Berita Biologi
LIPI
Jurnal ilmu-ilmu hayati

Berita Biologi 23(3): 421–430 (2024)
https://ejournal.brin.go.id/berita_biologi ISSN 2077 7019
DOI: 10.55981/beritabiologi.2024.5800

P-ISSN 0126-1754
E-ISSN 2337-8751

ARTICLE

ASSESMENT OF THE DNA BARCODES CHARACTERISTIC AND EVALUATION OF PHYLOGENETIC RELATIONSHIP OF *Castanopsis argentea* (Blume) A. DC.

[*Karakterisasi DNA Barcoding dan Hubungan Kekerabatan Castanopsis argentea (Blume) A. DC.*]

Syifara Chika[1*], Shofiyyatuz Zahro[2]

[1]Master Program of Biology, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia
[2]Plant Biology Study Program, Graduate School, Department of Biology, Faculty of Mathematics and Natural Sciences, IPB University, Bogor, Indonesia

ABSRACT

Castanopsis, the third largest genus under Fagaceae, is widespread in tropical and subtropical areas in East and South Asia. This plant is one of the woody plants that has the potential to be developed because it is helpful for wildlife for nesting and is used in land reforestation activities. Based on data from the Red List of the International Union for the Conservation of Nature and Natural Resources (IUCN), it is reported that the *Castanopsis argentea* species is threatened with extinction. Research based on genetic aspects of this species is also rarely carried out. This research aims to conduct an in silico study and analyze the kinship relationships of C*astanopsis argentea.* The method used in this research is the in silico method, which takes *Castanopsis argentea* nucleotide data from NCBI in the *mat*K region. Based on this research, the phylogenetic results show that the results of the phylogenetic tree reconstruction show that the *mat*K region is divided into two ingroup clades and one outgroup. The *mat*K region in *Castanopsis argentea* is highly conserved because it only has three genetic variations namely N282T, C285T, and C422A. In this study, the *mat*K gene can group species from the same genus and separate species from different genus. This is in line with the results of the phylogenetic tree, which shows that members of the same genus successfully grouped into one clade. More research on DNA barcoding of *Castanopsis argentea* must be carried out and developed because more genetic data still needs to be collected at NCBI. The genetic data of a species is essential to research and is stored in NCBI's Genbank for genetic conservation purposes.

Keywords: *Castanopsis argentea*, DNA barcoding, in silico, phylogenetic tree

Submitted: 5 Agustus 2024; **Revision:** 2 Oktober 2024; **Accepted:** 17 Oktober 2024
***Corresponding author:** E-mail - syifarachika7@gmail.com

ABSTRAK

Castanopsis sebagai genus terbesar ketiga di bawah Fagaceae, tersebar luas di daerah tropis dan subtropis di Asia Timur dan Selatan. Tumbuhan ini termasuk salah satu tumbuhan berkayu yang mempunyai potensi untuk dikembangkan karena bermanfaat bagi satwa liar untuk bersarang, dan digunakan dalam kegiatan reboisasi lahan. Berdasarkan data dari The red list of the International Union for the Conservation of Nature and Natural Resources (IUCN) melaporkan bahwa spesies Castanopsis argentea berada pada status terancam punah. Penelitian berdasarkan aspek genetik tentang spesies ini juga masih jarang dilakukan. Penelitian ini bertujuan untuk melakukan studi in silico dan menganalisis hubungan kekerabatan Castanopsis argentea. Metode yang digunakan dalam penelitian ini adalah metode in silico dengan mengambil data nukleotida Castanopsis argentea dari NCBI pada region matK. Hasil rekonstruksi pohon filogenetik menunjukkan bahwa pada region matK terbagi menjadi dua clade ingroup dan satu outgrup. Wilayah DNA barcode matK pada Castanopsis argentea bersifat sangat conserved karena hanya memiliki tiga variasi genetik yaitu N282T, C285T, dan C422A jika dibandingkan dengan sekuen pembanding lainnya. Pada penelitian ini, gen matK memiliki kemampuan untuk mengelompokkan spesies yang berasal dari genus yang sama dan memisahkan spesies yang berasal dari genus yang berbeda. Hal tersebut selaras dengan hasil pohon filogenetik yang menunjukkan bahwa anggota spesies dari genus yang sama berhasil mengelompok pada satu clade. Penelitian tentang DNA barcoding Castanopsis argentea harus lebih banyak dilakukan dan dikembangkan karena data genetik di NCBI masih sangat sedikit. Data genetik suatu spesies penting untuk diteliti dan disimpan di Genbank NCBI untuk tujuan konservasi genetik.

***Kata kunci**: Castanopsis argentea, DNA barcoding, in silico, pohon filogenetik*

INTRODUCTION

Castanopsis is the third-largest genus of the Fagaceae family. One *Castanopsis* member with quite a high potential is *Castanopsis argentea*, commonly known by the local name Saninten. *Castanopsis argentea*, as an indigenous species, plays an important role in mountain ecosystems with wide canopies. Previous research revealed that *Castanopsis* species are abundantly distributed, especially in Southern China and China (Chen *et al*., 2013; Cheuk & Fischer, 2021). The distribution of *Castanopsis argentea* in Indonesia can be found on the islands of Sumatra and Java (Whitmore TC, 1986). This tree is a place for wildlife, especially birds and mammals, to find food, rest, and nest. This plant is used for reforestation activities on land with a high stone content (Wibowo, 2006). Saninten wood is often used to build houses in West Java, and saninten bark can be used as a natural dye for rattan. The seeds of this plant can be used as food by boiling and burning. Saninten is one of the plants with potential for revegetation activities on ex-mining land (Mansur, 2013).

Castanopsis argentea tree has the morphological characteristics of a wide tree canopy, grayish brown bark, smaller leaves, shorter fruit spines, and a pyramid seed shape, in one fruit there are 3-5 seeds. Medium-sized Saninten trees reach 30 m in height and 60 cm in diameter, the surface of the trunk has longitudinal grooves, the bark is rough and cracked black. Leaves are single, elongated, alternate or spirally arranged, oblong or oval (7-12 cm x 2-3.5 cm), the upper surface of the leaf is smooth and coated with wax, the lower surface is silvery gray. Panicle inflorescences, unisexual. The fruit is clustered, egg-shaped and covered with sharp spines. Each fruit contains three rounded seeds (Pennida *et al*., 2024).

This plant naturally grows in protected forests. However, the number of *Castanopsis argentea* is currently decreasing due to long growth rates, logging for wood extraction, and little natural regeneration. The government has tried to introduce the potential of the *Castanopsis argentea* plant by issuing Minister of Forestry Regulation Number P.57/Menhut-II/2008 concerning Strategic Directions for National Species Conservation 2008-2018. This plant is used as a keystone, a species with great potential to be cultivated and developed because of its broad benefits (Chika *et al*., 2022). Based on data from the red list of the International Union for the Conservation of Nature and Natural Resources (IUCN), it is reported that the species *Castanopsis argentea* is endangered (Barstow & Kartawinata, 2018; Putri *et al*., 2022).

The development of this type as plantation forest or agroforestry outside its natural distribution (ex-situ germplasm conservation) can reduce pressure on these protected forests. This means it is necessary to conserve this species both in-situ and ex-situ (Hilwan *et al*., 2018). Ex-situ

conservation can be carried out at Genbank or botanical gardens. Besides maintaining the number of species populations, conservation activities must also consider genetic aspects (Chika *et al*., 2024). The collection and preservation of genetic resources at Genbank serves as a defense for plant life against the impacts of climate change. Management of plant genetic material is essential for the long-term preservation and utilization of plant genetic resources. Genbank is a repository that stores genetic material from various organisms and can be used for scientific research, conservation, and plant breeding. Genetic conservation aims to facilitate the use of genetic resources and preserve and protect genetic diversity (Aribi, 2024).

Genetic analysis and conservation require the involvement of DNA barcoding. Since 2003, DNA barcoding has been used intensively for organism identification (Chika & Zahro, 2024). DNA barcoding is used as a molecular tool to identify unknown organisms from a small number of tissues processed through standard genome region sequencing (Cowan *et al*., 2006). According to the theory put forward by Pang *et al*. (2012), DNA barcoding is based on sequence diversity in short and standard gene regions to differentiate species. DNA barcoding is a molecular approach using DNA barcode sequences such as rbcL, ITS, trnL-trnF, matK, and psbA-trnH (Meilina *et al*., 2024). DNA barcode markers for plant groups are two chloroplast coding sites, part of the genes *rbcL* and *mat*K (De Vere *et al*., 2015). The *mat*K gene is more widely used in various studies than the *rbcL* gene because its level of accuracy is more specific at the species level (Kalangi *et al*., 2014). More research on DNA barcoding of *Castanopsis argentea* must be carried out and developed because more genetic data still needs to be collected at NCBI. This research aims to explore the genetic data of *Castanopsis argentea* in NCBI Genbank, analyzing genetic variation and kinship relationships in silico.

MATERIALS AND METHODS

Sample collection

DNA sequences were collected from the GenBank database (NCBI) using a nucleotide-based search feature with the species name and gene of interest specified (e.g., *Castanopsis argentea mat*K). The DNA sequences obtained were then selected based on inclusion and exclusion criteria. The sequence data collected includes the accession code, nucleotide length, and whether the sequence is partial or complete. This data file is then saved in a Microsoft Excel file. Nucleotide base sequence information was saved in a Notepad file for further analysis to determine genes or regions to identify *C.argentea* via an in silico approach based on the molecular marker *mat*K from a bioinformatics database (Wathon *et al*., 2023).

Sequence alignment and genetic variation analysis

All DNA sequences stored in Notepad files were collected from NCBI and aligned using MEGA 11 software with the ClustalW method (Hall, B.G., 2013; Tamura *et al*., 2021). This is done to determine the similarities and differences in nucleotide base sequences between these sequences and to identify potential barcode sequences. The barcode potential sequence is different and unique compared to others. Genetic variation analysis was done using the website-based application MULTALIN (Multiple Sequence Alignment by Florence Corpet) via the link http://multalin.toulouse.inra.fr/multalin/.

Construction of phylogenetic trees and genetic distances

The aligned DNA sequence data was then used to build a phylogenetic tree using MEGA 11 software with the Maximum Parsimony method and Subtree-Pruning-Regrafting (SPR) algorithm with a bootstrap value 1000x. Next, genetic distance analysis was carried out using the complete deletion method with the p-distance model in the MEGA 11 application (Felsenstein, 1985; Nei & Kumar, 2000; Tamura *et al*., 2021).

RESULTS

Percent Identity of *Castanopsis argentea*

The selection of comparative sequences was carried out on the NCBI GenBank website with the keyword *Castanopsis argentea mat*K. The sample that has the highest query cover and percent identity values was selected and downloaded. Ten species with the highest percent identity were selected for alignment and phylogenetic tree construction. The highest percent identity was chosen because the higher the percent identity, the more significant the homology (similarity) value. The list of sequences used in this study using the *mat*K DNA barcode is presented in Table 1.

Table 1. Top 10 Percent Identity of *Castanopsis argentea* sequences based on DNA barcode *mat*K (*10 Persen Identitas Teratas dari sekuens Castanopsis argentea berdasarkan barcode DNA matK*).

No.	Scientific Name (*Nama ilmiah*)	Query Cover	E-Value (*Nilai E*)	Percentage of similarity (*Persentase kemiripan*)	Accession No. (*No. aksesi*)
1.	*Castanopsis argentea*	100%	0.0	100.00%	LC736869.1
2.	*Castanopsis fargesii*	99%	0.0	99.26%	EF057133.1
3.	*Castanopsis amabilis*	99%	0.0	99.26%	EF057136.1
4.	*Castanopsis delavayi*	99%	0.0	99.26%	EF057138.1
5.	*Castanopsis sieboldii*	99%	0.0	99.15%	AB060055.1
6.	*Castanopsis uraiana*	99%	0.0	99.00%	EF057135.1
7.	*Quercus aliena*	98%	0.0	98.56%	MG772990.1
8.	*Quercus pubescens*	98%	0.0	98.56%	HE966980.1
9.	*Chrysolepsis chrysophylla*	98%	0.0	97.56%	KF419004.1
10.	*Lithocarpus glaber*	98%	0.0	97.40%	OM021804.1

Phylogenetic Tree Reconstruction

A total of 10 sequences that were obtained from NCBI were then processed to reconstruct a phylogenetic tree. The results of the phylogenetic tree reconstruction show that the *mat*K region is divided into two ingroup clades and one outgroup. *Castanopsis argentea* has a close relationship with *Castanopsis fargesii, Castanopsis amabilis, Castanopsis delavanyi, Castanopsis sieboldii, and Castanopsis uraiana*.

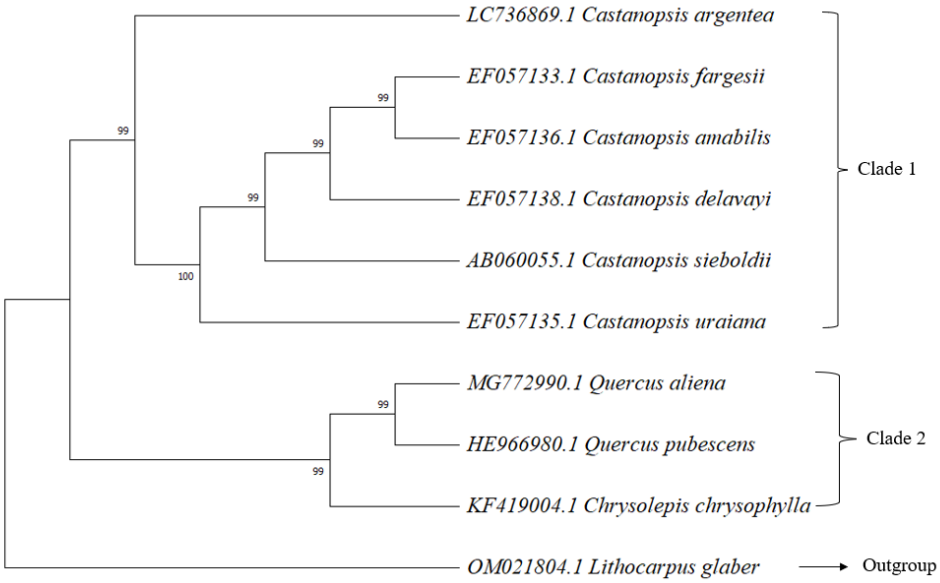

Figure 1. Reconstruction of the phylogenetic tree *Castanopsis argentea* based on the *mat*K DNA barcode using the Maximum Parsimony method and Subtree-Pruning-Regrafting (SPR) algorithm with a bootstrap value of 1000x (*Rekonstruksi pohon filogenetik Castanopsis argentea berdasarkan barcode DNA matK menggunakan metode Maximum Parsimony dan algoritma Subtree-Pruning-Regrafting (SPR) dengan nilai bootstrap 1000x*).

Genetic Distances of *Castanopsis argentea* and Comparative Sequences

Phylogenetic analysis is strengthened by the calculated value of genetic distance (pairwise distance). Genetic distance analysis was carried out to see kinship relationships. Genetic distance is a measure of genetic differences between species or between populations and species. Genetic distance analysis based on *mat*K DNA barcode is presented in Table 2.

Table 2. The genetic distance of *Castanopsis argentea* and comparison sequence using the *mat*K gene (*Jarak genetik Castanopsis argentea dan urutan perbandingan menggunakan gen matK*).

No	Species (*Spesies*)	1	2	3	4	5	6	7	8	9
1	*Castanopsis argentea*									
2	*Castanopsis fargesii*	0,001								
3	*Castanopsis amabilis*	0,001	0,000							
4	*Castanopsis delavanyi*	0,001	0,000	0,000						
5	*Castanopsis sielboldii*	0,001	0,000	0,000	0,000					
6	*Castanopsis uraiana*	0,001	0,000	0,000	0,000	0,000				
7	*Quercus Aliena*	0,003	0,003	0,003	0,003	0,003	0,000			
8	*Quercus pubescens*	0,003	0,003	0,003	0,003	0,003	0,000	0,000		
9	*Chrysolepsis chrysophylla*	0,003	0,003	0,003	0,003	0,003	0,001	0,001	0,001	
10	*Lithocarpus glaber*	0,004	0,004	0,004	0,004	0,004	0,004	0,001	0,001	0,001

Analysis of Genetic Variation of *Castanopsis argentea*

Analysis of genetic variation in *Castanopsis argentea* is presented in Figure 2. The length of the entire *Castanopsis argentea mat*K sequence is 678 bp. Based on this image, it shows that the *mat*K region in *Castanopsis argentea* is highly conserved because it only has three genetic variations, namely N282T, C285T, and C422A, when compared with comparative sequences. The genetic variation of *Castanopsis argentea* is presented in Figure 2.

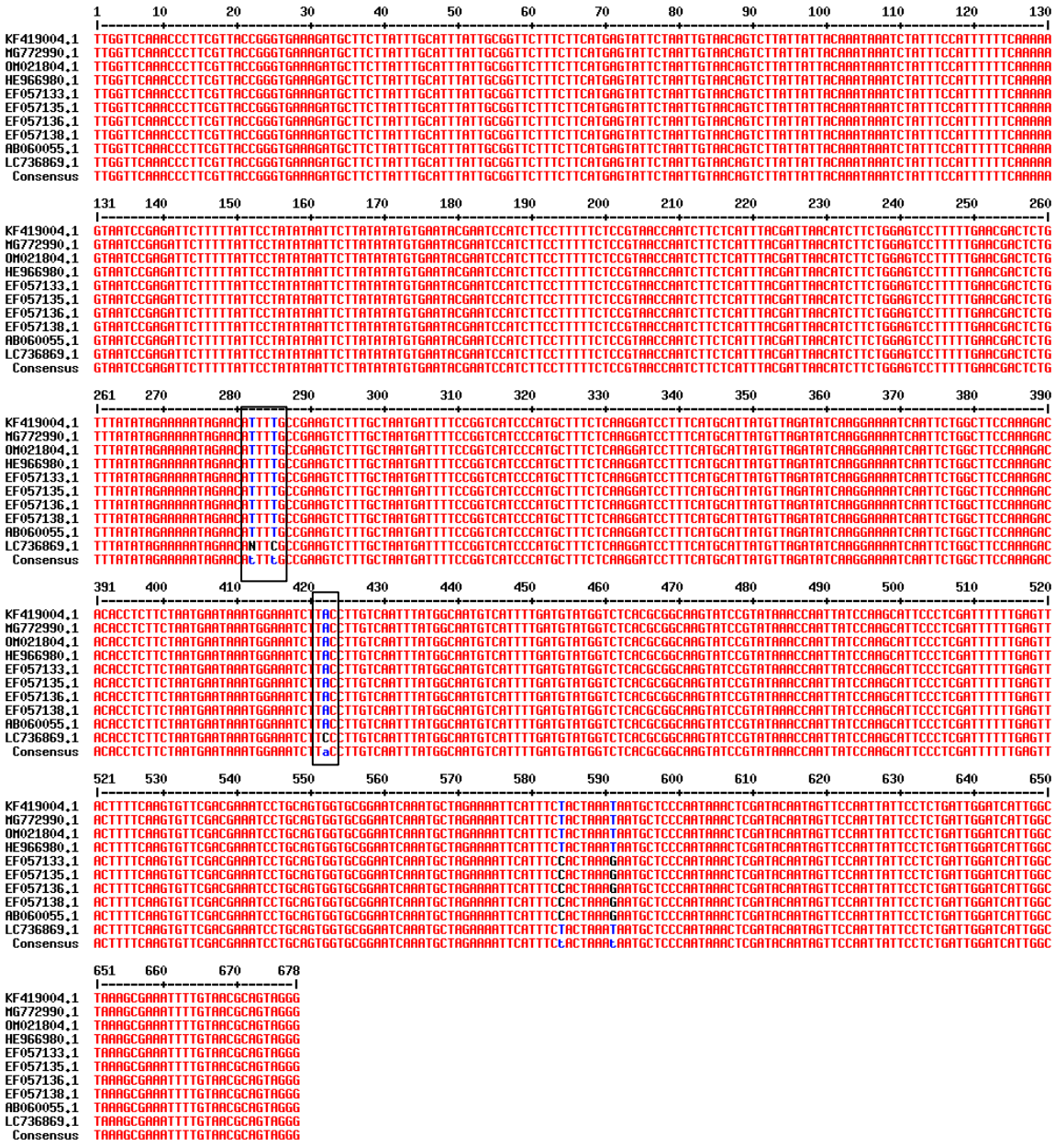

Figure 2. Genetic variation of *Castanopsis argentea* and comparative sequences based on the *mat*K gene analyzed using the MULTALIN (Multiple Sequence Alignment by Florence Corpet) Website (*Variasi genetik Castanopsis argentea dan urutan komparatif berdasarkan gen matK yang dianalisis menggunakan Situs Web MULTALIN (Multiple Sequence Alignment by Florence Corpet*).

		T(U)	C	A	G	Total
KF419004.1	Chrysolepis chrysophylla	36.9	19.6	28.8	14.7	678
LC736869.1	Castanopsis argentea	36.6	19.9	28.7	14.8	677
EF057133.1	Castanopsis fargesii	36.6	19.8	28.8	14.9	678
MG772990.1	Quercus aliena	36.9	19.6	28.8	14.7	678
EF057135.1	Castanopsis uraiana	36.6	19.8	28.8	14.9	678
OM021804.1	Lithocarpus glaber	36.9	19.6	28.8	14.7	678
HE966980.1	Quercus pubescens	36.9	19.6	28.8	14.7	678
AB060055.1	Castanopsis sieboldii	36.6	19.8	28.8	14.9	678
EF057138.1	Castanopsis delavayi	36.6	19.8	28.8	14.9	678
EF057136.1	Castanopsis amabilis	36.6	19.8	28.8	14.9	678
Avg.		36.7	19.7	28.8	14.8	677.

Figure 3. Nucleotide composition of *Castanopsis argentea* and comparative sequence using the *mat*K gene (*Komposisi nukleotida Castanopsis argentea dan urutan komparatif menggunakan gen matK*).

Nucleotide Composition of *Castanopsis argentea* using *mat*K Gene

The sequence is characterized to see the composition of the nucleotides. Nucleotides consist of several nitrogen bases, including adenine (A), guanine (G), thymine/uracil (T/U), cytosine (C). The composition of the nucleotide sequence of *Castanoposis argentea* consists of nucleotide bases with different percentages. The sequence composition of *Castanopsis argentea* can be seen in Figure 3.

DISCUSSION

The selection of ten sequences from NCBI Genbank in Table 1 is based on a high query cover value, namely 100%, for all sequences, the expectation value is low, namely zero (0.0); and a high percent identity value, namely in the range 100% to 99.56%. The query cover value is the percentage of nucleotide lengths that align with the database contained in BLAST. A good guess value is zero (0), or in the BLAST table, a value of 0.0; a guess value of zero indicates a very significant sequence alignment. This is by the statement by Frederick *et al*. (2003) said that the estimated value is significant if it reaches <0.05. Percent identity shows the highest percentage of sequence compatibility with the same sequence subject (Isda & Sofiyanti, 2019).

Reconstruction of the phylogenetic tree in Figure 1 was carried out using the Maximum Parsimony and Subtree-Pruning-Regrafting (SPR) algorithm with a bootstrap value of 1000x (Felsenstein, 1985). The Maximum Parsimony statistical method selects trees with the fewest evolutionary changes or the shortest overall branch lengths. This method is an excellent choice for phylogenetic analysis because the tree with the fewest substitutions explains the differences among the taxa studied. The Pruning-Regrafting (SPR) algorithm allows for the creation of better phylogenetic trees by avoiding non-informative sites so that the search for informative sites using the Maximum Parsimony statistical method can run more efficiently (Hordijk & Gascuel, 2005). These bootstrap replications were used to test the validity of the phylogenetic tree topology (Hoang *et al*., 2018).

The results of the phylogenetic tree reconstruction show that the *mat*K region is divided into two ingroup clades and one outgroup. *Castanopsis argentea* has a close relationship with *Castanopsis fargesii*, *Castanopsis amabilis*, *Castanopsis delavanyi*, *Castanopsis sieboldii*, and *Castanopsis uraiana* which are located in clade 1. This is because they come from the same genus, *Castanopsis*, and have many similarities in nucleotide sequences. Clade 2 consists of *Quercus aliena, Quercus pubescens*, and *Chrysolepsis chrysophylla*. The outgroup on the phylogenetic tree is *Lithocarpus glaber*. This shows that the genus *Castanopsis* has a close kinship with the genus *Quercus* and *Chrysolepsis*. The genus *Castanopsis* has a relatively distant kinship with the genus *Lithocarpus* compared to other comparative sequences shown from the phylogenetic tree line, with the furthest branch length located in the outgroup. The phylogenetic tree also shows high bootstrap values, ranging from 99-100 for all branches. This indicates that this species has a high level of confidence in the topology of the tree reconstruction results. The bootstrap value is a value for the confidence of a branch. The smaller the bootstrap value, the lower the level of confidence in the topology of the tree reconstruction results (Syahputra *et al*., 2017). So, a good bootstrap value can be close to 100.

In studying genetic variation and differentiation between populations, genetic distance can be calculated from the number of polymorphic base differences at a gene locus for each population based on the DNA sequence (Cavalli-Sforza *et al*., 1997). The lower the genetic distance value for a species, the closer the relationship is. Table 2 shows the *mat*K region with the lowest genetic distance of 0.000 between *Castanopsis fargesii, Castanopsis amabilis, Castanopsis delavanyi, Castanopsis sieboldii, Castanopsis uraiana, Quercus aliena,* and *Quercus pubescens*. The phylogenetic tree results show that these sequences are grouped in the same clade and indicate that the species are closely related. Based on this table, it is also known that the genetic distance between *Castanopsis argentea* and *Castanopsis fargesii, Castanopsis amabilis, Castanopsis delavanyi, Castanopsis sieboldii, Castanopsis uraiana* is 0.001, which means the species are closely related. This is in line with the results of the phylogenetic tree, which shows that the two species are in the same clade.

The length of the entire *Castanopsis argentea mat*K sequence is 677 bp. Figure 2 shows that the *mat*K region in *Castanopsis argentea* is highly conserved because it only has three genetic variations, namely N282T, C285T, and C422A when compared with comparative sequences. This is by the *mat*K, which is very conserved and has tiny nucleotide variations. This study used *Castanopsis argentea*, which was compared with another genus of *Castanopsis*, *Quercus*, *Chrysolepsis*, and *Lithocarpus* because there is only one accession number in the NCBI Genbank genetic data for *Castanopsis argentea* in the *matK* gene. Therefore, it is necessary to carry out further research so that more genetic data on *Castanopsis argentea* is available in NCBI Genbank so that more in-depth research can be carried out on genetic variations in *Castanopsis argentea*.

DNA Barcode is used for all groups of organisms and is available to aid in understanding, conserving, and using biodiversity. DNA barcode markers for plant groups are two chloroplast coding sites, part of the genes *rbcL* and *mat*K (De Vere *et al*., 2015). The *mat*K gene has a high evolutionary speed, so the *mat*K gene is considered better and more accurate in identifying and differentiating a species (Barthet, 2006; Kolondam *et al*., 2012). The *matK* gene encodes the maturase enzyme part of the K subunit found in chloroplasts in plants (Kalangi *et al*., 2014). This region of the nucleotide sequence of the *mat*K gene can produce approximately 1500 bp (base pairs) (Soltis *et al*., 1998). In this study, the *mat*K gene can group species from the same genus and separate species from different genus. This is in line with the results of the phylogenetic tree, which shows that members of the same genus successfully grouped into one clade.

Figure 3 shows that *Castanopsis argentea* and the comparative sequence in the *mat*K gene have an average nucleotide composition with the highest percentage of thymine (T) and adenine (A), namely 36.7% thymine and 28.8% adenine. This is to research by Huang *et al*. (2005), which stated that the chloroplast spacer region, such as *mat*K, has a higher thymine (T) and adenine (A) nucleotide composition. The nucleotide composition in a gene fragment is correlated with the composition of the genetic code contained in the gene fragment. This composition is also related to the composition of the amino acids coded by the gene, so changes in nucleotide composition cause changes in the genetic code. However, these changes do not always cause changes in the coded amino acids because there are amino acids that are coded for more than one genetic code (Suriana & Nasaruddin, 2016).

CONCLUSION

Based on this research, the phylogenetic results show that the results of the phylogenetic tree reconstruction show that the *mat*K region is divided into two ingroup clades and one outgroup. *Castanopsis argentea* has a close relationship with *Castanopsis fargesii, Castanopsis amabilis, Castanopsis delavanyi, Castanopsis sieboldii*, and *Castanopsis uraiana* which are located in clade 1. This is because they come from the same genus *Castanopsis*, and have many similarities in nucleotide sequences. The genetic distance between *Castanopsis argentea* and *Castanopsis fargesii, Castanopsis amabilis, Castanopsis delavanyi, Castanopsis sieboldii, Castanopsis uraiana* is 0.001, which means the two species are closely related. This is in line with the results of the phylogenetic tree, which shows that the species are in the same clade. The *mat*K region in *Castanopsis argentea* is highly conserved because it only has three genetic variations. *Castanopsis argentea* and the comparative sequence in the *mat*K gene have an average nucleotide composition with the highest percentage of thymine (T) and adenine (A), namely 36.7% thymine and 28.8% adenine.

AUTHOR CONTRIBUTIONS

All authors have contributed to this paper. SC and SZ: collecting research data and drafting the article; SC and SZ: revise manuscripts and final revision of manuscript; SC and SZ: create research concepts and revise manuscripts.

REFERENCES

Aribi, M.M. 2024. Plant Gene Banks: Conservation of Genetic Resources. *In Sustainable Utilization and Conservation of Plant Genetic Diversity*. pp.753-775. Singapore: Springer Nature Singapore.

Barstow, M., Kartawinata, K. 2018. *Castanopsis argentea*. The IUCN Red List of Threatened Species 2018: e.T62004506A62004510. Accessible online at https://www.iucnredlist.org/species/62004506/62004510

Barthet, M.M. 2006. *Expression and function of the chloroplastencoded gene matK. Virginia Polytechnic Institute and State University*. Blacksburg.

Cavalli-Sforza, L.L. 1997. Genes, peoples, and languages. *Proceedings of the National Academy of Sciences*, 94(15), pp.7719-7724.

Chen, L., Li, X.W., Li, J.Q. 2013. Taxonomic notes on *Castanopsis* (Fagaceae, Castaneoideae) from China. *Phytotaxa*, 146(2), pp.50-60.

Cheuk, M.L., Fischer, G.A. 2021. The impact of climate change on the distribution of *Castanopsis* (Fagaceae) species in South China and Indo-China region. *Global Ecol Conserv*, 26, p.e01388. DOI: 10.1016/j.gecco.2020.e01388

Chika, S., Ismaini, L., Armanda, D. T. (2022). Explant Sterilization Technique *Castanopsis argentea* (Blume) A. DC. with the addition of Ascorbic Acid and Sodium Hypochlorite (NaOCl) In Vitro. *Berkala Ilmiah Biologi*, 13(2), pp.32-41.

Chika, S., Meilina, T.D., Arfan, M.R., Febriana, A., Mukaromah, A.S. 2024. Genetic Variation of Wax Apple (*Syzygium samarangense* (Blume) Merr. & LM Perry) Cultivar using Internal Transcribed Spacer 2 (ITS2). *Floribunda*, 7(4), pp.153-165.

Chika, S., Zahro, S. 2024. An In Silico Approach for Evaluation of ITS, rbcL, and psbA-trnH for DNA Barcoding of *Eugenia* spp. *Biosfer: Jurnal Tadris Biologi*, 15(1), pp.153-167.

Cowan, R.S., Chase, M.W., Kress, W.J., Savolainen, V. 2006. 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. *Taxon*, 55(3), pp.611-616.

De Vere, N., Rich, T.C.G., Trinder, S.A., Long, C. 2015. DNA Barcoding for Plants In Batley, J. (Ed.). *Plant Genotyping: Methods and Protocols*. Methods in Molecular Biology Vol. 1245. Springer, New York.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39, pp.783-791.

Frederick, M.A., Williamson, K.E., Garcia, M.F., Ferretti, M.B., McCarthy, R.L., Donahue, G., Luzete Monteiro, E., Takenaka, N., Reynaga, J., Kadoch, C., Zaret, K.S. 2023. A pioneer factor locally opens compacted chromatin to enable targeted ATP-dependent nucleosome remodeling. *Nat Struct Mol Biol*, 30, pp.31–37. DOI: 10.1038/s41594-022-00886-5.

Hall, B.G. 2013. Building Phylogenetic Trees from Molecular Data with MEGA. *Molecular Biology and Evolution*, 30(5), pp.1229–1235. doi: 10.1093/molbev/mst012

Hilwan, I., Irfani, E. (2018). Pola Penyebaran dan Regenerasi Jenis Saninten (*Castanopsis argentea* Blume) di Resort Selabintana, Taman Nasional Gunung Gede Pangrango. *Silvikultur Trop-J Trop Silvic Sci Technol*, 9(1), p.5359.

Hoang, D.T., Vinh, L.S., Flouri, T., Stamatakis, A., Haeseler, A.V., Minh, B.Q. 2018. MPBoot: Fast Phylogenetic Maximum Parsimony Tree Inference and Bootstrap Approximation. *BMC Evolutionary Biology*,18 pp.1-11.

Hordijk, W., Gascuel, O. 2005. Improving The Efficiency of SPR Moves in Phylogenetic Tree Search Based on Maximum Likelihood. *Bioinformatics*, 21(24), pp.4338-4347.

Huang, S.S., Hwang, S.Y., Lin, T.P. 2002. Spatial pattern of chloroplast DNA variation of *Cyclobalanopsis glauca* in Taiwan and East Asia. *Molecular Ecology*, 11(11), pp.2349-2358.

Isda, N., Sofiyanti, M.N. 2019. Paku Kawat *Lycopodiella cernua* (L.) Pic. Serm. (Lycopdiaceae-Lycopodiales) dari Provinsi Riau – Kajian Morfologi dan Sekuen DNA berdasarkan Primer *rbcL*. *Jurnal Biologi Universitas Andalas*. 7(1), pp.43-50.

Kalangi, C., Kamu, V.S., Kumaunang, M. 2014. Barcode DNA tanaman leilem (*Clerodendrum minahassae* L.) berdasarkan gen *matK*. *Jurnal MIPA*, 3(2), pp.108-112.

Kolondam, B.J. 2015. Applying *matK* gene for identification of Liliopsida plant species from North Sulawesi through Bold System. *International Journal of Applied Biology and Pharmaceutical Technology*. 6(2), pp.242-245

Mansur. 2013. *Silvikultur untuk Reklamasi Lahan Bekas Tambang*. Seameo Biotrop.

Meilina, T.D., Chika, S., Hariri, M.R., Febriana, A., Mukaromah, A.S. 2024. psbA-trnH Intergenic Spacer profile of Wax Apple (*Syzygium samarangense* (Blume) Merr. & LM Perry) Cultivars. *Jurnal Ilmu Pertanian Indonesia*, 29(4), pp.548−553

Nei, M., Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Pang, X., Liu, C., Shi, L., Liu, R., Liang, D., Li, H., Cherny, S.S., Chen, S. 2012. Utility of the *trnH-psbA* Intergenic Spacer Region and Its Combinations as Plant DNA Barcodes: A Meta-Analysis. *PLoS ONE,* 7(11). DOI: 10.1371/journal.pone.0048833

Penidda, E.I., Jumari, J., Baskoro, K., Sahroni, D., Penidiyanti, P., Sujarwo, W. 2024. The Local Community Understanding of the Factors Influencing the Decline of Saninten (*Castanopsis argentea* (Blume) A. DC.) and Tungurut (*Castanopsis tungurrut* (Blume) A. DC.) in the Buffer Villages of the Gunung Halimun-Salak National Park. *Reinwardtia*, 23(1), pp.15-32.

Soltis, P. S., D. E. Soltis, and J. J. Doyle. 1998. *Molecular Systematics of Plants.* International Thomson Publishing, New York.

Suriana, N., Nasaruddin, N. 2016. Karakteristik Parsial Gen Sitokrom-C Oksidase Subunit-I Katak Pohon Suaka Marga Satwa Tanjung Peropa Moramo, Sulawesi Tenggara. *Jurnal Veteriner Desember*, 17(4), pp.517-523.

Syahputra, B., Bakti, D., Pinem, M.I., Prasetyo, A.E. 2017. Karakterisasi Molekuler *Elaedobius kamerunicus* Faust. (Coleoptera: Curculionidae) Asal Sumatera Utara Menggunakan Sekuen DNA Molekuler. *Jurnal Agroekoteknologi*, 5(3), pp.1–23.

Tamura K., Stecher G., Kumar S. 2021. MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* https://doi.org/10.1093/molbev/msab120

Wathon, S., Astikaningrum, D., Ardyah, N.P.C., Oktarianti, R., Senjarini, K. 2023. In silico exploration of the potential barcode DNA in *Anopheles* sp., a malarian vector from Indonesia. *Jurnal Biolokus*, 6(1), p.96. https://doi.org/10.30821/biolokus.v6i1.1749

Whitmore, T.C. 1986. *Tree Flora of Indonesia Check List for Sumatra* Compiled by Sidiyasa, K., Sutisna, U, Marfuah, Kalima, K., Whitmore, T.C. Ministry of Forestry. Agency for Forestry Research and Development. Forest Research & Development Centre.

Wibowo, C. 2006. *Hubungan antara Keberadaan Saninten (Castanopsis argenta Blume) dengan Beberapa Sifat Tanah: Kasus di Taman Nasional Gunung Gede Pangrango, Jawa Barat*. Disertasi. IPB University.