Indian J. Applied & Pure Bio. Vol. 40(1), 345-352 (2025). **ISSN: 0970-2091**
A Web of Science Journal

DNA Barcoding and phylogenetic analysis of fish *Oreochromis mossambicus* from Mallapura Lake, Chitradurga, India

C. N. Venksatesh, *Shashikanth H. Majagi, Rabiya Begum and Shahanawaz Ahemad

[1,2*]Department of Studies in Zoology, Vijayanagar Sri Krishnadevaraya University, Ballari - 583105 (India)
[3,4]Department of Zoology, Tumkur University. Tumkur - 572103 (India)
Email: smajgi@rediffmail.com
*Corresponding author

Abstract

India is one of the major contributors for fish biodiversity and these are largest group of organisms which shows varied phenotypical changes throughout the development. The Chitradurga region forms a continental bridge between the Vedavathi River & Krishna River basin. Hence, fresh water fauna in Mallapura lake consist of unique populations of species that originates from different range of distributions. The identification & discoveries of new species can be made easily through barcoding technique. Leading to resolve traditional morphological taxon identification for overlapping sister species. The present study aimed to apply the same technique of mitochondrial cytochrome oxidase I gene for accuracy with identification.The tilapia fish samples were collected from Mallapura lake, and DNA amplification results showed COI gene fragment sequences, characterized by a length of 654 base pairs. The nucleotide composition, polymorphic sites, haplotype grouping, BLAST analysis, and phylogenetic tree constructed by collected individual were used to categorize the samples into *Oreochromis mossambicus* species. The results indicated DNA barcodes as an effective identification approach for tilapia fish, and the results have a potential for application in aquaculture and during the management of fisheries resource in India. The detailed investigation provides a barcoding database and promotes further studying of the fish species in the region and their ecology with emphasis on monitoring and conservation of freshwater fish biodiversity.

Key words : *Oreochromis mossambicus,* COI gene, Bar coding, Phylogeny, Mallapur Lake, Chitradurga.

The recent advances with urbanization, anthropogenic activities and rapid growth in small scale industries lead to the loads of pollution towards the lakes of India, the reports with high concentrations of heavy metals into the water bodies directly links to the ecological food chain and resulting towards the greater concern to the ecological researchers[19]. Strategizing measures for conservation and improvement of stocks can be made possible by Understanding the level of genetic diversity in any population[6].

The identification of fish species mainly relies on morphometric and meristic characters[14,17]. Fish have remarkable diversity of morphological characteristics and most fish go through ontogenetic metamorphism. Many morphometric characteristics change during the stages of ontogenetic development[20]. Convergent and divergent adaptation also lead to changes in the morphological characteristics of fish species, imposing great challenges to morphological taxonomy, in which species identification is mainly based on morphological characteristics, and the classification of many species has thus also been controversial[9]. The limitations inherent in morphology-based identification systems and the declining number of taxonomists call for a molecular approach to identify species[13,20].

In 2003, Hebert *et al.,*[7] proposed DNA bar-coding technology, in which the Mitochondrial Cytochrome Oxidase subunit I (COI) gene sequence was used as a barcode for species identification with the expectation of bar-coding all species for the purpose of species identification and classification[5,18].

Tilapia, Tilapiine fishes have a huge species diversity, and are grouped into three main genera: Oreochromis (arena-spawning maternal mouthbrooders), Sarotherodon (paternal or biparental mouthbrooders) and Tilapia (substrate spawners)[3,16]. Previous reports show extensive investigation on Tilapiine fish diversity, using both morphological methods [10,16] and molecular markers[1,15]. These discrepancies in procedure have led to contradictory patterns in species description and identification through morphometric as well as meristic characters have been implicated in misidentification, taxon ambiguities andfluctuation in total species number. Furthermore, the main culprits for this phenomenon include phenotypic and genotypic plasticity, cryptic diversity or possible hidden species and variation in colour pattern at different stages of life within the same species[2].

The fisheries management and aquaculture studies majorly depend on investigations on fish taxonomy as it's a basic component for any of the further studies. The present study focusses on the identification of tilapia fish existing in Mallapura lake on the basis of nucleotide composition, polymorphic sites, haplotype grouping, nucleotide BLAST, and phylogenetic tree analysis for cytochrome C oxidase I gene. This required a DNA barcoding process, which involves theproduction of PCR amplicons from COI gene to generate a sequence data, which is subsequently used toascertain and distinct the organism from other species.

The identification of fish species still stands as one of the most basic but important issues in fisheries management. This study aims to investigate the COI gene sequence of

mt DNA for molecular identification and phylogenetic analysis of a tilapia species (*Oreochromis mossambicus*) from Mallapur Lake in Chitradurga.

Study area :

The Mallapur Lake situated towards the North-East of Chitradurga which is having about 100 acres of total area, & depth of 3-4mts. The lake caters the agricultural needs of neighboring villages. The Geographic co-ordinates of lake is 14.23' N latitude, 76.4' E longitude 735mts above MSL (Fig 1).

The specimen was collected and morphologically identified by visual inspection. Taxonomic classification was done with the aid of the field guide to Indian freshwater fish and freshwater fishes of India by Jayram[8].

DNA extraction :

DNA extraction was carried out using Zymo Research Genomic DNA™-Tissue Miniprep extraction kit, and the manufacturer protocol was followed.

Polymerase Chain Reaction (PCR) amplification :

The CO1 gene located in the mitochondrial genome was amplified using the pair of primers (Appendix. 1.) (Fig. 2).

Appendix. 1. Primer details

Fish F2	TGTAAAACGACGGCCAGTCGACTAATCATAAAGATATCGGCAC
Fish R2	CAGGAAACAGCTATGACACTTCAGGGTGACCGAAGAATCAGAA

PCR amplification was performed with Eppendorf Master cycler personal in 50 µl total reaction volume using OneTaq Quick-Load 2X Master Mix with Standard Buffer (New England Biolabs Inc) ready-to-use kit. The PCR was carried out under the following conditions: an initial denaturation of 94°C for 30 s, followed by 35 cycles of 94°C for 30 s, 60°C for 1 min, 68°C for 1 min; followed by final extension of 68°C for 10 min and held at 20°C. The PCR products were visualized on 1% agarose gel electrophoresis stained with ethidium bromide.

Purification of PCR products and sequencing:

The PCR products were purified using the Zymo Research DNA clean and concentrator TM-5. In a 1.5 ml micro centrifuge tube, a 5:1 ratio of DNA biding buffer was added to the PCR products. In a collection tube, the mixture was transferred to a Zymo-spinTM column. The flow-through was discarded after the sample was centrifuged for 30 seconds at 10,000x g. 200 l of DNA wash buffer was added to the spin column and centrifuged at 10,000x g for 30 seconds. The washing process was repeated. A 15-l volume of DNA elution buffer was added directly to the column matrix and incubated at room temperature for one minute. This was transferred to a 1.5 l micro-centrifuge tube and centrifuged at 10,000x g for 30 seconds. The purified PCR products were sequenced using Genetic Analyzer (Applied Biosystems 3130 XL).

Figure 1: Satellite image of the Mallapur Lake, Chitradurga, Karnataka, India
Sample collection and initial identification

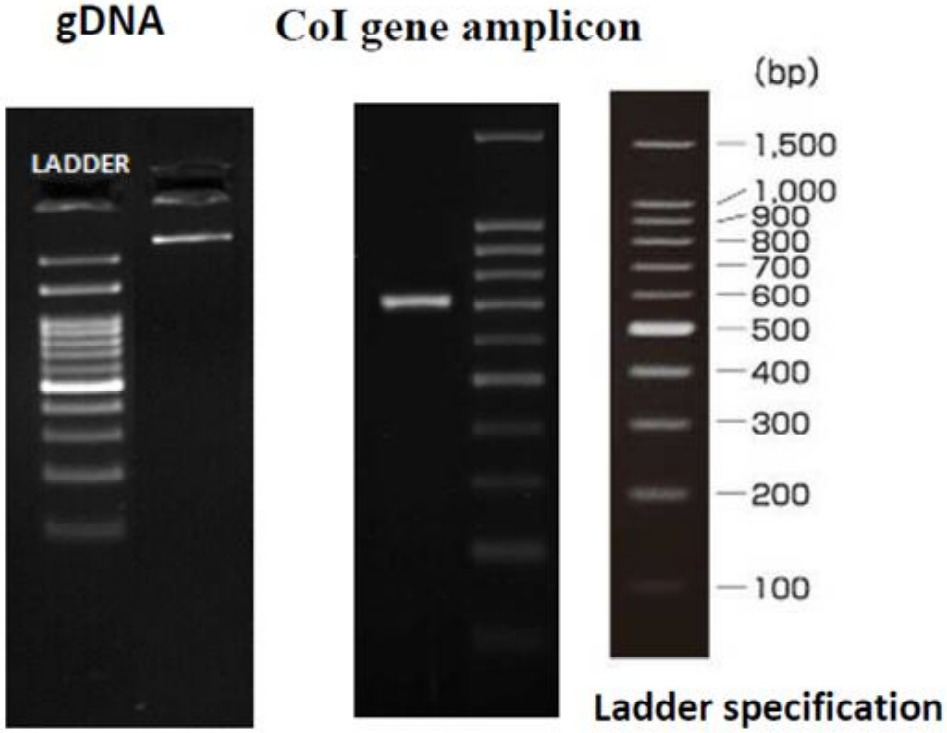

Figure 2: Amplification of mitochondrial COI locus (550-700bp)

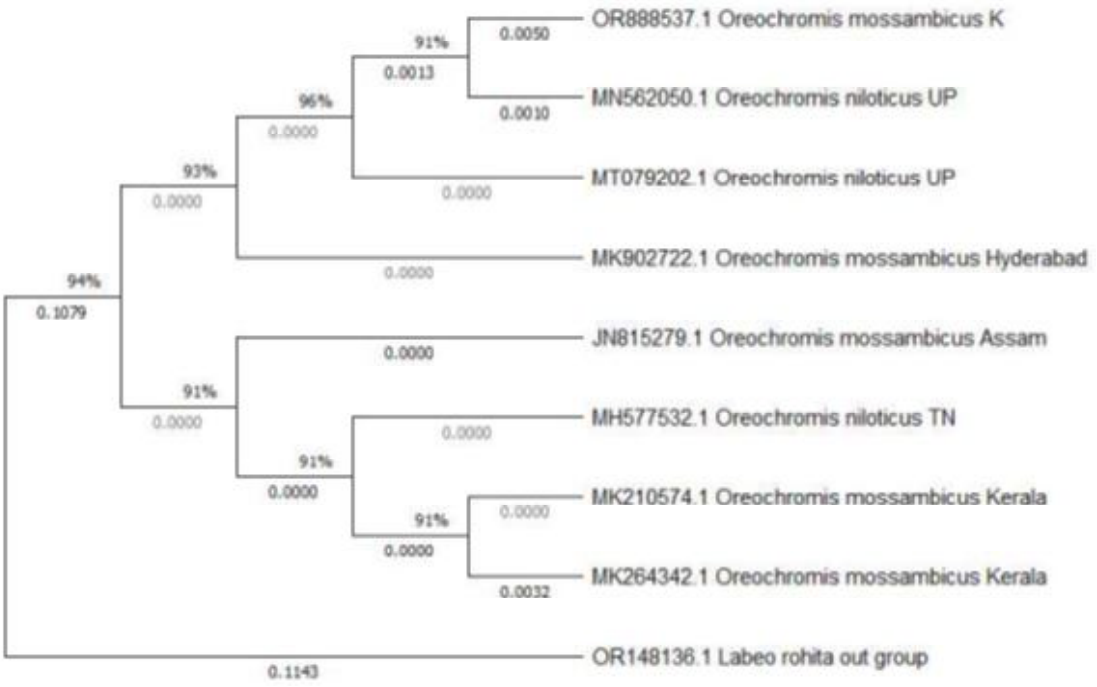

Figure 3: Phylogenetic analysis of *Oreochromis mossambicus* comparing Chitradurga tilapia barcode sequences and the sequences of other tilapia from India. Numbers attached to species name indicate GenBank ID.

Sequence analysis :

The sequencing (Figure 2.) was subjected to the Barcode of Life Database Identification (BOLD-ID) and BLAST algorithms from the National Centre of Biotechnology Information, Bethesda, MD, USA (http:// www.ncbi.nlm.nih.gov) to search for homologous sequences in Gen Bank. BioEdit[12] was used to edit the sequences, and Clustal X 2.1 [11] was used to align them. For phylogenetic and analysis purposes, MEGA X was used A Maximum Likelihood tree of K2P distances was created to provide a graphic pattern of species divergence A minimum of 1000 bootstrap replications were used for distance generation. To identify samples based on sequence similarity, the Barcoding of Life Database (BOLD) (http://www. barcodinglife.org) and GenBank (http://www.ncbi.nlm.nih.gov) were used[11].

Nucleotide bases in the genomic DNA of Oreochromis mossambicus from Chitradurga *:*

Number of nucleotide bases A, C, G, T, in the genomic DNA of *Oreochromis mossambicus* was estimated as153 (24.4 %), 202 (29.5%), 119 (17.4 %) and 197 (28.8 %), respectively. We found relatively high A+T content of 53.2 % compared to 46.9 % of G+C content. Percent of AT content at the 1st, 2nd

Table-1. The list of COI nucleotide sequences of the *O. mossambicus* from different locations of India including one of the present studies (highlighted).

State	Location	Accession No. / Voucher ID
Uttar Pradesh	Lucknow	MN562050.1
Uttar Pradesh	Lucknow	MT079202.1
Hyderabad	Attapur	MK902722.1
Assam	Silchar	JN815279.1
Tamil Nadu	Sirkazhi	MH577532.1
Kerala	Ernakulam	MK210574.1
Kerala	Ernakulam	MK264342.1
Karnataka	Chitradurga	OR888537.1

Table-2. The list of COI nucleotide frequencies of the *O. mossambicus* from different locations of India including one of the present studies (highlighted).

	T(U)	C	A	G	Total
MH577532 TAMILNADU	28.9	29.4	24.5	17.3	654
MK264342 KERALA	28.7	29.5	25.2	16.6	651
MN562050 UTTARPRADESH	28.7	29.8	24.6	16.8	641
OR888537 KARNATAKA	28.8	29.6	24.2	17.4	683
MK902722 HYDERABAD	28.9	29.4	24.2	17.5	657
MT079202 UTTARPRADESH	28.4	29.4	25.0	17.3	701
JN815279 ASSAM	28.7	29.6	24.5	17.3	642
MK210574 KERALA	28.7	29.5	24.2	17.6	654

and 3rd codon positions of *Oreochromis mossambicus* was 43.5 %; 57.2 % and 58.8 %, respectively,with slight variations.

Comparative analysis of Oreochromis mossambicus (OR888537.1) from different regions of India :

A phylogenetic analysis was performed by using neighbor joining method with Mega (Version 7.0). The barcode sequences of *Oreochromis mossambicus* from Chitradurga regionand those of other *Oreochromis mossambicus* from India, available in BOLD were compared to reveal the locus and phylogenetic relationship (Table-1).

The distance analysis of eight nucleotide sequences representing Uttar Pradesh, Hyderabad, Assam, Tamil Nadu, Kerala and Karnataka using MEGA X software (https://www.megasoftware.net) revealed frequency and percent of the nucleotides were: A: 1294 (24.2%); C: 1558 (29.5%); G: 929 (17.2%) and T: 1516 (28.7%). The sequences are comprised of 2810 and 2487 bp of A+T and G+C, respectively, which involves 53.04% and

46.95% of A+T and G+C content with an average length of 610 bp per submission. The COI nucleotide frequencies of *Oreochromis mossambicus* can be seen in suppl. mat. (Table 2). Periods in the sequence letters represents conserved bases between populations from each location. Conserved bases are indicative of similarities between populations whereas the different bases account for separation of the different populations of *Oreochromis mossambicus*.

A phylogenetic analysis revealed two clusters (excluding the out-group). The first cluster is consisted with *Oreochromis mossambicus* Kerala, Tamil Nadu and Assam. The second cluster consisted of *Oreochromis mossambicus* from Hyderabad, Uttar Pradesh and Karnataka (including the Chitradurga population) (Figure 3). *Oreochromis mossambicus* from Chitradurga, Karnataka forms sister group clade from *Oreochromis mossambicus* of Lucknow, Uttar Pradesh and other Lucknow population forming an independent clade.

These results indicate that a DNA barcode could be used to resolve the ambiguity in species identification.

The present investigation gives complete report on the single species of *O. musambicus* with detailed genomic information. Molecular COI gene sequence data of *O. musambicus* from the Mallapur Lake of Chitradurga, India shows similarity with gene sequences of this species from other parts of India suggesting that this Tilapia species maintain greater genetic similarity across varied climatic regions.

Funding: No funding agency.

Consent for publication: We have consent for publication.

Conflict of interest: The authors declare that they have no conflict of interest.

'Declaration of Generative AI and AI-assisted technologies in the writing process'

AI and its assisted technologies not used for preparation of the manuscript.

Details of author contributions :

Author 1: Involved in complete study plan & working protocol.

Author 2: Helped in reviewing of article.
Author 3: Helped in writing & reviewing of article.
Author 4: Helped in analysis.

References :

1. Arifin O. Z., E. Nugroho, and R. Gustiano (2007). *Berita Biologi 8*(6): 465-471.
2. Barman A.S., M. Singh, and S.K. Singh (2018). *Sci Rep* 8, 8579.https://doi.org/10.1038/s41598-018-26976-3
3. Canonico Gabrielle., Arthington Angela., Mccrary Jeffrey., and Thieme Michele., (2005). *Aquatic Conservation: Marine and Freshwater Ecosystems*. *15:* 10.1002/aqc.699.
4. Dailami Muhammad., Rahmawati Aulia., Saleky Dandi., Toha Abdul., Hamid Abdul Toha A. (2021). *AACL Bioflux. 14:* 849-858.
5. Doña J., J. Diaz-Real, S. Mironov, P. Bazaga, D. Serrano, and R. Jovani (2015) *Mol Ecol Resour.*1216-25. doi: 10.1111/1755-

0998.12384.

6. Ekerette E., E. Ikpeme, O. Udensi, M. Ozoje, O. Etukudo, A. Umoyen, S. Durosaro and M. Wheto (2018) *American Journal of Molecular Biology*, *8:* 39-57. doi: 10.4236/ajmb.2018.81004.
7. Hebert PD., A. Cywinska, SL. Ball, and JR. deWaard (2003) *Proc Biol Sci*. Feb *7:* 270(1512): 313-21. doi: 10.1098/rspb.2002.2218.
8. Jayaram K.C. (1981) The Freshwater Fishes of India. Hand Book, *Zoological Survey of India, Calcutta*.
9. Keskİn E., and H.H. Atar, (2013), *Mol EcolResour*, *13:* 788-797. https://doi.org/10.1111/1755-0998.12120
10. Ndiwa, T.C., D.W. Nyingi, and J. Claude, (2016). *Environ Biol Fish* 99: 473–485 https://doi.org/10.1007/s10641-016-0492-y
11. O.A. Sogbesan., M.K. Sanda., J.N. Jaafar., and H.A. Adedeji. (2017).*J. Biotec. Biomat., 7*(4): pp. 1-4, 10.4172/2155-952X.1000277
12. Sikes DS., M. Bowser, JM. Morton, C. Bickford, S. Meierotto and K. Hildebrandt (2017). *Genome. 60*(3): 248–259. doi: 10.1139/gen-2015-0203.
13. Steinke Dirk., Zemlak Tyler., Boutillier James., and Hebert Paul. (2009). *Marine Biology*. 156: 2641-2647. 10.1007/s00227-009-1284-0.
14. Strauss RE, and CE. Bond (1990). Taxonomic methods: Morphology. In: Schreck CB, Moyle PB. editors. Methods for fish biology. Bethesda, MD: *American Fisheries Society.* 109–140.
15. Tibihika Papius Dias., Manuel. Curto, Esayas. Alemayehu, H. Waidbacher, Charles. Masembe, Peter. Akoll, and Harald. Meimberg (2020). *BMC Evolutionary Biology*. *20:* 10.1186/s12862-020-1583-0.
16. Trewavas E. (1983) Tilapine Fishes of the Genera Sarotherodon, Oreochromis and Danakilia. *London British Museum (Natural History)*, 583 p.
17. Triantafyllidis A., D. Bobori, C. Koliamitra, E. Gbandi, M. Mpanti, O. Petriki, and N. Karaiskou (2011) *Mitochondrial DNA*. *1:* 37-42. doi: 10.3109/19401736.2010.542242.
18. Weigt LA., CC. Baldwin, A. Driskell, DG. Smith, A. Ormos, and EA. Reyier (2012) *PLoS One*. *7*(7): doi: 10.1371/journal.pone.0041059.
19. Zaidi J., and A. Pal (2017). *African Journal of Environmental Science and Technology*, *11*(6): 255-265.
20. Zhang Junbin., and Robert. Hanner (2011). *Biochemical Systematics and Ecology*. *39:* 31-42. 10.1016/j.bse.2010.12.017.t