



Molecular Identification of Dominant Fish Species Using DNA Barcoding from Two Different Ecosystems of Sindhudurga.

Nandu Hedulkar¹, Abhay Morajkar^{2*}, Vikas Jha³

Abstract:

DNA barcoding has emerged as a universal molecular tool for accurate species identification using mitochondrial cytochrome c oxidase subunit I (COI) gene sequences. This study focuses on identifying dominant fish species from two distinct aquatic ecosystems of Sindhudurga - lentic (lake) and lotic (river) environments located in of Tarkarli village of Malvan taluka & Sindhudurg district of Maharashtra, India. Fish samples were collected from Chorage wadi lake (16.082126°N, 73.724685°E) and the Karli River near Sateri Temple (16.043124°N, 73.69119°E). Genomic DNA was extracted using the CTAB method, and the COI gene was amplified and sequenced using Sanger sequencing. The sequences were analyzed through BLAST (Basic Local Alignment Search Tool) to confirm taxonomic identity by comparison with reference sequences in GenBank. Three dominant species — *Eleotris fusca*, *Aplocheilus lineatus*, and *Rasbora daniconius* — were identified with >98% sequence similarity and deposited under GenBank accession numbers PV235273–PV235275. These findings reaffirm the effectiveness of DNA barcoding as a rapid, reliable method for species identification, offering essential data for biodiversity monitoring and ecosystem management in freshwater and semi estuarine habitats.

¹ Anandibai Raorane Arts, Commerce, and Science College, Vaibhavwadi- 416810, India. (ORCID ID: 0000-0002-2613-4475)

²VPM's B. N. Bandodkar College of Science (Autonomous), Thane- 400601, India.

³ National Facility for Biopharmaceuticals, Guru Nanak Khalsa College, Mumbai- 400019, India(ORCID ID: 0000-0002-4008-1592)

Corresponding Author: asmorajkar@vpmthane.org

Keywords: DNA barcoding, COI gene, BLAST, fish biodiversity, lentic and lotic ecosystems, molecular identification

Introduction:

Aquatic biodiversity, particularly ichthyofaunal diversity, is crucial for maintaining ecological balance and ecosystem functioning.¹ Traditional morphological taxonomy, while foundational, often faces challenges such as cryptic morphology, ontogenetic variation, and specimen damage during collection.² DNA barcoding based on the mitochondrial cytochrome c oxidase subunit I (COI) gene has become a global standard for molecular identification due to its high interspecific variability and minimal intraspecific divergence.³ This molecular tool supports precise species authentication, population studies, and biodiversity conservation.^{4,5} The present study employs COI-based DNA barcoding to identify dominant fish species from two ecologically contrasting systems - lentic (pond) and lotic (river) - within the tarkarli region of Maharashtra. The goal is to document molecularly validated species records for regional biodiversity assessment without conducting comparative or phylogenetic inference analyses. This study aims to extract high-quality genomic DNA from dominant fish species sampled from lentic and lotic systems in tarkarli. Further amplifying the sequences of COI gene using standard molecular barcoding protocols. Lastly to perform BLAST-based molecular identification by comparing sequences with the NCBI GenBank database with deposition of validated COI sequences in GenBank as a molecular reference for future biodiversity and ecological studies.

Materials and Methods:

Genomic DNA was extracted from fish muscle tissue using the CTAB method following standardized protocols. Approximately 20–30 mg of tissue was homogenized with liquid nitrogen and digested with CTAB buffer and Proteinase K at 60°C for 30 minutes. DNA was purified using chloroform: isoamyl alcohol extraction and ethanol precipitation, washed, and dissolved in TE buffer. DNA integrity was evaluated using 0.7% agarose gel electrophoresis and quantified by Nanodrop spectrophotometry.⁶ Working concentrations (10–100 ng/μL) were used for PCR amplification. PCR reactions were prepared in 50 μL volumes containing 1X Taq buffer, 2.5 mM MgCl₂, 0.2 mM dNTPs, 0.5 μM primers, and 5 U Taq polymerase. The COI gene was amplified using thermal cycling conditions: 96°C for 5 min, 35 cycles of 95°C (45 s), 54°C (45 s), 72°C (60 s), and final extension at 72°C for 5 min. PCR products (~900 bp) were visualized on agarose gel, purified using the Qiagen PCR purification kit, and subjected to Sanger sequencing (Beckman Coulter CEQ 8000+).⁷ Sequence chromatograms were analyzed in FinchTV, trimmed for quality, and compared using BLASTn against the NCBI GenBank database. Species identification was based on highest sequence similarity (>98%). Verified sequences were deposited in GenBank with accession numbers PV235273 (*Eleotris fusca*), PV235274 (*Aplocheilus lineatus*), and PV235275 (*Rasbora daniconius*).

Results

The raw sequences obtained from Sanger sequencing were analyzed using Finch TV software to assess sequence quality. Low-quality bases, particularly at the 5' and 3' ends, were manually trimmed to enhance sequence accuracy. After trimming, sequence integrity was verified, and high-quality contigs were assembled. The processed sequences were then subjected to BLAST (Basic Local Alignment Search Tool) analysis against the NCBI GenBank database to identify the closest homologous sequences. Species identification was based on the highest percentage of sequence similarity with reference sequences available in the database.

The validated sequences were formatted according to NCBI submission guidelines and deposited in GenBank under the accession numbers PV235273 (*Eleotris fusca*), PV235274 (*Aplocheilus lineatus*), and PV235275 (*Rasbora daniconius*). These accession numbers were recorded for future reference and comparative analysis.

Multiple Sequence alignment

Focusing on the multiple sequence alignment of *Eleotris fusca*, *Aplocheilus lineatus*, and *Rasbora daniconius*, with *Mugil cephalus* as the outgroup, we observe distinct patterns of sequence conservation and variation. *E. fusca* displays the highest degree of sequence divergence, with numerous nucleotide differences compared to the other species. This is consistent with its phylogenetic position as a distinct and early diverging lineage. In contrast, *A. lineatus* and *R. daniconius* exhibit greater sequence similarity, particularly in certain regions, supporting their closer evolutionary relationship as indicated in the phylogenetic tree. *M. cephalus*, serving as the outgroup, shows considerable sequence divergence from all three species, confirming its distant relationship and validating its use as an outgroup for rooting the tree. These observations from the multiple sequence alignment corroborate the phylogenetic analysis and emphasize the value of DNA barcoding in revealing evolutionary relationships among fish species.

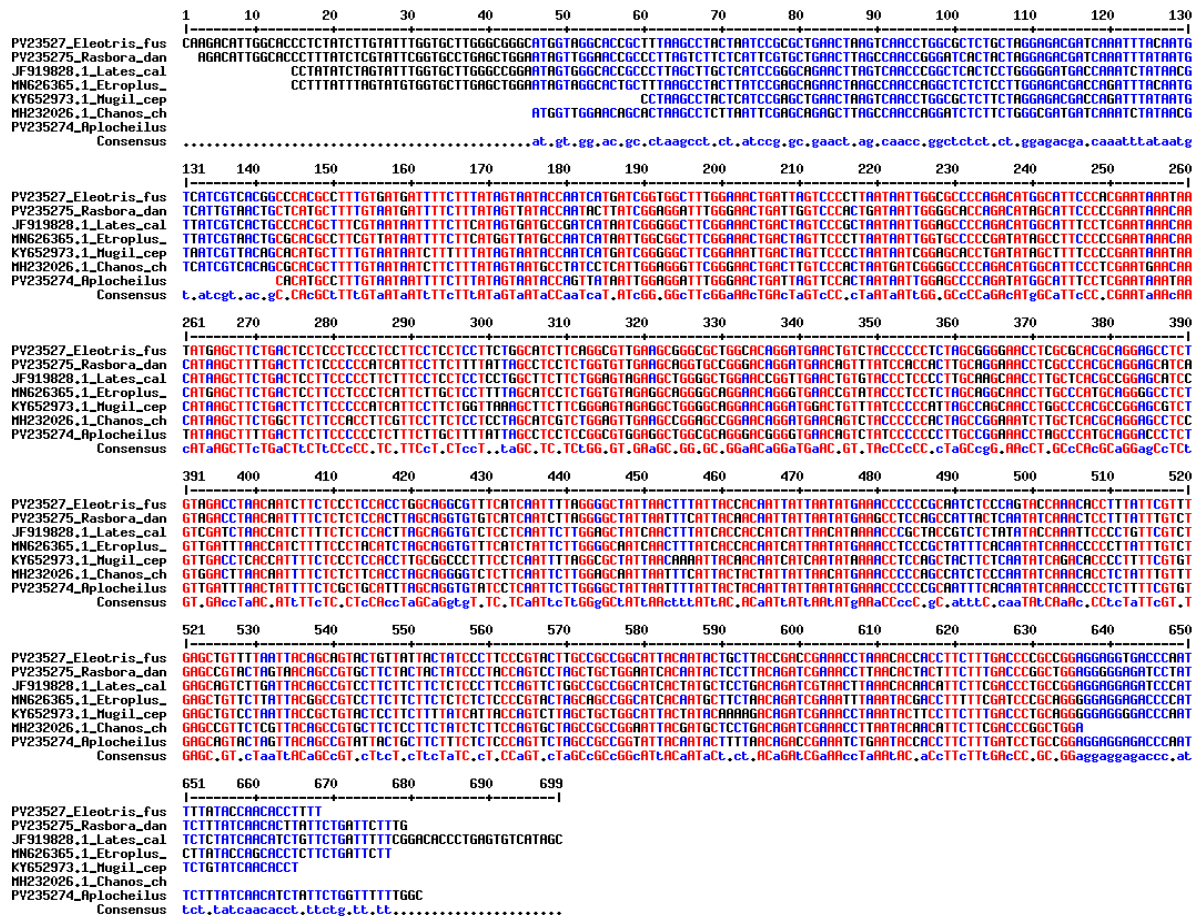


Figure 1. Multiple sequence alignment of mitochondrial COI gene sequences from selected teleost species.

The alignment shows nucleotide sequence comparison among *Eleotris fusca*, *Rasbora daniconius*, *Lates calcarifer*, *Etroplus suratensis*, *Mugil cephalus*, *Chanos chanos*, and *Rhopileus* sp. Conserved regions are indicated by identical residues highlighted in red, while variable nucleotides are shown in blue. Dots in the consensus line represent identical bases, and gaps (dashes) indicate insertions or deletions introduced to optimize alignment. The alignment demonstrates high sequence conservation across the species, with a few regions exhibiting variability suggestive of species-specific divergence.

Molecular Clock Test:

This test involved a comparison of the Maximum Likelihood (ML) values with and without molecular clock constraints, employing the kimura 2-parameter. Evolutionary rate variations among sites were accommodated, with the null hypothesis of equal evolutionary rate throughout the tree not being rejected at a 5% significance level (p=3.264E-001). The analytical procedure encompassed 7 coding nucleotide sequences using 1st, 2nd, 3rd and non-coding positions with 699 positions in the final dataset. Evolutionary analyses were conducted in MEGA12 utilizing up to 8 parallel computing threads.

	InL	Parameters	(+G)	(+I)
With Clock	-3321.959	7	n/a	n/a
Without Clock	-3318.489	12	n/a	n/a

Table 1: Result of Molecular clock test using ML method

Pairwise Alignment:

	1	2	3	4	5	6	7
1. PV23527 <i>Eleotris fusca</i>							
2. PV235274 <i>Aplocheilus lineatus</i>	0.246						
3. PV235275 <i>Rasbora daniconius</i>	0.246	0.222					
4. MH232026.1 <i>Chanos chanos</i>	0.256	0.248	0.216				
5. JF919828.1 <i>Lates calcarifer</i>	0.262	0.288	0.283	0.237			
6. MN626365.1 <i>Etroplus suratensis</i>	0.234	0.245	0.289	0.280	0.243		
7. KY652973.1 <i>Mugil cephalus</i>	0.264	0.280	0.275	0.313	0.293	0.280	

Pairwise genetic distances calculated from COI gene sequences ranged from 0.222 to 0.313. As expected, *Mugil cephalus* exhibited the greatest genetic distances to all other species (0.264 - 0.313), confirming its suitability as an out group. *Aplocheilus lineatus* and *Rasbora daniconius* showed the smallest genetic distance (0.222), corroborating their close relationship observed in the phylogenetic tree. *Eleotris fusca* displayed relatively large distances to all other species (0.234 - 0.264), consistent with its position as an early diverging lineage. The pairwise distances support the overall topology of the phylogenetic tree and provide a quantitative measure of the evolutionary divergence among the studied fish species.

Discussion:

The success of DNA barcoding for fish identification depends on robust genetic amplification and accurate reference sequence comparison. In this study, molecular identification through BLAST analysis of COI gene sequences effectively validated three dominant species from tarkarli: *Eleotris fusca*, *Aplocheilus lineatus*, and *Rasbora daniconius*. All sequences exhibited >98% similarity with GenBank reference entries, confirming species identity with high accuracy. Such precision is especially valuable in regions with overlapping habitats where morphological traits alone may be insufficient for accurate classification.^{8,9} The COI gene remains the most reliable barcode marker due to its universality and robust amplification efficiency across teleost species.^{10,11}

The two sampling sites represented contrasting aquatic environments: the lentic (Chorge wadi lake) and the lotic (Karli River). Differences in hydrological dynamics influence local biodiversity and often contribute to habitat-specific adaptation.^{12,13} *Eleotris fusca*'s identification from both systems suggests ecological tolerance,¹⁴ while *Aplocheilus lineatus* and *Rasbora daniconius* predominantly represent freshwater niches.¹⁵ These identifications, confirmed through molecular tools, provide baseline data for future environmental monitoring and ecological assessments.¹⁶

BLAST-based molecular identification is a cornerstone in modern barcoding workflows. Sequence similarity percentages greater than 98% are considered definitive for species-level classification.¹⁷ In this study, BLASTn analyses returned high confidence matches, validating sample identification and ensuring sequence reliability. Bioinformatic platforms such as BOLD and NCBI GenBank provide curated datasets that enhance global DNA-based taxonomic precision.^{18,19} The integration of AI-based sequence alignment tools further enhances efficiency in DNA-based taxonomy.²⁰

This study demonstrates that standard molecular workflows - CTAB extraction, PCR amplification, and Sanger sequencing - are sufficient for accurate fish identification from field-collected samples. Similar success rates have been reported across Indian river systems, emphasizing the reproducibility of these techniques.^{21,22} Furthermore, the availability of validated COI reference sequences facilitates expansion of DNA barcode libraries, contributing to biodiversity informatics and national conservation databases.^{23,24,25} Incorporating molecular data into ecosystem monitoring frameworks strengthens conservation decision-making, particularly for regions affected by anthropogenic pressure.²⁶

In summary, this work reaffirms the reliability of COI-based DNA barcoding for molecular identification of fish species in ecologically distinct environments. By focusing on BLAST analysis rather than phylogenetic reconstruction, the study emphasizes rapid and practical application of molecular techniques for biodiversity documentation. Future directions should involve the integration of next-generation sequencing (NGS) and eDNA-based surveillance to capture broader community-level diversity.^{27,28} Such molecular frameworks will be essential for long-term monitoring, ecological forecasting, and management of freshwater-marine transitional ecosystems in coastal India.

References:

1. Panja S, Podder A, Homechaudhuri S. Evaluation of aquatic ecological systems through dynamics of ichthyofaunal diversity in a Himalayan torrential river, Murti. *Limnologica*. 2020;82(March):125779. doi:10.1016/j.limno.2020.125779
2. Walton LN, Quindazzi M, Gauthier S, Stevens C. Fish ID face-off: A comparison of genetic barcoding and otolith shape analysis for streamlining species identification of mesopelagic fishes. *Fish Res*. 2025;281(May 2024). doi:10.1016/j.fishres.2024.107254
3. Arif IA, Khan HA, Shobrak M, Williams J. Cytochrome c oxidase subunit I barcoding of the green bee-eater (*Merops orientalis*). *Genet Mol Res*. 2011;10(4):3992-3998. doi:10.4238/2011.October.21.2
4. Theissinger K, Fernandes C, Formenti G, et al. How genomics can help biodiversity conservation. *Trends Genet*. 2023;39(7):545-559. doi:10.1016/j.tig.2023.01.005
5. Trivedi S. Role of DNA barcoding in marine biodiversity assessment and conservation : An update. *Saudi J Biol Sci*. 2016;23(2):161-171. doi:10.1016/j.sjbs.2015.01.001
6. Quiñones KJO, Gentallan RP, Timog EBS, et al. Liquid-nitrogen-free CTAB DNA extraction method from silica-dried specimens for next-generation sequencing and assembly. *MethodsX*. 2024;12(May):10-17. doi:10.1016/j.mex.2024.102758
7. Gusmao ACB, van Dijk RL, Girard EB, et al. Exploring the potential of the COI gene marker for DNA barcoding of planktonic foraminifera. *Sci Rep*. 2025;15(1):1-10. doi:10.1038/s41598-025-03842-7
8. Panprommin D, Soontornprasit K, Tuncharoen S, Pithakpol S, Keereelang J. DNA barcodes for the identification of species diversity in fish from Kwan Phayao, Thailand. *J Asia-Pacific Biodivers*. 2019;12(3):382-389. doi:10.1016/j.japb.2019.05.003
9. Kundu S, Chandra K, Tyagi K, Pakrashi A, Kumar V. DNA barcoding of freshwater fishes from Brahmaputra River in Eastern Himalaya biodiversity hotspot. *Mitochondrial DNA Part B Resour*. 2019;4(2):2411-2419. doi:10.1080/23802359.2019.1637290
10. Nahavandi R. Fish DNA barcoding: A review. *Int J Biol Reports*. 2024;2(2):17-33.
11. Jennings WB, Ruschi PA, Ferraro G, et al. Barcoding the Neotropical freshwater fish fauna using a new pair of universal COI primers with a discussion of primer dimers and M13 primer tails. *Genome*. 2019;62(2):77-83. doi:10.1139/gen-2018-0145
12. Uno H, Utsumi S, Morita K, Kishida O, Alam MK, Negishi J. Hydrological Connectivity and Local Environment Alternately Drive Spatial Structure of Floodplain Aquatic Community Across Seasons. *Ecol Evol*. 2025;15(2). doi:10.1002/ece3.70880
13. Qiu J, Yuan S, Tang H, Deng Y, Gualtieri C. Hydrodynamic Responses of Fish Community Dynamics in Large-Scale Morphologically Complex River Systems. *Water Resour Res*. 2025;61(7). doi:10.1029/2025WR040690
14. Mennesson MI, Tabouret H, Pécheyran C, Feunteun E, Keith P. Amphidromous life cycle of *Eleotris fusca* (Gobioidei: Eleotridae), a widespread species from the Indo-Pacific studied by otolith analyses. *Cybium*. 2015;39(4):249-260.
15. Darshan A, Abujam S, Das DN. *Systematic Account*. Vol 37.; 2019. doi:10.1016/b978-0-12-815556-1.00004-2
16. Çevik T, Çevik N. Environmental DNA (eDNA): A review of ecosystem biodiversity detection and applications. *Biodivers Conserv*. 2025;34(9):2999-3035. doi:10.1007/s10531-025-03112-y
17. Kumar G, Reaume AM, Farrell E, Gaither MR. Comparing eDNA metabarcoding primers for assessing fish communities in a biodiverse estuary. *PLoS One*. 2022;17(6 June):1-20. doi:10.1371/journal.pone.0266720
18. Baena-Bejarano N, Reina C, Martínez-Revelo DE, et al. Taxonomic identification accuracy from BOLD and GenBank databases using over a thousand insect DNA barcodes from Colombia. *PLoS One*. 2023;18(4 4):1-19.

doi:10.1371/journal.pone.0277379

19. Alam M, Abbas K, Usmani N, et al. A Comprehensive review on DNA barcoding for species identification across diverse taxa. *Munis Entomol Zool J.* 2024;19(2):1057-1072. <https://www.munisentzool.org/>
20. Nanni L, Cuza D, Brahmam S. AI-Powered Biodiversity Assessment: Species Classification via DNA Barcoding and Deep Learning. *Technologies.* 2024;12(12):1-15. doi:10.3390/technologies12120240
21. Kawato M, Yoshida T, Miya M, et al. Optimization of environmental DNA extraction and amplification methods for metabarcoding of deep-sea fish. *MethodsX.* 2021;8:101238. doi:10.1016/j.mex.2021.101238
22. Rathipriya A, Karal Marx K, Jeyashakila R. Molecular identification and phylogenetic relationship of flying fishes of Tamil Nadu coast for fishery management purposes. *Mitochondrial DNA Part A DNA Mapping, Seq Anal.* 2019;30(3):500-510. doi:10.1080/24701394.2018.1558220
23. Jabot F, Auger G, Bonnal P, et al. Use of massive DNA barcoding to monitor biodiversity: A test on forest soil macrofauna. *For Ecol Manage.* 2025;595(July). doi:10.1016/j.foreco.2025.123004
24. Zhou Y, Trujillo-González A, Nicol S, Huerlimann R, Sarre SD, Gleeson D. Evaluation of DNA barcoding reference databases for marine species in the western and central Pacific Ocean. *PeerJ.* 2025;13:1-25. doi:10.7717/peerj.19674
25. Shelake SK, Deshpande VY. Dna Barcoding of Fish Diversity From Indian Riverine Ecosystem: a Review Uttar Pradesh Journal of Zoology Dna Barcoding of Fish Diversity From Indian Riverine Ecosystem: a Review. 2021;(January). <https://www.researchgate.net/publication/357524896>
26. Nielsen ES, Hanson JO, Carvalho SB, et al. Molecular ecology meets systematic conservation planning. *Trends Ecol Evol.* 2023;38(2):143-155. doi:10.1016/j.tree.2022.09.006
27. Chakraborty M, Ghosh SK. An assessment of the DNA barcodes of Indian freshwater fishes. *Gene.* 2014;537(1):20-28. doi:10.1016/j.gene.2013.12.047
28. Derocles SAP, Bohan DA, Dumbrell AJ, et al. Biomonitoring for the 21st Century: Integrating Next-Generation Sequencing Into Ecological Network Analysis. *Adv Ecol Res.* 2018;58:1-62. doi:10.1016/bs.aacr.2017.12.001