Asian Journal of Plant Sciences

ISSN 1682-3974

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS **Asian Journal of Plant Sciences**

ISSN 1682-3974
DOI: 10.3923/ajps.2023.444.451

Research Article
Genetic Diversity in the *matK* Gene of *Dimocarpus longan* Varieties in the Mekong Delta

Nguyen Pham Anh Thi, Tran Thanh Tam and Do Tan Khang

Institute of Food and Biotechnology, Can Tho University, Ninh Kieu, Can Tho, Vietnam

Abstract

Background and Objective: *Dimocarpus longan* has been known widely due to its many nutritional values for human health. Therefore, this study aimed to evaluate genetic diversity in the *matK* gene of *D. longan* varieties in the Mekong Delta. **Materials and Methods:** Eleven *D. longan* varieties were extracted by modified CTAB procedure and amplified *matK* gene by *matK*-4600/ *trnK*-2R and tamF/tamR primers. The DNA sequences were treated by BioEdit 7.0, MEGA 11, Geneious 7.0, Denovo, ClustalW, NCBI/BLAST tool and compared with *D. longan*-NC_037447.1 on NCBI. The phylogenetic tree was analyzed by the maximum-likelihood method. **Results:** The results successfully amplified the *matK* gene with a size of ~1500 bp. The longan varieties had 3 SNPs that were respectively updated in GenBank (NCBI) with accession numbers including OP819669, OP819670, OP819671, OP819672, OP819673, OP819674 and OP819676. Compared with *D. longan*-NC_037447.1, all sequences showed similarity ranging from 99-100%. The results of the phylogenetic tree analysis showed that the bootstrap values were $\geq$75% in all branches. **Conclusion:** The sequences of the *matK* gene were relatively phylogenetically distinguishable between the longan varieties of the *Dimocarpus* genus.

Key words: *Dimocarpus longan*, DNA sequencing, longan varieties, SNP, *matK* gene

Citation: Thi, N.P.A., T.T. Tam and D.T. Khang, 2023. Genetic diversity in the *matK* gene of *Dimocarpus longan* varieties in the Mekong Delta. Asian J. Plant Sci., 22: 444-451.

Corresponding Author: Nguyen Pham Anh Thi, Institute of Food and Biotechnology, Can Tho University, Ninh Kieu, Can Tho, Vietnam

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Dimocarpus longan belongs to the *Dimocarpus* genus, Sapindaceae family, Sapindales order came from Southern China, then introduced to Australia in the mid-1800s, Thailand in the 1800s, Hawaii and Florida in the 1900s. Recently, this species has become one of the main export items in Vietnam. However, the emergence of different varieties of longan with different delicious flavors makes new challenges for export. Therefore, through DNA sequencing, it is possible to construct specific DNA barcodes that can be used to identify the source of different cultivars[1-3].

The DNA barcodes are essential to respond the demand of new development directions of today's era. The DNA barcodes can be used as a method of identification using DNA sequences from specific gene regions of the genome to aid in new species discovery or subspecies and species identification[4,5]. One of the gene regions that has been proven to be highly effective in identifying plant species in many domestic and international studies is the *matK* gene[6-10]. The highly conserved chloroplast Maturase K (*matK*) gene in plant systems is involved in group II intron splicing, ~1500 bp in size, located in the noncoding region of the *trnK* gene[11,12]. Identifying the origin of longan varieties through barcoded DNA can help to increase their commercial value and optimize the export of longan to potential consumption markets. Therefore, the objective of this study was to decode the gene sequence in the *matK* gene and contribute to demonstrating the genetic diversity of *Dimocarpus longan* varieties in the Mekong Delta.

MATERIALS AND METHODS

Study duration and location: The study was implemented in the Molecular Biology Laboratory of the Institute of Food and Biotechnology at Can Tho University from June, 2022 to November, 2022.

Plant sampling: Eleven different varieties of *Dimocarpus longan* were collected at the longan garden (Phong Nam commune, Ke Sach District, Soc Trang Province) and the longan cooperative (Thoi An ward, O Mon District, Can Tho City) in June, 2022. *Dimocarpus longan* cultivars collected in this study including "Thach kiet" longan, "Thanh" longan, "Idol" longan, "Long" longan, "Xuong" longan in Can Tho City, "Purple" longan, "Xuong com vang" longan, "Thach kiet" longan, "Long" longan, "Da bo" longan, "Xuong vang" longan in Soc Trang Province (Fig. 1a-k).

Amplification of DNA markers: The DNA extraction from leaves of longan fruits was performed according to the modified CTAB procedure. The obtained DNA was eluted with 50 µL of 0.1X TE solution. The solution containing DNA was stored in the freezer (-20°C) for the next procedure. The extracted DNA was treated by a Nanodrop 2000°C spectrophotometer. Then the quality of DNA was checked by 1% agarose electrophores.

To obtain the complete DNA sequence of the *matK* gene with a size of ~1500 bp, two pairs of primers (*matK*-4600/*trnK*-2R and tamF/tamR) were used. Which, tamF/tamR primer pair was designed in this study (Fig. 2). The tamF/tamR primer pair was synthesized at Phu Sa Biochem One Member Co., Ltd. (Vo Nguyen Giap Street, Phu Thu Ward, Cai Rang District, Can Tho City).

Each PCR reaction was performed in a volume of 50 µL containing 25 µL of ddH_2O, 20 µL of Mastermix, 3 µL of DNA template, 1 µL of forward primer (20 pmol μL^{-1}) and 1 µL of reverse primer (20 pmol μL^{-1}) (Table 1). The concentration of the DNA template was given in Table 2.

The PCR thermal cycles to amplify DNA barcodes was shown in Table 3. The quality of the PCR product was checked by 2% agarose electrophores to observe the specific bands. The selected PCR products sequenced their DNA sequences at The Institute of DNA Technology and Genetic Analysis-Genlab (Yen Hoa ward, Cau Giay District, Hanoi City).

Statistical analysis: The obtained DNA sequences were corrected by BioEdit 7.0 Software to check the peak of nucleotides. The sequences of PCR products after processing were paired by using the Denovo alignment tool. The *matK* sequences were analyzed by using BioEdit software, aligned with the ClustalW tool and compared on the GenBank by using the NCBI/BLAST tool integrated into the Geneious 7.0 analysis software[14]. The SNPs (Single nucleotide Polymorphisms) and amino acids (dN/dS) were calculated by using the MEGA7 software[15]. The access numbers of the *matK* gene having the different sequences were registered on the GenBank (NCBI).

The MEGA11 software was applied (The Molecular Evolution Genetics Analysis) with a bootstrap coefficient of 1000 to determine the distance between sequences when comparing nucleotide sequences in the *matK* gene of longan varieties. The phylogenetic tree was constructed by using the maximum-likelihood method in order to show the genetic relationship between different cultivars based on the sequences of the *matK* gene.

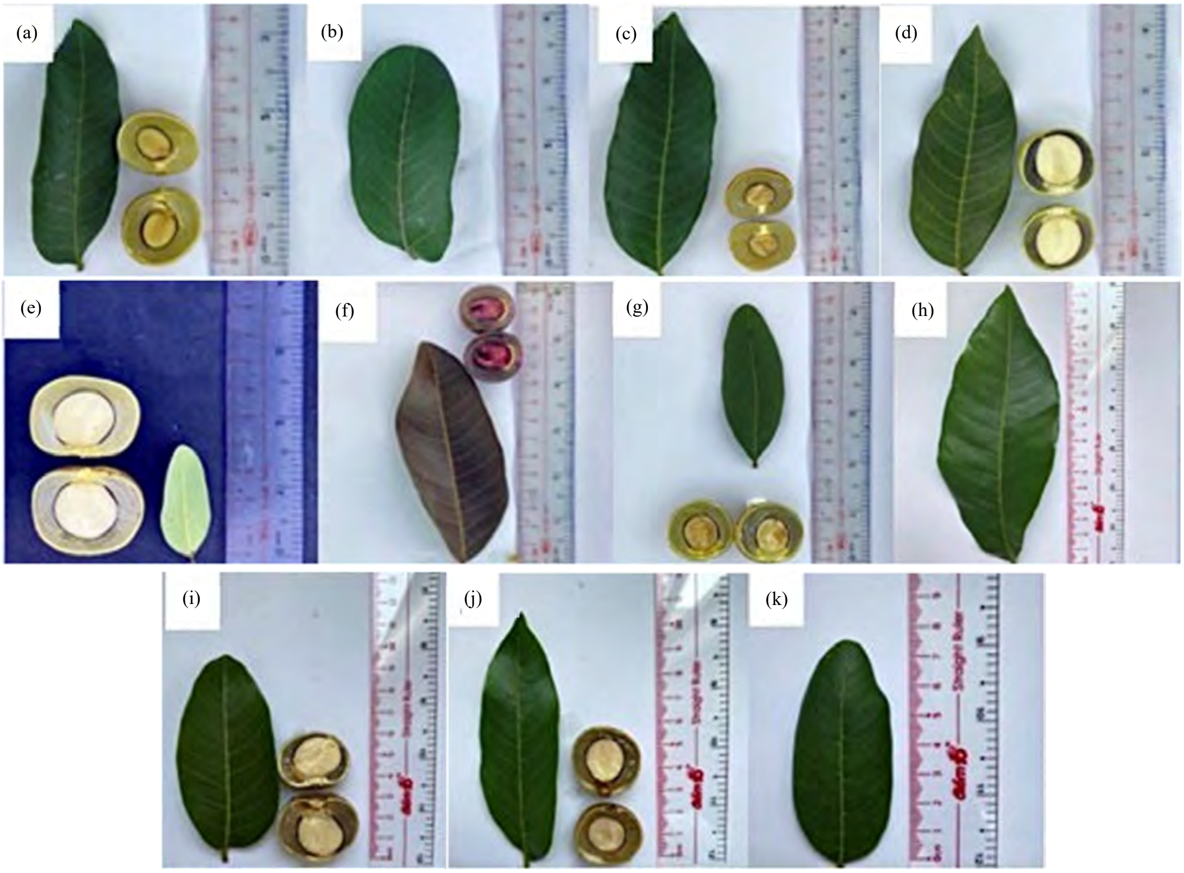

Fig. 1(a-k): Some fruits and leaf morphology of 11 collected longan varieties, (a) "Thach kiet" longan (Can Tho), (b) "Thanh" longan (Can Tho), (c) "Idol" longan (Can Tho), (d) "Long" longan (Can Tho), (e) "Xuong" longan (Can Tho), (f) "Purple" longan (Soc Trang), (g) "Xuong com vang" longan (Soc Trang), (h) "Thach kiet" longan (Soc Trang), (i) "Long" longan (Soc Trang), (j) "Da bo" longan (Soc Trang) and (k)"Xuong vang" longan (Soc Trang)

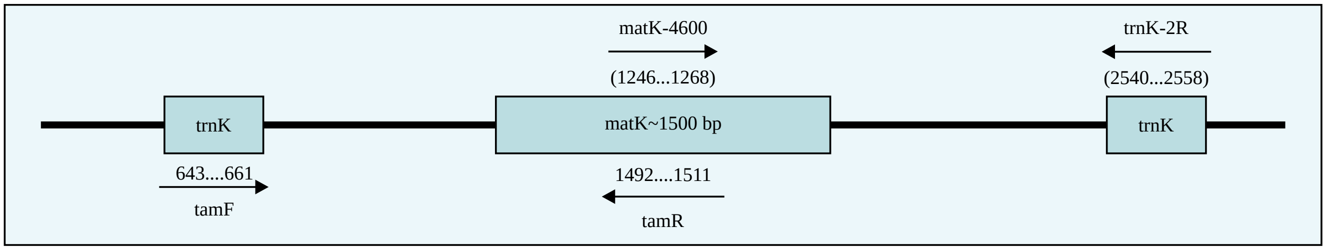

Fig. 2: Primer location diagram

Table 1: Primer sequences for DNA barcode loci

Primers	Nucleotide sequence (5'-3')	Temp. (°C)	References
matK-4600	GAAATCTTGGTTCAAACCTTTCG	53	Buerki *et al.*[13]
trnK-2R	AACTAGTCGGATGGAGTAG	53	
tamF	CGTTTCCAAGGTATCCGCT	55	
tamR	TGGATTCGCTCAAGGAGGAC	55	

Table 2: DNA concentration of longan varieties was performed through PCR amplification

No.	Name	DNA concentration (ng μL^{-1})
1	"Thach kiet" longan (Can Tho)	18.0
2	"Thanh" longan (Can Tho)	30.7
3	"Idol" longan (Can Tho)	14.8
4	"Long" longan (Can Tho)	18.7
5	"Xuong" longan (Can Tho)	84.4
6	"Purple" longan (Soc Trang)	18.1
7	"Xuong com vang" longan (Soc Trang)	67.0
8	"Thach kiet" longan (Soc Trang)	56.4
9	"Long" longan (Soc Trang)	98.6
10	"Da bo" longan (Soc Trang)	128.2
11	"Xuong vang" longan (Soc Trang)	90.5

Table 3: Thermal cycles for amplification of *matK* gene by two primer pairs

		Thermal cycles				
Primer	Initial denaturation	Denaturation	Annealing	Extension	Final extension	Storage
*matK*4600/*trnK*-2R		35 cycles				20°C
	94°C	94°C	53°C	72°C	72°C	
	4 min	40 sec	35 sec	1 min	10 min	
tamF/tamR		35 cycles				
	94°C	94°C	55°C	72°C	72°C	
	4 min	40 sec	30 sec	1 min	10 min	

RESULTS AND DISCUSSION

DNA barcode amplification: Electrophoresis results showed that the bands were approximately 1500 bp in size. Daniell *et al.*[16] and Steane[17] found that the length of the *matK* gene in the chloroplasts of plants, have ~1500 bp in size. Therefore, the amplified DNA band size of ~1500 bp was relatively consistent with the theory to conduct sequencing of the amplified gene region with two pairs of primers (*matK*-4600/*trnK*-2R and tamF/tamR). According to Fig. 3a-b, the PCR products after using *matK*4600/*trnK*-2R were ~1200 and ~1000 bp, respectively. Their sequences are continuously treated by BioEdit software to remove unclear nucleotides at two ends of each sequence. Then those sequences were applied to the Denovo alignment tool to form many consensuses which had the complete *matK* gene of ~1500 bp in size (Table 4). Therefore, the size of the *matK* gene region (~1500 bp) in this research was relatively consistent with the theory presented by many previous studies.

***matK* gene analysis:** Through Fig. 4, the quality of the obtained *matK* gene sequences was checked by Bioedit software and showed that most of the sequences have clear fluorescence signals and nucleotides could be identified in all samples. However, at the two ends of the sequencing results, some unclear signals appeared. Therefore, to avoid leading to inaccurate sequencing results, segments of the sequence at both ends of the obtained sequences would be removed before aligning. The analysis results after aligning consensus sequences of the complete *matK* gene region were shown in Table 4.

The *matK* gene size of *Dimocarpus longan* cultivars ranged from 1524-1526 bp. Sugita *et al.*[18] and Turmel *et al.*[19] also introduced that the *matK* gene size was ~1500 bp. Thus, the figures in this study were appropriate to the above previous research. Some cultivars had the highest nucleotide positions including "Thanh" longan, "Long" longan, "Xuong" longan (Can Tho), "Purple" longan, "Xuong com vang" longan, "Long" longan, "Xuong vang" longan (Soc Trang) with 3 SNPs. The similarities ranged from 99-100% when comparing all sequences in this study with *Dimocarpus longan*-NC_037447.1 on NCBI. Lan *et al.*[14] showed that the similarity level when studying the *matK* gene sequence at 31 different longan varieties in the *Dimocarpus* genus varies from 99-100% so there was a correlation between the results of this study and previous studies. Significantly, the *matK* gene of seven cultivars was updated on NCBI including OP819669, OP819670, OP819671, OP819672, OP819673, OP819674 and OP819676 (Table 5).

The dN/dS ratio, in which, dN is non-synonymous substitutions and dS is synonymous substitutions. The result of the dN/dS ratio was recorded as 1.09. Statistically, the dN/dS ratio was higher than 1.0 could be assumed that the change of nucleotide sequences lead to the change of amino acid sequences. According to the results of amino acid sequence analysis, two positions of amino acid variations at 47 and 132 were discovered after translation in "Thanh" longan, "Long" longan, "Xuong" longan (Can Tho), "Purple" longan, "Xuong com vang" longan, "Long" longan, "Xuong vang" longan (Soc Trang) (Fig. 5).

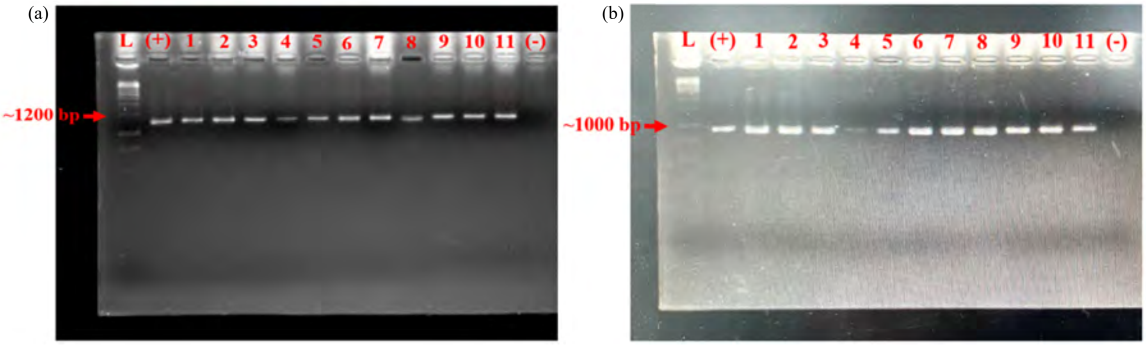

Fig. 3(a-b): Electrophoresis results after amplifying *matK* gene, (a) Amplified by *matK*4600/*trnK*-2R and (b) Amplified by tamF/tamR

L: Ladder 100 bp, (+): Positive control sample, 1: "Thach kiet" longan (Can Tho), 2: "Thanh" longan (Can Tho), 3: "Idol" longan (Can Tho), 4: "Long" longan (Can Tho), 5: "Xuong" longan (Can Tho), 6: "Purple" longan (Soc Trang), 7: "Xuong com vang" longan (Soc Trang), 8: "Thach kiet" longan (Soc Trang), 9: "Long" longan (Soc Trang), 10: "Da bo" longan (Soc Trang), 11: "Xuong vang" longan (Soc Trang) and (-): Negative control sample

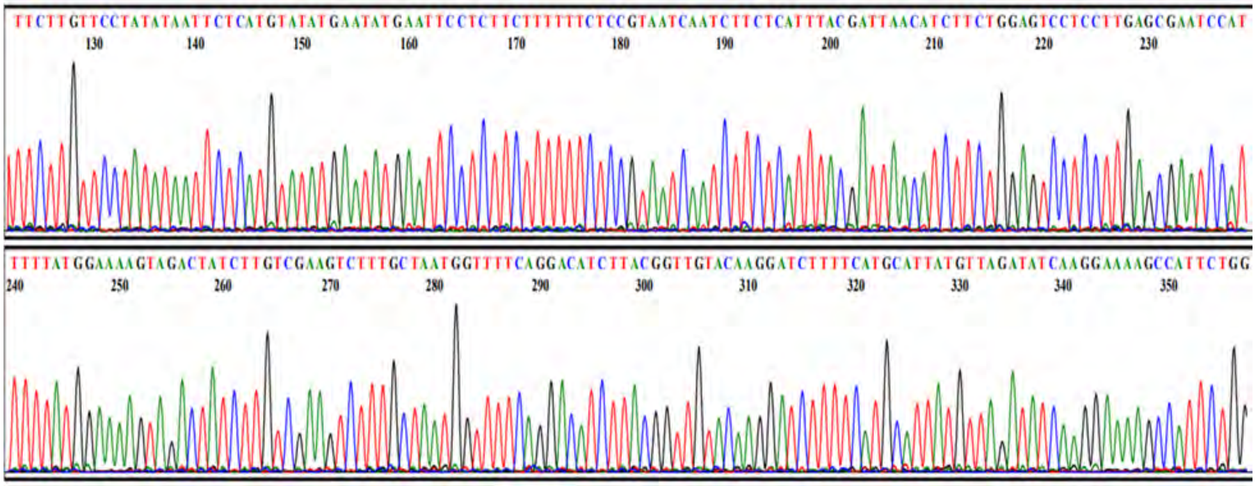

Fig. 4: A fragment of sequencing results for amplified product using *matK*-4600 primer in purple longan (Soc Trang)

```
                  1       10        20        30        40        50        60        70
0matK-Dimoca...   MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENGGYDNKSSSLSGKRLITLLYQQIHL
1-ThachkietCan... MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENGGYDNKSSSLSGKRLITLLYQQIHL
2-ThanhnhanC...   MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENEGYDNKSSSLSGKRLITLLYQQIHL
3-IdolCanTho tr... MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENGGYDNKSSSLSGKRLITLLYQQIHL
4-LongCanTho ...  MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENEGYDNKSSSLSGKRLITLLYQQIHL
5-XuongCanTh...   MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENEGYDNKSSSLSGKRLITLLYQQIHL
6-TimSocTrang ... MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENEGYDNKSSSLSGKRLITLLYQQIHL
7-Xuongcomva...   MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENEGYDNKSSSLSGKRLITLLYQQIHL
8-ThachkietSoc... MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENGGYDNKSSSLSGKRLITLLYQQIHL
9-LongSocTran...  MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENEGYDNKSSSLSGKRLITLLYQQIHL
10-DaboSocTra...  MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENGGYDNKSSSLSGKRLITLLYQQIHL
11-XuongvangS...  MEKFQIYLELDGSQQHNFLYPLLFREYIYALAQDHGLNRSTISLENEGYDNKSSSLSGKRLITLLYQQIHL
                          80        90        100       110       120       130       140
0matK-Dimoca...   SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSIFPFFEDNFSH
1-ThachkietCan... SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSIFPFFEDNFSH
2-ThanhnhanC...   SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSVFPFFEDNFSH
3-IdolCanTho tr... SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSIFPFFEDNFSH
4-LongCanTho ...  SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSVFPFFEDNFSH
5-XuongCanTh...   SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSVFPFFEDNFSH
6-TimSocTrang ... SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSVFPFFEDNFSH
7-Xuongcomva...   SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSVFPFFEDNFSH
8-ThachkietSoc... SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSIFPFFEDNFSH
9-LongSocTran...  SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSVFPFFEDNFSH
10-DaboSocTra...  SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSIFPFFEDNFSH
11-XuongvangS...  SISANDSNPNQFIGHNNNLYSQMISEGFAVIVEIPFSLQLVSFLQGKEMAKSHNFQSIHSVFPFFEDNFSH
```

Fig. 5: Amino acid polymorphisms after translating the nucleotide sequences of the *matK* gene

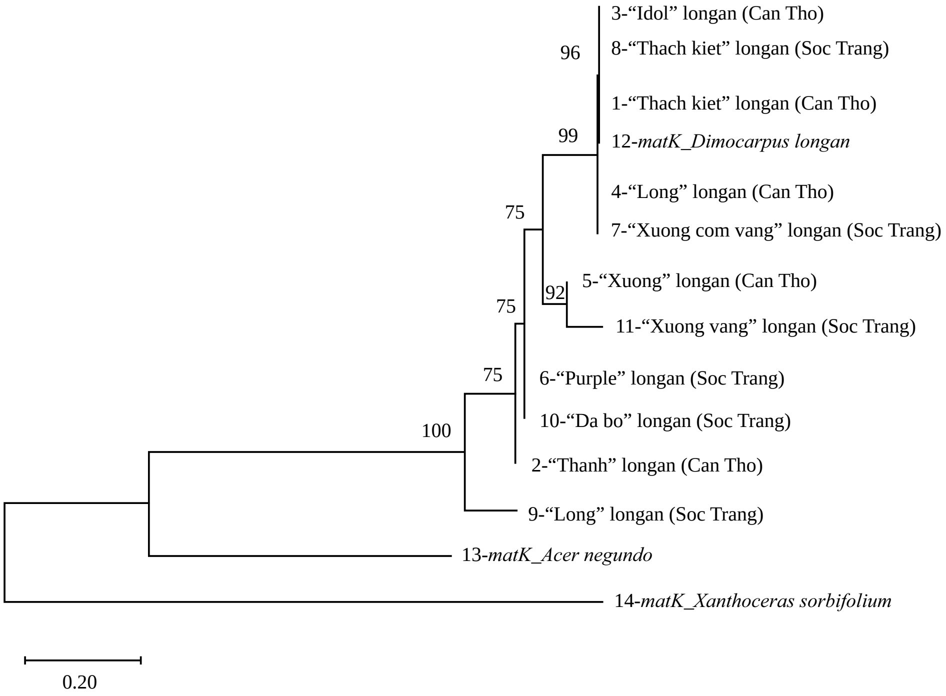

Fig. 6: Phylogenetic tree of *Dimocarpus longan* variations based on the *matK* gene by maximum-likelihood (ML) method

Data at each branch showed the bootstrap values, *Dimocarpus longan*-NC_037447.1: Reference sequence, *Acer negundo*-NC_051957.1: Outgroup sequence and *Xanthoceras sorbifolium*-NC_037448.1: Outgroup sequence

Table 4: Nucleotide variations of complete *matK* gene between *Dimocarpus longan* based on NCBI database and other cultivars

		Nucleotide position			
Cultivar	Size (bp)	140	394	940	SNPs
*matK**	1524	G	A	T	
Consensus 1	1524	.	.	.	0
Consensus 2	1525	A	G	G	3
Consensus 3	1524	.	.	.	0
Consensus 4	1524	A	G	G	3
Consensus 5	1526	A	G	G	3
Consensus 6	1525	A	G	G	3
Consensus 7	1524	A	G	G	3
Consensus 8	1524	.	.	.	0
Consensus 9	1525	A	G	G	3
Consensus 10	1525	.	.	.	0
Consensus 11	1526	A	G	G	3

".": Exhibited a nucleotide similar to that of the first sample, *: *matK-Dimocarpus longan* (Based on database of NCBI), 1: "Thach kiet" longan (Can Tho), 2: "Thanh" longan (Can Tho), 3: "Idol" longan (Can Tho), 4: "Long" longan (Can Tho), 5: "Xuong" longan (Can Tho), 6: "Purple" longan (Soc Trang), 7: "Xuong com vang" longan (Soc Trang), 8: "Thach kiet" longan (Soc Trang), 9: "Long" longan (Soc Trang), 10: "Da bo" longan (Soc Trang) and 11: "Xuong vang" longan (Soc Trang)

The *matK* sequences in some other subjects previously studied on NCBI were used to combine the construction of the phylogenetic tree for this research. Specifically, the *matK* gene sequences in plants of *Acer* genus (*Acer negundo*-NC_051957.1), *Xanthoceras* genus (*Xanthoceras sorbifolium*-NC_037448.1) of the family Sapindaceae were selected as outgroup sequences.

By the method of maximum-likelihood (ML), the *matK* gene sequences of longan varieties in the *Dimocarpus* genus were grouped separately from cultivars in the *Acer* genus and *Xanthoceras* genus in the Sapindaceae family (Fig. 6). For eleven studied longan sequences, some longan varieties such as "Thach kiet" longan (Can Tho), "Idol" longan (Can Tho) and "Thach kiet" longan (Soc Trang) had the most correlation

Table 5: Accession number of longan variations updated on NCBI

No.	Name	Accession number
1	"Thanh" longan (Can Tho)	OP819669
2	"Long" longan (Can Tho)	OP819670
3	"Xuong" longan (Can Tho)	OP819671
4	"Purple" longan (Soc Trang)	OP819672
5	"Xuong com vang" longan (Soc Trang)	OP819673
6	"Long" longan (Soc Trang)	OP819674
7	"Xuong vang" longan (Soc Trang)	OP819676

when they were located in the same clade with *matK* reference sequence on NCBI (*Dimocarpus longan*-NC_037447.1). The above varieties were grouped with "Long" longan (Can Tho) and "Xuong com vang" longan (Soc Trang). "Xuong" longan (Can Tho) and "Xuong vang" longan (Soc Trang) belonged to the same group and had a relatively close relationship with the above longan varieties. Then the varieties of "Purple" longan (Soc Trang) and "Da bo" longan (Soc Trang) were grouped, while "Thanh" longan (Can Tho) and "Long" longan (Soc Trang) were analyzed that they had a more different genetic relationship. Thus, the results of the phylogenetic tree based on the nucleotide sequences of the *matK* gene had taxonomic significance to the species level in many different longan varieties.

Schmitz-Linneweber *et al.*[20] evaluated that the *matK* gene is one of the genes with a high evolutionary rate in the chloroplast genome. Therefore, this gene has been widely applied to genealogical studies at the species level. Yesson *et al.*[21] carried out that the species recognition efficiency of the *matK* gene has also been evaluated in 528 species of the Cactaceae family, of which 75% were native succulents. This group of authors concluded that the *matK* gene sequencing was able to correctly identify 77% of the collected species, these studies have shown that the *matK* gene has been used to differentiate many species effectively. In this study, by the maximum-likelihood (ML) method, the phylogenetic tree was built with the phylogenetic branches and the bootstrap values recorded in all branches were higher than 75%.

The results of this study was reliable to conclude about phylogenetic resources based on the *matK* gene of longan varieties at the species level. It can be applied to marking varieties of longan through barcodes on labels, helping to understand the origin of longan varieties and improving their commercial value when exporting. The result was not clear in nucleotide sequence differences of the longan varieties in this study. However, it can be promising in the future if some further research is carried out in other genomic regions to determine the differences in gene sequences of longan varieties.

CONCLUSION

The genetic diversity of eleven *D. longan* varieties was detected based on molecular biology in Mekong Delta. The results of this study contributed to distinguishing longan varieties at the species level based on nucleotide sequences after analyzing the DNA sequencing results. However, these findings did not clearly show the significant differences in genetic diversity between the longan varieties. Therefore, the study could be extended by further research in other genomic regions of the chloroplast genomes of several longan varieties.

SIGNIFICANCE STATEMENT

The objective of this study was to investigate the genetic diversity of the *matK* gene of the longan varieties in the Mekong Delta. It is one of the first molecular biology studies including "purple" longan. Based on the results of this study, the nucleotide sequences in the *matK* gene were different in many longan varieties. Therefore, this study has made several important contributions to the research of the *matK*'s genetic diversity to improve the commercial value of the longan varieties. Furthermore, the study was as a premise for further applied studies to build DNA barcodes to help distinguish different cultivars.

ACKNOWLEDGMENTS

The fund for this study was invested in the Scientific Research Project of Can Tho University (T2022-131). We thank the Institute of Food and Biotechnology at Can Tho University for its instrumental support.

REFERENCES

1. Cios, K.J., R.W. Swiniarski, W. Pedrycz and L.A. Kurgan, 2007. Data Mining: A Knowledge Discovery Approach. 1st Edn., Springer, New York, ISBN: 978-0-387-36795-8, Pages: 606.
2. Mai, D.T.H., H.N. Hon, T.G. Huy, N.H.B. Lieu, T.T. Men and D.T. Khang, 2020. Morphological and genetic characteristics of three species of the genus (*Mimosa*) [In Vietnamese]. Tap chiKhoa hoc Truong Dai hoc Can Tho, 56: 78-86.
3. Duc, H.H., N.T. Giang, D.H. Xo, H.T. Loan, P.D. Yen, T.T. Tuan and D.D. Giap, 2019. Research using some DNA barcodes in genetic analysis and identification of some species of orchid (*Anoectochilus* spp.) [In Vietnamese]. J. Sci. Can Tho Univ., 55: 14-23.

4. Nguyen, T.D., Q.N. Nguyen, N.L. Tran, T.T. Nguyen and T.P. Ninh *et al*., 2018. Morphological characteristics and DNA barcoding of the one-flowered Seven-leaf clover, *Paris vietnamensis* (Takht.) H.Li, in Vietnam [In Vietnamese]. Vietnam J. Agric. Sci., 16: 282-289.
5. Hebert, P.D.N. and T.R. Gregory, 2005. The promise of DNA barcoding for taxonomy. Syst. Biol., 54: 852-859.
6. Abbas, B., I. Tjolli and Munarti, 2020. Genetic diversity of sago palm (*Metroxylon sagu*) accessions based on plastid cpDNA matK gene as DNA barcoding. Biodiversitas J. Biol. Diversity, 21: 219-225.
7. Hilu, K.W. and H. Liang, 1997. The matK gene: Sequence variation and application in plant systematics. Am. J. Bot., 84: 830-839.
8. Hue, H.T.T., N.T.B. Ngoc, D.Q. Ha, L.T.T. Hien and N.X. Canh, 2022. Analysis of its and matk DNA instructions on some species of lenses of Vietnam [In Vietnamese]. TNU J. Sci. Technol., 227: 318-325.
9. Sil, S., K.K. De and A. Ghosh, 2021. Phylogenetic analysis of six different species of *Saraca* L. (Fabaceae, Caesalpinioideae) based on chloroplast *matK* gene. Biodiversitas J. Biol. Diversity, 22: 3880-3889.
10. Young, N.D. and C.W. dePamphilis, 2000. Purifying selection detected in the plastid gene *matK* and flanking ribozyme regions within a group II intron of nonphotosynthetic plants. Mol. Biol. Evol., 17: 1933-1941.
11. Selvaraj, D., R.K. Sarma and R. Sathishkumar, 2008. Phylogenetic analysis of Chloroplast *matK* gene from Zingiberaceae for Plant DNA barcoding. Bioinformation, 3: 24-27.
12. Skuza, L., I. Szućko, E. Filip and A. Adamczyk, 2019. DNA barcoding in selected species and subspecies of rye (*Secale*) using three chloroplast loci (*matK*, *rbcL*, *trnH-psbA*). Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 47: 54-62.
13. Buerki, S., F. Forest, P. Acevedo-Rodriguez, M.W. Callmander and J.A.A. Nylander *et al*., 2009. Plastid and nuclear DNA markers reveal intricate relationships at subfamilial and tribal levels in the soapberry family (Sapindaceae). Mol. Phylogenet. Evol., 51: 238-258.
14. Lan, N.T.N., N.T.L. Hoa, N.T.T. Thuy and L.T. Nghia, 2019. Study on genetic diversity of *matK* gene fragment in some Vietnamese longan gene sources [In Vietnamese]. Khoa Hoc Nong Nghiep, 61: 61-64.
15. Mathur, R., B.S. Rana and A.K. Jha, 2018. Single Nucleotide Polymorphism (SNP). In: Encyclopedia of Animal Cognition and Behavior, Vonk, J. and T. Shackelford (Eds.), Springer, Cham, Switzerland, ISBN: 978-3-319-47829-6, pp: 1-4.
16. Daniell, H., S.B. Lee, J. Grevich, C. Saski and T. Quesada-Vargas *et al*., 2006. Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other Solanaceae genomes. Theor. Appl. Genet., 112: 1503-1518.
17. Steane, D.A., 2005. Complete nucleotide sequence of the chloroplast genome from the tasmanian blue gum, *Eucalyptus globulus* (Myrtaceae). DNA Res., 12: 215-220.
18. Sugita, M., K. Shinozaki and M. Sugiura, 1985. Tobacco chloroplast tRNA^{Lys}(UUU) gene contains a 2.5-kilobase-pair intron: An open reading frame and a conserved boundary sequence in the intron. Proc. Natl. Acad. Sci. U.S.A., 82: 3557-3561.
19. Turmel, M., C. Otis and C. Lemieux, 2006. The chloroplast genome sequence of *Chara vulgaris* sheds new light into the closest green algal relatives of land plants. Mol. Biol. Evol., 23: 1324-1338.
20. Schmitz-Linneweber, C., M.K. Lampe, L.D. Sultan and O. Ostersetzer-Biran, 2015. Organellar maturases: A window into the evolution of the spliceosome. Biochim. Biophys. Acta (BBA)-Bioenerg., 1847: 798-808.
21. Yesson, C., R.T. Bárcenas, H.M. Hernández, M. de La Luz Ruiz-Maqueda, A. Prado, V.M. Rodríguez and J.A. Hawkins, 2011. DNA barcodes for Mexican Cactaceae, plants under pressure from wild collecting. Mol. Ecol. Resour., 11: 775-783.