acad. Sci. USA 102 (23), 8245–8250.
Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J. Comput. Chem. 31, 455–461.
Vardhan, S., Sahoo, S.K., 2020. In silico ADMET and molecular docking study on searching potential inhibitors from limonoids and triterpenoids for COVID-19. Comput. Biol. Med. 124, 103936. https://doi.org/10.1016/j.compbiomed.2020.103936.
Wang, Y., Xie, B., Wan, F., Xiao, Q., Dai, L., 2007. Application of ROC curve analysis in evaluating the performance of alien species potential distribution models. Biodivers. Sci. 15 (4), 365–372.
Wei, W., Cherukupalli, S., Jing, L., Liu, X., Zhan, P., 2020. Fsp3: A new parameter for drug-likeness. Drug Discov. Today 25 (10), 1839–1845.