Molecular phylogenetic analysis of *Ricinus communis* L. using single loci matK (cpDNA)

Ruchira Javkar[1] Meenakshi Vaidya[2] Priyanka Shukla[2] Darshika Acharya[2] Farzan Afkham[2] Ankit Yadav[2]

[1] *H.V.P.S. Ramniranjan Jhunjhunwala College of Arts, Science & Commerce (Empowered Autonomous), Ghatkopar-(W), Mumbai-400086, Maharashtra, India*

[2]*S.V. K. M's Mithibai College of Arts, Chauhan Institute of Science & Amrutben Jivanlal College of Commerce and Economics (Empowered Autonomous) Affiliated to University of Mumbai, Vile-Parle (West), Mumbai 400056, India*
E-mail: ruchijavkar12@gmail.com

Abstract

The present study explores the molecular phylogeny and species identification of *Ricinus communis* L. using the chloroplast-encoded matK gene as a DNA barcode. The matK gene, due to its high sequence variability and discriminatory power, is widely used for plant identification and phylogenetic analysis. In this study, DNA was extracted from *R. communis* leaves and the matK region was successfully amplified, sequenced, and submitted to GenBank (Accession No. PQ863321). Phylogenetic analysis using the Neighbor-Joining method revealed that *R. communis* shares a close evolutionary relationship with *Acalypha* species, while *Macaranga* species formed a separate clade within Euphorbiaceae. This confirms the matK gene's effectiveness in distinguishing species and resolving evolutionary relationships. The findings support the application of DNA barcoding in plant systematics, biodiversity conservation, and genetic resource management, and highlight the potential of matK in identifying and differentiating economically and medicinally valuable plant species.

Keywords: *Ricinus communis* L., matK gene, DNA barcoding, Phylogenetics analysis, Euphorbiaceae

INTRODUCTION

Ricinus communis L., or castor bean, is a monotypic species belonging to the Euphorbiaceae family and has significant economic, medicinal, and industrial value because of its rich content of ricinoleic acid. Despite its global production and use, the phylogenetic position of *R. communis* within the Malpighiales order is still debated, partially due to conflicting phylogenies as determined by different molecular markers (Xi *et al*., 2012; Cai *et al*., 2019). Chloroplast DNA (cpDNA) has been a cornerstone for many phylogenetic studies due to its uniparental inheritance, moderate mutation rate, and conserved structure (Ye *et al*., 2024). Within the cpDNA loci, the matK gene, or maturase/coding related to the trnK intron, has been recognized as valuable for resolving plant phylogenetics across a variety of taxonomic levels due to its high substitution rates and informative sequence variability (Hilu *et al*., 2003; Selvaraj *et al*., 2008).

Earlier phylogenetic analysis of *R. communis* have mostly used multi-locus methods that combine plastid and nuclear markers (Tokuoka, 2007; Wurdack and Davis, 2009). Single-locus analyses using matK can provide complementary resolution, particularly for closely related species, even though these studies have shed light on the broader evolutionary relationships

within the Euphorbiaceae (Yu *et al.*, 2021). In other Euphorbiaceae genera, including *Jatropha* and *Manihot*, the matK gene has been effectively used to establish species boundaries and infer evolutionary histories (Frajman and Geltman, 2021). Nevertheless, there are still gaps in our knowledge of *R. communis*'s exact position in relation to important lineages within Malpighiales due to the lack of a focused matK-based phylogenetic analysis.

This study aims to reconstruct phylogenetic relationships of *R. communis* from matK sequences and evaluate *R. communis*'s genetic distance from closely related taxa in Euphorbiaceae and other members in Malpighiales. By examining the sequence polymorphisms of *R. communis*, along with potential insertion-deletion events and substitutions, we can evaluate the value of matK in distinguishing species-level phylogeny and resolving taxonomic problems. Given *R. communis*'s significance in agriculture and biotechnology, breeding initiatives, conservation plans, and comparative genomic research will all benefited from a strong phylogenetic framework (Rivarola *et al.,* 2011).

Additionally, this study adds to the broader discussion about the effectiveness of single-locus cpDNA markers in plant systematics. Single-gene analyses are still useful for initial evaluations, particularly in understudied taxa, even though multi-gene approaches are becoming more and more popular (Ye *et al.*, 2024). This study highlights the ongoing value of chloroplast markers in resolving plant phylogeny and offers updated insights into the evolutionary history of *R. communis* by combining new phylogenetic analyses with recent matK sequence data from public databases.

MATERIALS AND METHODS

Sample Collection and Authentication of plant material

Fresh leaves of *Ricinus communis* L. were collected from Sanjay Gandhi National Park (SGNP), Borivali (West), Mumbai, and authenticated. A voucher specimen (Voucher ID: MIT0177) was prepared and deposited in the Department of Botany at Mithibai College, Mumbai, for future reference.

DNA extraction

Genomic DNA was extracted from fresh leaves using a modified cetyltrimethylammonium bromide (CTAB) method (Doyle *et al*., 1990). The leaves were ground to a fine powder in a mortar with liquid nitrogen and subsequently mixed with CTAB isolation buffer. The quality of the extracted DNA was assessed by electrophoresis on a 0.8% agarose gel. Quantification of the DNA was performed using a NanoDrop spectrophotometer (Thermo Fisher).

PCR amplification

PCR amplification was conducted to target plastid gene regions maturase kinase (matK). Each PCR reaction was performed in a 25 µL reaction volume, containing 90–100 ng of DNA template, 2.5 U Taq DNA polymerase (Thermo Fisher), 2.5 mM $MgCl_2$, 1X Taq DNA polymerase buffer, 0.2 mM dNTPs, and 0.5 µM primers procured from Eurofins Genomics India Pvt. Ltd. The primer sequences and their respective annealing temperatures are detailed in Table 1 and Table 2. The amplified products were resolved on a 1.5% agarose gel stained with 0.1 µg/mL ethidium bromide using 1X TAE buffer. The separated amplicons were purified to remove potential contaminants using the Qiagen PCR clean-up kit and subsequently sent to Eurofins Pvt. Ltd. (India) for sequencing.

Table 1. Primer used for PCR

Gene (matK)	Primer Sequence	Primer Reference
Forward primer	CGATCTATTCATTCAATATTTC	(Cuénoud *et al.*, 2002)
Reverse primer	TCTAGCACACGAAAGTCGAAGT	(Levin *et al*., 2003; Kress and Erickson, 2007)

RESULTS AND DISCUSSIONS

PCR result and Quality Assessment

The successful amplification of the matK region of *Ricinus communis* through PCR demonstrated high-quality results. The primer set used in this study effectively amplified the target region, producing a distinct and sharp band on agarose gel electrophoresis, confirming its suitability for downstream DNA sequencing. The obtained nucleotide sequences were subsequently submitted to the NCBI GenBank database and assigned the accession number PQ863321 (Fig.1).

>PQ863321 *Ricinus communis* voucher MIT0177 maturase K (matK) gene, partial cds; chloroplast

ATTTAAATTATGTGTCAGATGTATTAATACCTTACCCCATCCATCTAGAAAAATTGGTTCAAATCCT
TCGCTATTGGGTGAAAGATCCCTCTTCTTTGCATTTATTACGACTCTTTCTTCATGAGTATTGGAAT
TGGAACAGTTTTATTATTCCAAAGAAATCAATTTCTATTTTTACAAAAAGTAATCCAAGATTTTTCG
TGTTCCTATATAATTCTCATGTATATGAATATGAATCCCTCTTCTTTTTTCTCCGTAACCAATCCTTT
CATTTACGATCAACATTTTCTCGAGTACTTCTTGAACGAATTTTTTTCTATGGAAAAATAGAACATT
TTGCGGAAGTCTTTGCTAATGATTTTCAGGCCATCCTATGGTTGTTCAAGGACCCTTTCATGCATTA
TGTTAGATATCAAGGAAAATCTGTTTTGGCTTCAAAAGATGGGCCTCTTCTGATGAAAAAATGGAA
ATATTACCTTGTCCATTTATGTCAATGTCATTTTTATGTGTGGTTTCAACCGGAAAAGATCTATATA
AATTCATTATCTAAGCATTCTCTCAACTTTTTGGGCTATCTTTCAAATGTACAATTTAATCCTTCGTT
GGTACGGAGTCAAATGATAGAAAATTCATTTATAATAGATAAAGATAATACTATGAAGAAACTCG
ATACAATAGTTCCAATTATTCCTTTAATTAGATCATTGGCAAAAATGAAATTTTGTAACGCAGCAG
GACATCCCATTAGTAAACCGACCTGGGCGGATTCGGCAGATTCTGAGAT

Fig 1. matK barcode sequence of *Ricinus communis* (GenBank accession number PQ863321)

Sequence Analysis

Evolutionary analysis by Neighbor-Joining method

The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Negi, 1987), and the optimal tree is presented. Bootstrap analysis (1000 replicates) was performed to assess the reliability of the tree, with the percentage of replicate trees supporting each cluster indicated next to the branches (Felsenstein, 1985). Evolutionary distances were calculated using the Maximum Composite Likelihood method (Tamura *et al.,* 2004) and are expressed as the number of base substitutions per site. The proportion of sites containing at least one unambiguous base in at least one sequence per descendant clade is shown next to each internal

node. This analysis included nine nucleotide sequences, incorporating all codon positions (1st, 2nd, 3rd, and noncoding). Ambiguous positions were removed using the pairwise deletion option. The final dataset comprised 868 positions. All evolutionary analyses were conducted using MEGA11 (Tamura *et al.,* 2021) (Fig. 2).

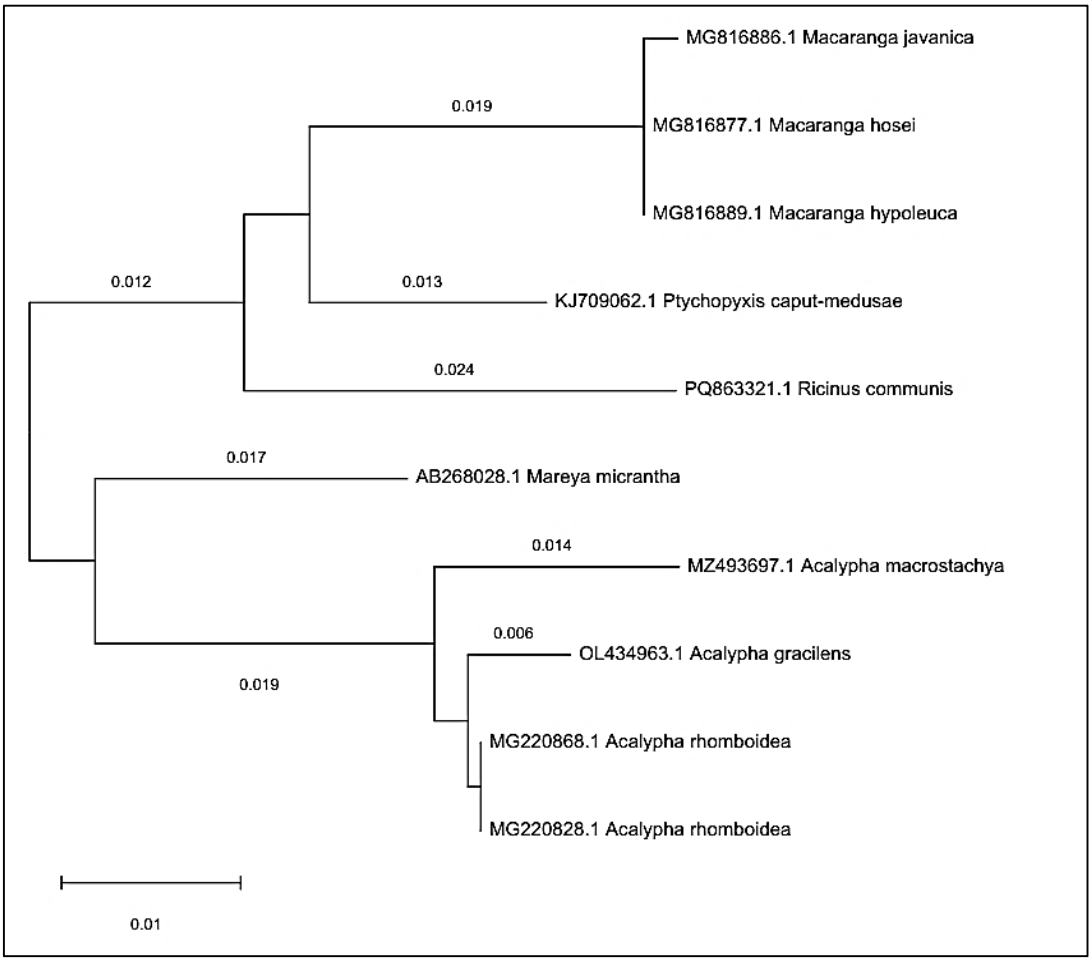

Fig. 2. Phylogenetic tree constructed using the Neighbor-Joining method based on partial matK gene sequences showing the evolutionary relationships among selected Euphorbiaceae species. *Ptychopyxis caput-medusae* was used as the outgroup.

In this study, the phylogenetic analysis includes several representative taxa from the Euphorbiaceae family, with a focus on evaluating the evolutionary position of *Ricinus communis* using the matK gene. The genetic sequences retrieved from GenBank and aligned using MEGA11 software, reveal evolutionary relationships based on variations in their matK sequences.

The phylogenetic analysis based on the matK chloroplast gene revealed distinct genetic divergence patterns among the studied taxa (Fig. 2). The distance matrix displayed pairwise genetic distances ranging from 0.006 to 0.024, with the lowest divergence observed between *Macaranga hosei* (MG816877) and *Macaranga hypoleuca* (MG816889) (0.006), suggesting a close evolutionary relationship. In contrast, *Ricinus communis* (PQ863321) exhibited higher divergence from *Macaranga* species (0.017–0.024), reveling its distinct phylogenetic position within Euphorbiaceae.

Table 2. Taxa Included in the Phylogenetic Tree

Species Name	GenBank Accession Number
Ricinus communis	PQ863321
Acalypha macrostachya	MZ493697
Acalypha gracilens	OL434963
Acalypha rhomboidea	MG220868
Acalypha rhomboidea	MG220828
Macaranga javanica	MG816886
Macaranga hosei	MG816877
Macaranga hypoleuca	MG816889
Ptychopyxis caput-medusae	KJ709062
Mareya micrantha	AB268028

In accordance with its established taxonomic distinction from the ingroup taxa, *Ptychopyxis caput-medusae* (KJ709062) was clearly resolved as the outgroup in the neighbor-joining (NJ) tree derived from the distance matrix (Wurdack and Davis, 2009).

The analysis revealed several key phylogenetic relationships: First, the *Macaranga* species (*M. javanica*, *M. hosei,* and *M. hypoleuca*) formed a well-established monophyletic clade (bootstrap >70%), demonstrating their close genetic affinity. Second, *Ricinus communis* occupied a distinct phylogenetic position, clustering separately from both *Macaranga* and *Acalypha* species, which is consistent with previous studies recognizing its unique evolutionary lineage within Euphorbiaceae (Xi *et al.*, 2012). Third, the *Acalypha* species exhibited varying levels of genetic divergence, with the two accessions of *A. rhomboidea* (MG220828 and MG220868) showing particularly low genetic distance (0.008), suggesting either recent speciation events or intraspecific variation. In contrast, *A. gracilens* (OL434963) and *A. macrostachya* (MZ493697) displayed slightly higher genetic distances (0.010-0.014). These results collectively provide important insights into the phylogenetic structure of these taxa while highlighting the utility of the matK marker for resolving relationships within Euphorbiaceae.

The matK gene was effective in distinguishing between genera, with interspecific distances exceeding the intraspecific variation limits suggested for barcoding (Li *et al.*, 2015). However, the low divergence between accessions of *A. rhomboidea* suggests that matK alone should not be relied upon for identifying species in recently radiated groups.

CONCLUSION

Overall, this study highlights the value of the matK gene as a reliable molecular tool for species identification and phylogenetic inference. It facilitates a deeper understanding of genetic relationships within Euphorbiaceae and provides a scientific basis for further research in plant systematics, conservation, and applied plant sciences.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

1. Cai, L., Xi, Z., Amorim, A.M., Sugumaran, M., Rest, J.S., Liu, L. and Davis, C.C., 2019. Widespread ancient whole-genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytol.*, 221(1): 565–576.
2. Cuénoud, P., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J. and Chase, M.W., 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. *Am. J. Bot.*, 89(1): 132–144.
3. Doyle, J.J., Doyle, J.L. and Brown, A.H., 1990. Chloroplast DNA phylogenetic affinities of newly described species in *Glycine* (Leguminosae: Phaseoleae). *Syst. Bot.*, 15: 466–471.
4. Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4): 783–791.
5. Frajman, B. and Geltman, D., 2021. Evolutionary origin and systematic position of *Euphorbia normannii* (Euphorbiaceae), an intersectional hybrid and local endemic of the Stavropol Heights (Northern Caucasus, Russia). *Plant Syst. Evol.*, 307(2): 20.
6. Hilu, K.W., Borsch, T., Müller, K., Soltis, D.E., Soltis, P.S., Savolainen, V. and Chatrou, L.W., 2003. Angiosperm phylogeny based on matK sequence information. *Am. J. Bot.*, 90(12): 1758–1776.
7. Kress, W.J. and Erickson, D.L., 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS ONE*, 2(6): e508.
8. Levin, R.A., Wagner, W.L., Hoch, P.C., Nepokroeff, M., Pires, J.C., Zimmer, E.A. and Sytsma, K.J., 2003. Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF data. *Am. J. Bot.*, 90(1): 107–115.
9. Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y. and Chen, S., 2015. Plant DNA barcoding: from gene to genome. *Biol. Rev.*, 90(1): 157–166.
10. Rivarola, M., Foster, J.T., Chan, A.P., Williams, A.L., Rice, D.W., Liu, X. and Rabinowicz, P.D., 2011. Castor bean organelle genome sequencing and worldwide genetic diversity analysis. *PLoS ONE*, 6(7): e21743.

11. Saitou, N. and Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 4(4): 406–425.

12. Selvaraj, D., Sarma, R.K. and Sathishkumar, R., 2008. Phylogenetic analysis of chloroplast matK gene from Zingiberaceae for plant DNA barcoding. *Bioinformation*, 3(1): 24–27.

13. Tamura, K., Nei, M. and Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA*, 101(30): 11030–11035.

14. Tamura, K., Stecher, G. and Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.*, 38(7): 3022–3027.

15. Tokuoka, T., 2007. Molecular phylogenetic analysis of Euphorbiaceae sensu stricto based on plastid and nuclear DNA sequences and ovule and seed character evolution. *J. Plant Res.*, 120: 511–522.

16. Wurdack, K.J. and Davis, C.C., 2009. Malpighiales phylogenetics: gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. *Am. J. Bot.*, 96(8): 1551–1570.

17. Xi, Z., Ruhfel, B.R., Schaefer, H., Amorim, A.M., Sugumaran, M., Wurdack, K.J. and Davis, C.C., 2012. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proc. Natl. Acad. Sci. USA*, 109(43): 17519–17524.

18. Ye, J., Luo, Q., Lang, Y., Ding, N., Jian, Y.Q., Wu, Z.K., Wei, S.H. and Yan, F.L., 2024. Analysis of chloroplast genome structure and phylogeny of the traditional medicinal *Ardisia crispa* (Myrsinaceae). *Sci. Rep.*, 14(1): 19045.

19. Yu, J., Wu, X.I., Liu, C., Newmaster, S., Ragupathy, S. and Kress, W.J., 2021. Progress in the use of DNA barcodes in the identification and classification of medicinal plants. *Ecotoxicol. Environ. Saf.*, 208: 111691.