RESEARCH ARTICLE

SABRAO Journal of Breeding and Genetics
57 (5) 1838-1850, 2025
http://doi.org/10.54910/sabrao2025.57.5.6
http://sabraojournal.org/
pISSN 1029-7073; eISSN 2224-8978

GENETIC MAKEUP OF THE MEDICINAL PLANT, *VITEX SP.*, BASED ON DNA BARCODING: AN IN-SILICO ANALYSIS

B.K. SARI[1], N.R. HARTANTO[1], L.H. MANDIRI[1], T. RAHAYU[1], R. MUNAWAROH[2], M. FAWWAZ[3], B. PURWONO[4], and Y. SIDIQ[1*]

[1]Department of Biology Education, Universitas Muhammadiyah Surakarta, Central Java, Indonesia
[2]Department of Pharmacy, Faculty of Pharmacy, Universitas Muhammadiyah Surakarta, Central Java, Indonesia
[3]Department of Pharmaceutical Sciences, Universitas Muslim Indonesia, Makassar, Indonesia
[4]Laboratory of Organic Chemistry, Department of Chemistry, Universitas Gadjah Mada, Yogyakarta, Indonesia
[*]Corresponding author's email: ys120@ums.ac.id
Email addresses of co-authors: a420210037@student.ums.ac.id, a420210038@student.ums.ac.id, listyhastim@gmail.com, tr124@ums.ac.id, rima.munawaroh@ums.ac.id, muammar.fawwaz@umi.ac.id, purwono.bambang@ugm.ac.id

SUMMARY

Vitex species belong to a pivotal genus of the medicinal plants; however, research on their genetic diversity is still limited. Therefore, this study aimed to determine the genetic variation in *Vitex* species based on the DNA barcoding genes *matK* and *rbcL*. The collected data for the *matK* and *rbcL* complete sequences came from the gene bank, with 31 *matK* and 23 *rbcL* accessions analyzed, comprising five geolocations, including China, Korea, Thailand, the Philippines, and Malaysia. The data's analysis used the MEGA 11 to reveal the genetic variation and generate phylogenetic trees. The *Vitex* haplotype network generation utilized the DnaSP and PopART. The results showed the genetic variation within the *matK* gene in *Vitex* species was greater than that in the *rbcL* gene. Overall, 75 and 40 nucleotide polymorphisms succeeded in detecting in *matK* and *rbcL* genes, respectively. Based on the phylogenetic trees of *Vitex* species, *matK* and *rbcL* genes could effectively separate the species into three general clades. The haplotype network analysis revealed 14 and nine haplotypes based on *matK* and *rbcL*, respectively. Genetic variation within *Vitex* species could be useful for the development of specific DNA barcoding for species authentication, conservation, and comprehensive assessment of genetic diversity across the regions.

Keywords: DNA barcoding, genetic variation, nucleotide polymorphism, *matK*, *rbcL*, *Vitex* species

Communicating Editor: Dr. Quaid Hussain

Manuscript received: March 07, 2025; Accepted: May 03, 2025.

Citation: Sari BK, Hartanto NR, Mandiri LH, Rahayu T, Munawaroh R, Fawwaz M, Purwono B, Sidiq Y (2025). Genetic makeup of the medicinal plant, *Vitex sp.*, based on DNA barcoding: An in-silico analysis. *SABRAO J. Breed. Genet.* 57(5): 1838-1850. http://doi.org/10.54910/sabrao2025.57.5.6.

Key findings: This study detected 75 and 40 nucleotide polymorphisms on the *matK* and *rbcL*, respectively, in *Vitex* species. The haplotype network analysis revealed 14 and nine haplotypes based on *matK* and *rbcL*, respectively.

INTRODUCTION

Vitex species are members of the Lamiaceae family, widely used as traditional medicinal plants. Among the 236 genera of the Lamiaceae family, the genus *Vitex* is the most widely desirable as ethnomedicinal plants (Zaki and Salleh*,* 2020). Among several disease symptoms, the most frequent utilization of *Vitex* plants applies to antimicrobial treatment, which comprised 10 *Vitex* plants (Islam *et al.,* 2024). In addition, metabolite-based comprehensive studies on *Vitex* plants have confirmed their potential antioxidant, anticancer (Yan *et al.,* 2023), and antimalaria functions (Arpiwi *et al.,* 2020). The species *Vitex trifolia* belonged to the genus *Vitex,* which was examined for its pharmacological activities to confirm its benefits as a traditional medicine and explore its bioactive compounds (Meng *et al.,* 2023). Some beneficial metabolites derived from *V. trifolia* gained identification, including flavonoids and terpenoids (Mottaghipisheh *et al*., 2024; Fawwaz *et al*., 2024).

The species *Vitex negundo* is a widely used treatment for its anti-inflammatory, analgesic, antioxidant, antimicrobial, antihistaminic, and hepatoprotective activities (Koirala *et al.,* 2020). Moreover, *Vitex agnus-castus* alleviated premenstrual syndrome and menopausal symptoms, with antimicrobial and antioxidant properties. Although genetic engineering is a promising approach for improving the potential of medicinal plants (Hasan *et al.,* 2024), studies on the genetic variation of *Vitex* species are still insufficient.

In medicinal plants, genetic engineering can begin by revealing their genetic variations using DNA barcoding. As several medicinal plants have similar morphological characteristics, the DNA barcoding approach using chloroplast nucleotide sequences can authenticate their scientific name and avoid misidentification. Inaccurate species identification can result in the use of ineffective or potentially harmful substitutes, undermining the efficacy and safety of herbal medicines, as well as contributing to the unintentional exploitation of rare or endangered taxa (Hollingsworth *et al*., 2016). DNA barcoding can also be beneficial to determine the genetic relationship among the numerous closely related species (Antil *et al.,* 2023). Among the plant chloroplast genes, *maturase K* (*matK*) and the large subunit of *ribulose-1,5-bisphosphate carboxylase* (*rbcL*) are the most employed genes. Various complete chloroplast sequences of *Vitex* species have succeeded in the submission and deposited in the National Center of Biotechnology Information (NCBI).

Complete chloroplast sequences, *rbcL*, *atpF-H*, *rpoB*, *rpoC1*, *ndhF*, *matK, trnH-psbA*, r*ps16-trnQ, rpl32-trnL,* and *trnL-F* are powerful barcodes for distinguishing plant species (Ashour *et al.,* 2023). Two of them, *matK* and *rbcL,* can be effective to distinguish various species and subspecies of plants owing to their variability in sequences, including orchid species (Worthy *et al.,* 2022), *Paeonia* taxa (Cetiz *et al.,* 2023), Apiaceae family (Abdelaziz *et al.,* 2024), and ginseng plants (Kim *et al.,* 2025).

The Consortium for the Barcode of Life, or CBOL, has endorsed these two loci due to their superior species discrimination and effective sequence recovery (Utama *et al.,* 2024). The *matK* and *rcbL* can differentiate the species due to their conserved and variable region sequences. In some cases, *rbcL* exhibits low resolution with high universality, whereas *matK* reveals high resolution with low universality across the different species (Ismail *et al.,* 2020). Moreover, *matK* and *rbcL* are crucial for the authentication of medicinal plants, ensuring the correct species, which is the primary requirement for both conservation and therapeutic effectiveness. The *matK* and *rbcL* employment exceeds other common DNA barcoding genes, such as *ITS* and *trnH-psbA*, due to their superior universality and

amplification success, coupled with balanced inter- and intraspecific resolutions. The nuclear ITS region provides higher interspecific discrimination but often exhibits paralogous variants and amplification difficulties. Conversely, the plastid trnH-psbA spacer, while highly variable, frequently contains indels that complicate alignment and phylogenetic inference (Shaw *et al*., 2014).

Although *matK* and *rbcL* are crucial for species identification and plant authentication, the conserved and variable regions of these genes in *Vitex* species have not reached detection. Therefore, the presented study sought to determine the genetic variation in *Vitex* species based on *matK* and *rbcL* genes to identify the conserved and variable regions, as well as the genetic relationship among the *Vitex* species. The promising results will contribute to the initial exploration of genetic variation in *Vitex* species. These findings can also be a basis for further studies on *Vitex* species using DNA-based technologies, including species authentication for medicinal and conservation efforts.

MATERIALS AND METHODS

This study conducted a bioinformatics analysis of *matK* and *rbcL* sequence data in *Vitex* species. These comprised only complete, high-quality sequences with clear geographic metadata to ensure a robust phylogeographic analysis.

Genetic material

This research employed bioinformatics analysis of sequence data of *matK* and *rbcL* genes in *Vitex* species plants. The sequences comprised 31 accessions of *matK* and 23 accessions of *rbcL* in 14 different *Vitex* species. Sequence data collection also came from several countries, including the Philippines, Malaysia, Korea, Thailand, and China.

Data collection

The sequence data collection resulted in accessing the NCBI homepage (https://www.ncbi.nlm.nih.gov/) and searching *matK* and *rbcL* sequences of *Vitex* species. This study used '*matK Vitex* complete' and '*rbcL Vitex* complete' keywords in searching the sequences. Collected sequences underwent subsequent filtering, with only complete sequences of both *matK* and *rbcL* genes included for further analysis. Having geolocation data was also a criterion for sequence inclusion.

Data analysis

Alignment data sequences

Collected *matK* and *rbcL* sequences' organization as FASTA files preceded importing into MEGA 11 (Tamura *et al*., 2021). The multiple sequence alignment, as performed, used MEGA's built-in ClustalW algorithm with default nucleotide settings (gap opening penalty = 15.0; gap extension penalty = 6.66; IUB weight matrix; transition weight = 0.5). After the initial alignment, excluding ambiguous positions—those with gaps or uncertain base calls—continued via pairwise deletion, with the remaining regions manually inspected and trimmed to eliminate poorly aligned segments. Then, the final, curated alignments' exporting in the meg format succeeded for downstream analyses.

Phylogenetic tree construction

Phylogenetic relationships' interpretation in MEGA 11 ensued by selecting the Neighbor-Joining method with the Kimura 2-parameter model of nucleotide substitution. Bootstrap support evaluation with 500 replicates occurred, eliminating positions with less than 95% site coverage to avoid bias from missing data. Ambiguous nucleotide sites received

handling by the pairwise deletion during the tree reconstruction. The resulting tree's visualization and editing transpired within MEGA before exporting as a high-resolution image (Ningrum and Chasani, 2021).

Nucleotide polymorphic site

Aligned *matK* and *rbcL* sequences continued to be imported into MEGA 11 (Tamura *et al*., 2021), incurring analysis under the Analyze > Polymorphism toolkit. Parsimony-informative sites' identification used the Highlight Parsimony-Informative Sites function with default settings. Ambiguous positions—those containing gaps or undetermined bases ("N")—sustained treatment from pairwise deletion to ensure only high-confidence sites contributed to downstream assessments. Identified sites proceeded to be exported via Data > Export Table, selecting "Excel" as the output format and setting "Sites per Line" to the total number of variable positions, ensuring each column corresponded to a single polymorphic site. The exported spreadsheet, containing only highlighted informative sites, underwent importing into the DnaSP v5.0 to calculate the total number of segregating (polymorphic) sites (S), nucleotide diversity (π), and haplotype diversity (Hd) under default neutral evolution assumptions.

Haplotype networks

The polymorphism summary file generated by DnaSP reached conversion to Nexus format before loading it into PopART v1.7 (Leigh and Bryant, 2015). Haplotype relationships deduction used the TCS network algorithm with a 95% connection limit and epsilon = 0 to minimize reticulations. Geographic origin encoding of each accession as a population tag helped visualize clustering by the sampling location. The resulting median-joining network, displayed with node sizes proportional to haplotype frequency and branch lengths, attained scaling to the number of mutational steps, with exported figures saved in SVG format for incorporation into the manuscript.

RESULTS

Gene sequence data of *matK* and *rbcL*

This research commenced by collecting data from the NCBI website to obtain nucleotide sequence data from two gene loci, namely, *matK* and *rbcL*. The collected gene sequence data included 31 accessions of *matK* and 23 accessions of *rbcL* sequences. Thirty-one *matK* accessions comprised 14 *Vitex* species. The length of the *matK* genes ranged from 1512 to 1557 bases (Table 1). Among the 23 *rbcL* accessions, 11 different *Vitex* species existed, and the *rbcL* gene length ranged from 1440 to 1464 bp (Table 2).

Polymorphism in the *matK* and *rbcL* genes

This study employed the DnaSP v5.0 to identify the nucleotide polymorphisms. Among the 1512–1557 nucleotide bases of the *matK* gene in the *Vitex* species, variations appeared in the 28-nucleotide number to 1537 with 75 polymorphic sites (Figure 1). The *matK* sequences also showed 75 variable sites, while the *rbcL* sequences revealed only 40 variable sites (Figure 2). These findings are consistent with those of Ismail *et al.* (2020), who reported that the *matK* gene had a higher resolution and lower universality than the *rbcL* gene in *Acacia* species.

Phylogenetic trees of *Vitex* species

The phylogenetic tree construction engaged the Neighbor-Joining (NJ) method in MEGA 11, which estimates evolutionary relationships by minimizing total branch lengths based on pairwise genetic distances. Bootstrap analysis performed with 500 replicates provided statistical support for each branch. This combination of the NJ method and bootstrap validation offers an efficient and robust approach for inferring evolutionary relationships among the analyzed sequences. The presented results revealed three clades of *Vitex* species based on both *matK* and *rbcL* genes (Figure 3). Both genes discriminated

Table 1. The collected *matK* gene sequences from NCBI GenBank.

No.	Species	NCBI ID	Length (base)	Geolocation	Reference
1	*Vitex negundo*	MW689264.1	1530	China	unpublished
2	*Vitex negundo*	NC_057235.1	1530	China	unpublished
3	*Vitex negundo*	MW366787.1	1530	China	unpublished
4	*Vitex negundo*	MT473783.1	1530	China	Zhao *et al.,* 2021
5	*Vitex quinata*	MT473784.1	1530	China	Zhao *et al.,* 2021
6	*Vitex tripinnata*	MT473785.1	1530	China	Zhao *et al.,* 2021
7	*Vitex yunnanensis*	MT473786.1	1530	China	Zhao *et al.,* 2021
8	*Vitex trifolia*	ON711030.1	1530	China	unpublished
9	*Vitex rotundifolia*	NC_050991.1	1530	Korea	unpublished
10	*Vitex rotundifolia*	MT937186.1	1530	Korea	unpublished
11	*Vitex trifolia*	AB284175.1	1530	Thailand	unpublished
12	*Vitex agnus-castus*	AB284182.1	1530	Thailand	unpublished
13	*Vitex rotundifolia*	AB284177.1	1530	Thailand	unpublished
14	*Vitex limonifolia*	AB284180.1	1539	Thailand	unpublished
15	*Vitex canescens*	AB284181.1	1512	Thailand	unpublished
16	*Vitex carbunculorum*	AB284179.1	1539	Thailand	unpublished
17	*Vitex glabrata*	AB284178.1	1530	Thailand	unpublished
18	*Vitex negundo*	AB284176.1	1530	Thailand	unpublished
19	*Vitex trifolia*	NC_062602.1	1530	Philippines	Gentallan *et al.,* 2022
20	*Vitex trifolia*	OM868083.1	1530	Philippines	Gentallan *et al.,* 2022
21	*Vitex glabrata*	MT473782.1	1530	Philippines	Zhao *et al.,* 2021
22	*Vitex parviflora*	NC_065666.1	1530	Philippines	Bartolome *et al.,* 2023
23	*Vitex parviflora*	ON597620.1	1530	Philippines	Bartolome *et al.,* 2023
24	*Vitex agnus-castus*	PP584505.1	1530	Philippines	Gentallan *et al.,* 2024
25	*Vitex agnus-castus*	NC_088534.1	1530	Philippines	Gentallan *et al.,* 2024
26	*Vitex negundo*	PP584504.1	1530	Philippines	Unpublished
27	*Vitex bicolor*	ON526805.1	1530	Philippines	Unpublished
28	*Vitex bicolor*	NC_065871.1	1530	Philippines	Unpublished
29	*Vitex rotundifolia*	OQ942922.1	1530	Philippines	Unpublished
30	*Vitex longisepala*	NC_084072.1	1557	Malaysia	Unpublished
31	*Vitex longisepala*	OR536591.1	1557	Malaysia	Unpublished

Table 2. The collected *rbcL* gene sequences from NCBI GenBank.

No.	Species	NCBI ID	Length (base)	Geolocation	Reference
1	*Vitex negundo*	PP584504.1	1443	China	Gentallan *et al.,* 2024
2	*Vitex negundo*	NC_057235.1	1443	China	Unpublished
3	*Vitex negundo*	MW366787.1	1443	China	Unpublished
4	*Vitex tripinnata*	MT473785.1	1464	China	Zhao *et al.,* 2021
5	*Vitex yunnanensis*	MT473786.1	1455	China	Zhao *et al.,* 2021
6	*Vitex rotundifolia*	OQ942922.1	1443	China	Unpublished
7	*Vitex rotundifolia*	MT937186.1	1443	Korea	Unpublished
8	*Vitex rotundifolia*	NC_050991.1	1443	Korea	Unpublished
9	*Vitex agnus-castus*	PP584505.1	1443	Philippines	Gentallan *et al.,* 2024
10	*Vitex agnus-castus*	NC_088534.1	1443	Philippines	Gentallan *et al.,* 2024
11	*Vitex bicolor*	NC_065871.1	1443	Philippines	Unpublished
12	*Vitex bicolor*	ON526805.1	1443	Philippines	Unpublished
13	*Vitex glabrata*	MT473782.1	1461	Philippines	Zhao *et al.,* 2021
14	*Vitex negundo*	MW689264.1	1443	Philippines	Unpublished
15	*Vitex negundo*	MT473783.1	1464	Philippines	Zhao *et al.,* 2021
16	*Vitex parviflora*	NC_065666.1	1455	Philippines	Bartolome *et al.,* 2023
17	*Vitex parviflora*	ON597620.1	1455	Philippines	Bartolome *et al.,* 2023
18	*Vitex quinata*	MT473784.1	1461	Philippines	Zhao *et al.,* 2021
19	*Vitex trifolia*	NC_062602.1	1443	Philippines	Gentallan *et al.,* 2022
20	*Vitex trifolia*	ON711030.1	1464	Philippines	Unpublished
21	*Vitex trifolia*	OM868083.1	1443	Philippines	Gentallan *et al.,* 2022
22	*Vitex longisepala*	NC_084072.1	1440	Malaysia	Unpublished
23	*Vitex longisepala*	OR536591.1	1440	Malaysia	Unpublished

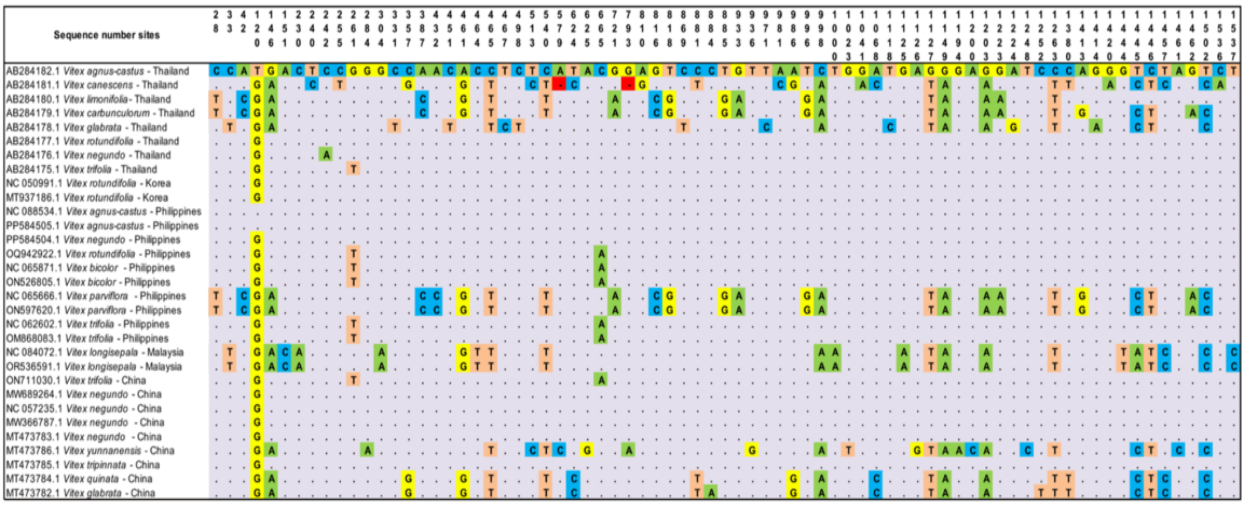

Figure 1. The nucleotide variations of *matK* genes in *Vitex* species. Each row corresponds to an individual *matK* sequence, and each column denotes a variable nucleotide position, numbered according to the reference alignment.

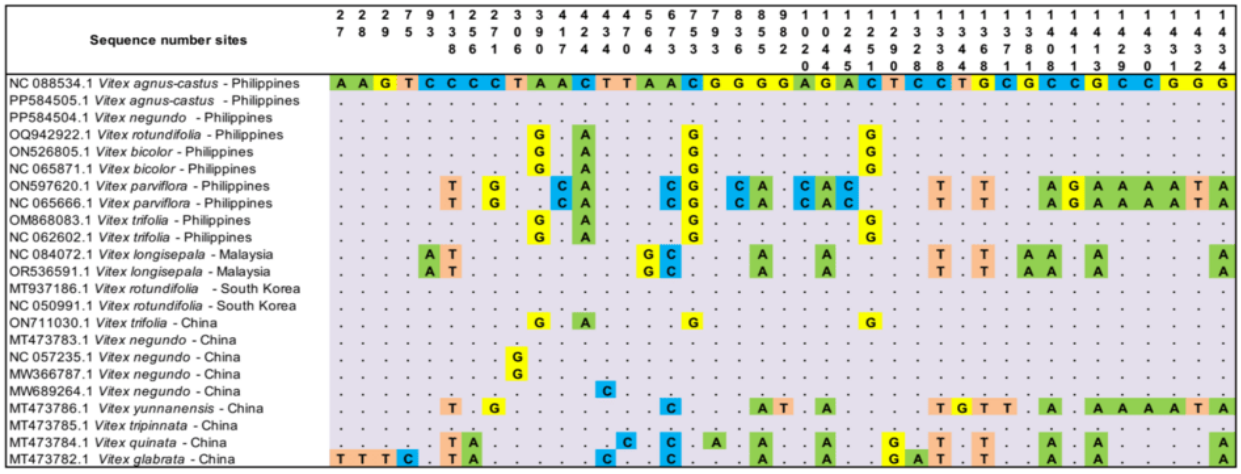

Figure 2. Nucleotide polymorphisms of *rbcL* genes in *Vitex* species. Each row corresponds to an individual *rbcL* sequence, while each column indicates a variable nucleotide position, numbered according to the reference alignment.

against *Vitex* species into three clades with slight differences within species members. This is because a bootstrap value of 90% in phylogeny using mitochondrial genes has an association with a 90% probability of obtaining the same clade in the phylogenetic analysis. Results of the phylogenetic analysis depended on nucleotide polymorphisms. The outcomes showed that *matK* was more reliable than *rbcL* in distinguishing the different species of *Vitex*.

Nucleotide substitution (%) of *matK* and *rbcL* genes

The study revealed minor differences in the genetic variations of *matK* and *rbcL* genes in *Vitex* species plants (Table 3). The substitution in the *rbcL* gene was smaller than that in the *matK*.

Haplotypes in *Vitex* species as per *matK* and *rbcL*

The haplotype results showed *matK* clustered 31 *Vitex* accessions into 14 groups, and H7 had a frequency of nine accessions, making it the largest haplotype, indicating H7 as the best source of evolutionary formation for other haplotypes (Table 4). The haplotype results for *rbcL* divided the 23 accessions into nine groups, with H6 having the frequency of seven accessions, identifying it as the largest haplotype and the best source of evolutionary formation for other haplotypes. Each accession has different alleles from one another. These differences were reliant upon the various combinations of alleles located on homologous chromosomes (Jin *et al.,* 2023).

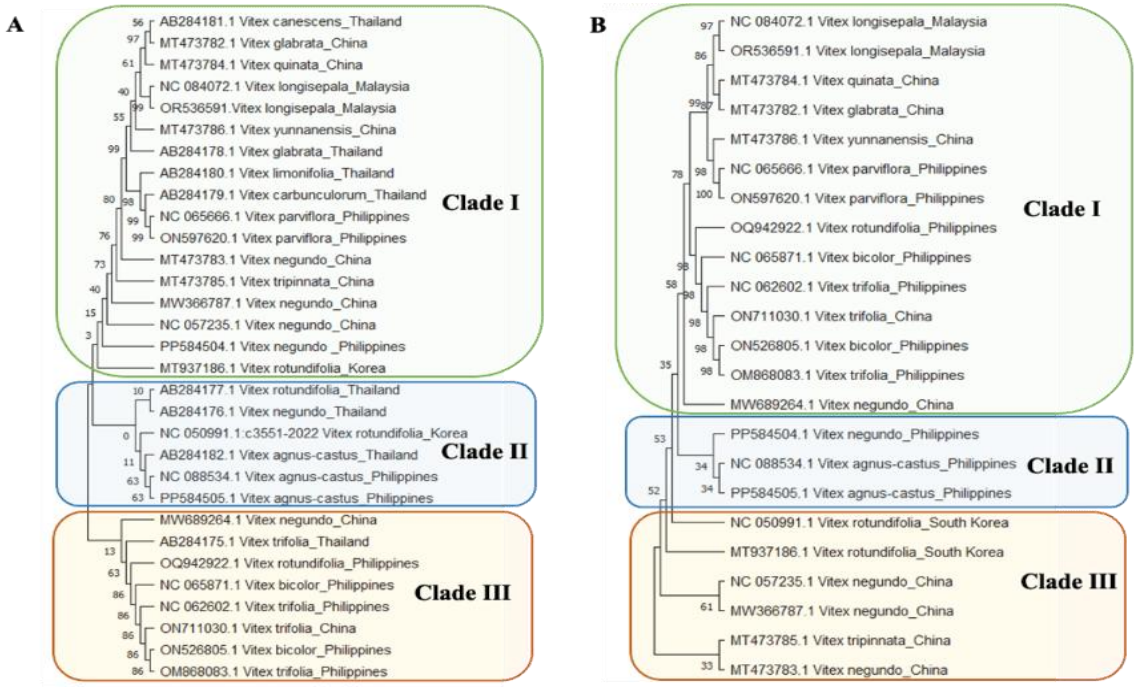

Figure 3. Phylogenetic trees of *Vitex* species constructed using matK (A) and rbcL (B) gene sequences. The trees illustrate the evolutionary relationships among *Vitex* accessions from various geographic regions, grouped into three main clades (Clade I, II, and III). Bootstrap values, shown at each node, represent the percentage of support from 500 replicates, indicating the statistical reliability of each branch; values closer to 100 suggest stronger support. The observed clustering highlights genetic diversity and potential biogeographical patterns within *Vitex* species. GenBank accession numbers and origins are indicated for each taxon.

Table 3. Nucleotide substitution (%) of *matK* and *rbcL* genes in the *Vitex* species.

	matK gene				*rbcL* gene			
	A	T	C	G	A	T	C	G
A	-	0.26	0.14	11.05	-	08.16	05.54	13.06
T	0.22	-	0.51	0.13	8	-	0.27	07.15
C	0.22	24.12	-	0.13	8	0.38	-	07.15
G	0.88	0.26	0.14	-	0.63	08.16	05.54	-

Table 4. List of haplotypes of *Vitex* plants according to their *matK* and *rbcL* loci genes.

MatK gene			*RbcL* gene		
Haplotypes	Number of samples	Accessions	Haplotypes	Number of samples	Accessions
H1	2	NC 065666.1	H1	1	MT473782.1
		ON597620.1	H2	1	MT473784.1
H2	1	AB284179.1	H3	2	NC 084072.1
H3	1	AB284180.1			OR536591.1
H4	1	AB284178.1	H4	1	MT473786.1
H5	1	MT473786.1	H5	2	NC 065666.1
H6	2	NC 084072.1			ON597620.1
		OR536591.1	H6	7	NC 088534.1
H7	9	AB284177.1			PP584505.1
		MT473785.1			PP584504.1
		MW366787.1			MT937186.1
		NC 057235.1			NC 050991.1
		MW689264.1			MT473783.1
		PP584504.1			MT473785.1
		MT937186.1	H7	6	OQ942922.1
		NC 050991.1			ON526805.1
		MT473783.1			OM868083.1
H8	3	AB284182.1			NC 065871.1
		NC 088534.1			NC 062602.1
		PP584505.1			ON711030.1
H9	1	AB284176.1	H8	2	NC 057235.1
H10	1	AB284175.1			MW366787.1
H11	6	OQ942922.1	H9	1	MW689264.1
		NC 065871.1			
		NC 062602.1			
		ON711030.1			
		ON526805.1			
		OM868083.1			
H12	1	MT473784.1			
H13	1	AB284181.1			
H14	1	MT473782.1			

DISCUSSION

The study included all the NCBI ID prefixes, namely, NC_ as reference nucleotides and other prefixes, such as MT, AB, PP, and ON, as the raw material submitted by various researchers. In this study, researchers collected all complete sequences of *rbcL* and *matK* genes from the *Vitex* species. All the *matK* accession sequences comprised five geolocations, namely, China, Korea, the Philippines, Thailand, and Malaysia. The *rbcL* accessions consisted of four geolocations, including China, Korea, the Philippines, and Malaysia. The results were consistent with the findings of Ahmed *et al*. (2023), who reported that *matK* showed higher species discrimination efficiency than *rbcL* due to its highly variable nature.

Nucleotide alignment determined that *Vitex agnus-castus* (AB284182.1), Thailand's accession, was the nucleotide reference. As shown in Figure 1, the same species with two accessions from the Philippines (NC_088534.1 and PP584505.1) were identical to the reference nucleotide. The same species also have close relationships and conserved nucleotide sequences across various countries. Different *Vitex* species showed at least one SNP difference, even within the same genus. The results disclosed that nucleotide polymorphism in the *rbcL* gene of *Vitex* species was lower than the same in the *matK* gene. This outcome corresponds to the findings of Patel *et al.* (2024), which stated that *matK*

gene barcoding exhibited robust and highly reproducible differentiation among the species of a genus.

The lower resolution of the *rbcL* gene may be due to its involvement in the photosynthetic process, as it encodes the rubisco enzyme. Therefore, *rbcL* is important for maintenance and the prevention of mutation. As DNA barcoding should discriminate among the species and lower taxa, finding more variation was better (Nair *et al.,* 2024). However, *matK* and *rbcL* varied based upon the plant species. In the case of *Prunus armeniaca* (*Rosaceae*), the use of *matK* and *rbcL* genes served to differentiate *P. armeniaca* from the other species (Sevindik *et al.,* 2024). Notably, the *rbcL* was more powerful than *matK* in discriminating the orchid species. Although most of the plant species can gain authentication by *matK* and *rbcL* genes (Nair *et al.,* 2024), these two genes do not discriminate among the *Cinnamomum* species (Chandrasekara *et al.,* 2021).

The genetic variation of *matK* and *rbcL* genes in *Vitex* species was applicable to identifying the evolutionary relationship. Evolutionary forces continually alter the genetic frequencies of plant species, affecting the genetic diversity of a population (Begna, 2021). The phylogenetic tree showed a genetic relationship and evolutionary history of the species being studied. Wang *et al.* (2023) reported the phylogenetic tree aligns with the evolutionary relationship among the species, demonstrating the consistency of traditional taxonomy with molecular classification. Varying branch lengths of the tree contain key evolutionary information. Longer branches indicate greater evolutionary distances, showing significant divergence over time (Suvorov and Schrider, 2024). Conversely, the shorter branches suggest a closer evolutionary relationship and more recent common ancestry. These branches are effective for molecular dating, reconciling gene trees with species trees, and measuring the biodiversity.

Based on the *matK* gene, the species *Vitex agnus-castus* accession AB284182.1 from Thailand was identical to the species *Vitex agnus-castus* accessions NC_088534.1 and PP584505.1 from the Philippines. Although these accessions came from two distinct countries, the phylogenetic analysis grouped the three accessions into a single clade with a very close relationship (Figure 3A). Contrastingly, the *rbcL* genes of the *Vitex* species accessions showed identical nucleotide sequences, albeit in different species. Based on the *rbcL* gene, the *Vitex* species *agnus-castus* from the Philippines (NC_088534.1), used as a nucleotide reference, was identical to *V. negundo* from the Philippines (PP584504.1), *V. rotundifolia* from Korea (MT937186.1), and the species *V. negundo* from China (MT473783.1) (Figure 2). Correspondingly, the phylogenetic tree revealed a similar pattern, where the species *V. agnus-castus* and *V. negundo* from the Philippines obtained grouping in the same clade with 34 distances from the ancestor. Meanwhile, the species *V. rotundifolia* and *V. negundo* reached placement in different clades at close distances of 52 and 61, respectively (Figure 3B). These results support the low sensitivity and high universality of *rbcL* in differentiating plant diversity (Ismail *et al.,* 2020). The phylogenetic analysis based on *matK* sequences demonstrates enhanced resolution at the species level, attributable to the gene's relatively higher mutation rate, which facilitates discrimination among closely related taxa. Conversely, the rbcL-based phylogeny, owing to the gene's conserved nature, yields strong support for relationships at higher taxonomic levels but exhibits limited resolution for distinguishing species within the genus.

The phylogenetic tree indicated that the three species (*V. agnus-castus, V. negundo,* and *V. rotundifolia*) were closest to the common ancestor. The three other species, *V. parviflora, V. carbunculorum,* and *V. longisepala,* were the most distant from their common ancestor (Figure 3A). Based on the *rbcL* gene, the species *V. agnus-castus*, *V. negundo*, and *V. rotundifolia* have a close relationship with a common ancestor. The *rbcL* gene showed little difference in the most distant species, while the m*atK* gene expressed the greater difference with the species *V. parviflora, V. longisepala*, and *V. trifolia* with a wider distance. The geographic origins of species also influence the clade composition.

For instance, the *rbcL* tree highlights regional groupings, such as species from the Philippines clustering within Clade II, whereas the *matK* tree reveals broader evolutionary patterns. However, some species consistently group together in both trees, strengthening their close evolutionary relationships. For instance, *Vitex longisepala* sets within Clade I in both the *matK* and *rbcL* trees. Even though both markers recover similar major clades, *matK* provides better species-level resolution due to its higher mutation rate and sequence variability. In contrast, *rbcL* is more conserved, supporting deeper nodes and higher-order relationships. Together, *matK* and *rbcL* complement each other, enhancing phylogenetic inference in *Vitex*.

Based on nucleotide substitution in the *matK* gene, the most common substitution among the *Vitex* species was C to T (24.12%), followed by A to G, C to A, and T to C. Meanwhile, in the *rbcL* gene, the highest percentage of substitution was 13.06% from A to G, followed by G to A, C to T, and T to C (Table 3). Specifically, in the plant genome, the nucleotide substitution rate in chloroplast genes was higher than that in nuclear and mitochondrial genes. This study corresponds to that of Dong *et al.* (2020), who stated that purine-to-purine or pyrimidine-to-pyrimidine substitutions (transitions) occur more frequently than transversions (purine-to-pyrimidine or vice versa).

Nucleotide variations and mutations in *matK* and *rbcL* genes can reveal the evolutionary pattern and gene flow across the geographical barriers. These variations can provide insights into migration patterns, population structure, and speciation events (Garcia *et al.,* 2021). Visualizations help in understanding microevolution and genetic diversity within the populations. This group comprises the individuals within a population who have a close genetic relationship. Such clustering allows for the easy classification of accessions into distinct populations. Short lines connecting different circles represent mutation points between the haplotypes, illustrating the occurrence of genetic variation.

The haplotype network grouped the *Vitex* species into 14 haplotypes of *matK* loci and nine haplotypes of *rbcL* loci (Table 4). The evolutionary process may have gained a boost from the nine accessions in haplotype 7 (H7), as the most prominent genetic sharing. The H7 branched into H8, H9, and H10, with a single mutation point in each line (Figure 4A). One branch, H11, comprising six accessions, developed from H10 with a single mutation. Other branches developed from H7 were prominent with multiple mutations in H3, H2, H1, H4, H5, H6, H12, H13, and H14. Interestingly, different *Vitex* species, such as *V. negundo* and *V. rotundifolia,* pooled genetic flow in H7. *Vitex* species from China, Korea, Thailand, and the Philippines shared the genetic variations in H7, H11, and H8. The species *Vitex longifolia* from Malaysia had discriminations from other geolocations.

Meanwhile, the *rbcL* genes evolved from H6 with seven accessions from China, Korea, and the Philippines to H8 and H9 with a single mutation point in each accession, two accessions, and one accession, respectively (Figure 4B). In accordance with *matK*, the *rbcL* distinguished between *V. longifolia* from Malaysia and the same species from other geolocations. Interestingly, based on *matK* and *rbcL* genes, grouping the same species from different geolocations can happen into haplotypes. This indicates gene flow across the different countries. In *Vitex* species, the gene flow may occur through selective movement by humans and natural separation by geographical factors. Moreover, the genetic flow of both *matK* and *rbcL* depended not only on the species but also on the location.

CONCLUSIONS

Nucleotide polymorphism of *matK* and *rbcL* genes in the *Vitex* species classified them into three main clades with higher confidence. The *matK* gene showed greater discriminatory power for species differentiation than the *rbcL*. Haplotype network analysis revealed gene flow, with species *Vitex longifolia* from Malaysia

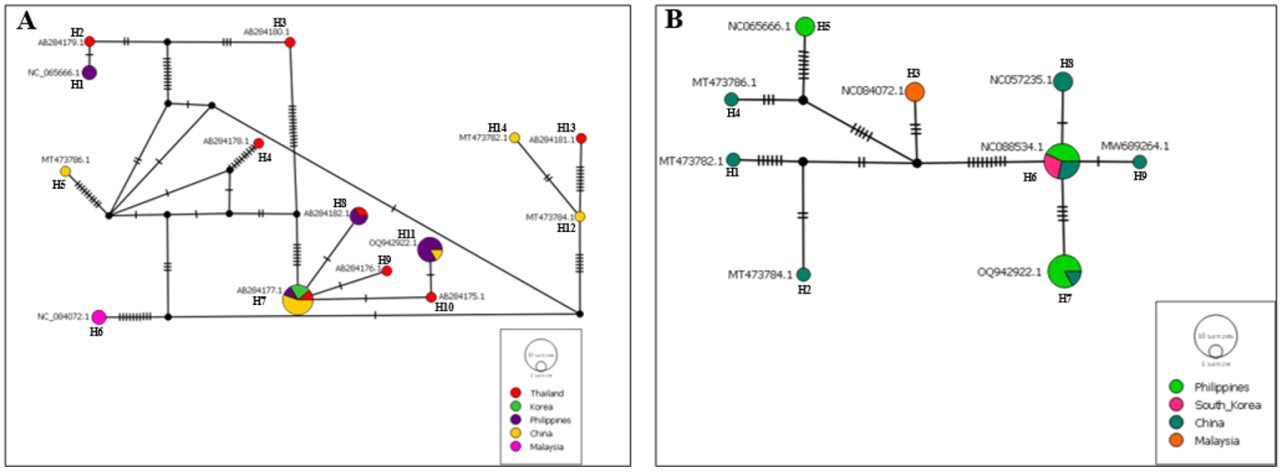

Figure 4. The haplotype network of *Vitex* plants based on (A) *matK* and (B) *rbcL* loci genes. The haplotype results showed that *matK* clustered 31 *Vitex* accessions into 14 groups. The haplotype results for *rbcL* divided the 23 accessions into nine groups. Haplotype serves as a method to determine the evolutionary proximity of species based on their geolocation. The letter H indicates haplotype. The circle sizes represent sample numbers, connecting lines denote the evolutionary pattern among haplotypes, and short lines indicate the nucleotide mutations.

forming a distinct lineage. The genetic flow of both *matK* and *rbcL* acquired influences from both species and geographic location. Identification of conserved and variable regions in *matK* and *rbcL* provides valuable insights for developing universal primers and molecular authentication of medicinal plants. Furthermore, DNA barcoding using *matK* and *rbcL* enhances quality control in the herbal industry by preventing adulteration and ensuring species authenticity, particularly when morphological identification is unreliable.

ACKNOWLEDGMENTS

This study received financial support from the Kolaborasi Riset Strategis (KATALIS) of Kementerian Pendidikan Tinggi, Sains, dan Teknologi, Indonesia, under contract number 003/LL6/P. Katalis/AL.04/2024 and 279.07/A.3-III/LRI/VIII/2024.

REFERENCES

Abdelaziz SA, Khaled, Rania, Al-Kordy MA, El-Domyati FM, Moghazee MM (2024). Comparison of four DNA barcoding loci to distinguish between some Apiaceae family species. *Beni-Suef Univ. J. Basic Appl. Sci*. 13: 1–12.

Ahmed SS, Rahman MO, Ali MA, Hemaid FA, Lee J (2023). Molecular phylogenetics and molecular dating of Arecaceae in Bangladesh inferred from *matK* and *rbcL* genes. *Bangladesh J. Plant Taxon* 30(2): 213–232.

Antil S, Abraham JS, Sripoorna S, Maurya S, Dagar J, Makhija S, Bhagat P, Gupta R, Sood U, Lal R, Toteja R (2023). DNA barcoding, an effective tool for species identification: A review. *Mol. Biol. Rep.* 50(1): 761–775.

Arpiwi NL, Muksin IK, Kriswiyanti E (2020). Essential oils from *Vitex trifolia* as an effective repellent for *Aedes aegypti*. *Biodiversitas* 21(10): 4536–4544.

Ashour BM, Rabiei M, Shiran B (2023). Intraspecific identification of some pomegranate (*Punica granatum* L.) genotypes based on DNA barcoding and morpho-biochemical characteristics. *J. Trees* 37(5): 1435–1442.

Bartolome MC, Gentallan Jr R, Quiñones KJ, Madayag R, Borromeo T, Timog EB (2023). The complete chloroplast genome of *Vitex parviflora* A. Juss. (Lamiaceae) and its comparison with *Vitex* species. *J. For. Res*. 28(3): 217–224.

Begna T (2021). Role and economic importance of crop genetic diversity in food security. *Int. J. Agric. Sci. Food Tech*. 7(1): 164–169.

Cetiz MV, Turumtay EA, Burnaz NA, Özhatay FN, Kaya E, Memon A, Turumtay H (2023). Phylogenetic analysis based on the ITS, *matK* and *rbcL* DNA barcodes and comparison of chemical contents of twelve *Paeonia* taxa in Türkiye. *Mol. Biol. Rep*. 50(6): 5195–5208.

Chandrasekara CHWMRB, Naranpanawa DNU, Bandusekara BS, Pushpakumara DKNG, Wijesundera DSA, Bandaranayake PCG (2021). Universal barcoding regions, *rbcL*, *matK* and *trnH-psbA* do not discriminate *Cinnamomum* species in Sri Lanka. *PloS One* 16(2): e0245592.

Dong W, Xu C, Wen J, Zhou S (2020). Evolutionary directions of single nucleotide substitutions and structural mutations in the chloroplast genomes of the family Calycanthaceae. *BMC Ecol. Evol.* 20: 96.

Fawwaz M, Purwono B, Sidiq Y, Arwansyah, Arsul MI, Fitriana, Pratama M, Fajriani A, Kusuma AT, Amirullah, Musdalifah AA, Sarfandi A, Baits M (2024). Anti-inflammatory, antioxidant, and antibacterial activities with molecular docking studies of *Vitex trifolia* L. targeting human COX-2 and Peroxiredoxin-5. *Chemistry Select* 9: e202403834.

Garcia E, Wright D, Gatins R, Roberts MB, Pinheiro HT, Salas E, Chen J-Y, Winnikoff JR, Bernardi G (2021). Haplotype network branch diversity, a new metric combining genetic and topological diversity to compare the complexity of haplotype networks. *PloS One* 16(6): e0251878.

Gentallan Jr RP, Quiñones KJ, Bartolome MC, Madayag RE, Vera-Cruz JR, Cirunay AT, Endonela LA, Borromeo TH, Altoveros NC, Lalusin AG, Alvaran BBS, Magtoltol JB, Cejalvo RDC (2022). The complete chloroplast genome of *Vitex trifolia* L. (Lamiaceae). *Mitochondrial DNA, Part B. Resour.* 7(7): 1316–1318.

Gentallan Jr RP, Sengun S, Bartolome MCB, Quiñones KJO, Coronado NB, Borromeo TH, Timog EBS (2024). The *Vitex trifolia* complex (Lamiaceae) in the Philippines. *Phyto Keys* 248: 1–40.

Hasan N, Laskar RA, Farooqui SA, Naaz N, Sharma N, Budakoti M, Joshi DC, Choudhary S, Bhinda MS (2024). Genetic improvement of medicinal and aromatic plant species: Breeding techniques, conservative practices and future prospects. *Crop Design* 3(4): 100080.

Hollingsworth PM, Graham SW, Little DP (2016). Choosing and using a plant DNA barcode. *PLoS One* 11(1):e0146390.

Islam Z, Caldeira GI, Caniça M, Islam N, Silva O (2024). *Vitex* genus as a source of antimicrobial agents. *Plants* 13(3): 401.

Ismail M, Ahmad A, Nadeem M, Javed MA, Khan SH, Khawaish I, Sthanadar AA, Qari SH, Alghanem SM, Khan KA, Khan MF, Qamer S (2020). Development of DNA barcodes for selected *Acacia* species by using *rbcL* and *matK* DNA markers. *Saudi J. Biol. Sci.* 27(12): 3735–3742.

Jin X, Du H, Zhu C, Wan H, Liu F, Ruan J, Zhu A (2023). Haplotype-resolved genomes of wild octoploid progenitors illuminate genomic diversifications from wild relatives to cultivated strawberry. *Nat. Plants 9*(8): 1252–1266.

Kim DY, Miranda-Romo D, Cate ART, Hellberg RS (2025). Use of a novel combination of multiplex PCR and DNA barcoding in assessing authenticity of ginseng products. *Food Control* 168: 110893.

Koirala N, Dhakal C, Munankarmi NN, Ali SW, Hameed A, Martins N, Sharifi-Rad J, Imran M, Arif AM, Hanif MS, Basnyat RC, Salehi B (2020). *Vitex negundo* Linn.: Phytochemical composition, nutritional analysis, and antioxidant and antimicrobial activity. *Cell Mol. Biol.* 66(4): 1–7.

Meng X, Wang H, Kuang Z, Wu Y, Su X, Wang J, Li L, Liu C, Jia M (2023). Traditional use, phytochemistry and pharmacology of *Viticis fructus*. *Heliyon* 9(9): e19144–e19144.

Mottaghipisheh J, Kamali M, Doustimotlagh AH, Nowroozzadeh MH, Rasekh F, Hashempur MH, Iraji A (2024). A comprehensive review of ethnomedicinal approaches, phytochemical analysis, and pharmacological potential of *Vitex trifolia* L. *Front. Pharmacol*. 15: 1322083.

Nair VD, Aseema P, Saini KC (2024). DNA barcoding: An effective molecular tool for species identification, molecular authentication and phylogeny studies in plant science research. *Plant Sci. Today* 11(4): 1–15.

Ningrum AM, Chasani AR (2021). Numerical phenetic and phylogenetic relationships in silico among brown seaweeds (*Phaeophyceae*) from Gunungkidul, Yogyakarta, Indonesia. *Biodiversitas* 22(6): 3057–3064.

Patel N, Otieno WO, Gawande ND, Parmar SHNK, Sankaranarayanan S (2024). DNA barcoding-based molecular profiling of Bougainvillea, Dianthus, and Plumeria using *matK* locus*. J. Appl. Biol. Biotechnol.* 13(1): 119–126.

Sevindik E, Korkom Y, Murathan ZT (2024). Evaluating DNA barcoding using cpDNA, *matK,* and *rbcL* for species identification and phylogenetic analysis of *Prunus armeniaca* L. (Rosaceae) genotypes. *Genet. Resour. Crop Evol.* 71(5): 1825–1835.

Shaw J, Lickey EB, Schilling EE, Small RL (2014). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: trnH-psbA is a promising candidate. *Mol. Biol. Evol.* 31(12):3033–3041.

Suvorov A, Schrider DR (2024). Reliable estimation of tree branch lengths using deep neural networks. *PLoS Comput. Biol.* 20(8): e1012337.

Tamura K, Stecher G, Kumar S (2021). MEGA 11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol*. 38(7): 3022–3027.

Utama MN, Etikawati N, Sugiyarto S, Susilowati A (2024). New specific primer *matK* and *rbcL* region for DNA barcode pitcher plant *Nepenthes spathulata*. *Biodiversitas 25*(6): 2515–2523.

Wang Y, Chen S, Chen J, Chen C, Lin X, Peng H, Zhao Q, Wang X (2023). Characterization and phylogenetic analysis of the complete mitochondrial genome sequence of *Photinia serratifolia*. *Sci. Rep*. 13(1): 770.

Worthy SJ, Bucalo K, Perry E, Reynolds A, Cruse-Sanders J, Pérez ÁJ, Burgess KS (2022). Ability of *rbcL* and *matK* DNA barcodes to discriminate between montane forest orchids. *Plant Syst. Evol*. 308(3): 19.

Yan CX, Wei YW, Li H, Xu K, Zhai RX, Meng DC, Fu XJ, Ren X (2023). *Vitex rotundifolia* L. f. and *Vitex trifolia* L.: A review on their traditional medicine, phytochemistry, pharmacology. *J. Ethnopharmacol.* 308: 116273.

Zaki FLM, Salleh WMNHW (2020). Essential oils and biological activities of the genus *Vitex* (Lamiaceae) – A review. *Nat. Vol. Essential Oils* 7(2): 13–21.

Zhao F, Chen YP, Salmaki Y, Drew BT, Wilson TC, Scheen AC, Celep F, Bräuchler C, Bendiksby M, Wang Q, Min DZ, Peng H, Olmstead RG, Li B, Xiang CL (2021). An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Boil*. 19: 1–27.