Further specimens and phylogenetic position of the recently described catfish species *Mystus prabini* (Siluriformes: Bagridae)

Shantanu KUNDU, Vikas KUMAR, Kaomud TYAGI, Kailash CHANDRA

Centre for DNA Taxonomy, Molecular Systematics Division, Zoological Survey of India, M Block, New Alipore, Kolkata-700053, India.

Corresponding author: *E-mail: vikasszi77@gmail.com*

**Abstract**

The study generated six DNA barcode data of the recently described catfish, *Mystus prabini* from different localities in northeast India. Further, we assess the phylogenetic position of *M. prabini* using partial sequence data of the mitochondrial cytochrome c oxidase subunit I (cox1) gene to confirm its genetic distinctiveness. Mitochondrial data strongly suggest a sister relationship of *M. prabini* with *M. bleekeri*, *M. ngasep*, and *M. cineraceus*. According to the genetic intimacy of *M. pabini* with *M. bleekeri*, we presumed the possible gene-flow within Ganges and Brahmaputra River basins.

**Keywords:** Freshwater Fish, *Mystus*, Ganges, Brahmaputra, Molecular Phylogeny.

**Zoobank:** urn:lsid:zoobank.org:pub:610B63ED-1FCF-4594-8840-5950EA83F107

**Introduction**

Catfishes of the genus *Mystus*, belonging to the family Bagridae, occur in fast flowing freshwaters in Asia and Southeast Asia (Ng 2004). This small to medium-size fishes is comprised 46 valid species, among them 21 species are found in India (Eschmeyer 2019). The members of *Mystus* are mainly differentiated from other Bagrids by the osteological features of the metapterygoid and lachrymal (Hollister 1934; Mo 1991). The *Mystus* species possess ornamental value; hence, these fishes are highly exploited from different riverine systems in their range distribution (Dhar and Ghosh 2015). Most of the *Mystus* species are categorized as least concern, whereas two species of *M. bocourti* and *M. malabaricus* are vulnerable and near threatened, respectively (IUCN 2019). Several taxonomic studies have been aimed to determine the species diversity of *Mystus* from India and other countries (Ng and Pethiyagoda 2013; Plamoottil and Abraham 2014). In recent past, a new species namely *M. prabini* has been discovered from Brahmaputra river basin in Arunachal Pradesh, Northeast India (Darshan et al. 2019).

Both morphological and molecular approaches have been aimed to identify and discriminate the species from few congeners. However, the actual diversity and genetic distinctiveness of *M. prabini* is still anonymous as comparison with other species. The partial segment of cytochrome c oxidase subunit 1 (cox1) gene in mitochondrial region has been largely used in the taxonomic assessment of fishes (Hebert et al. 2003; Laskar et al. 2013, 2018a, b). Further, the DNA barcode data of *Mystus* species has been largely generated from different geographical regions, including India. Till date more than 321 DNA sequences of partial cox1 gene are available in GenBank, among them 306 are labeled up to species level and remaining 15 are unclassified. Further, many barcode data are unreliable due to the taxonomic mistake and incorrect sequence handling. In the present study, we aimed to generate the DNA barcode data of the taxonomically identified recently described species, *M. prabini* from different localities in the northeast India. We also acquired the verified barcode data of all available *Mystus* congeners from GenBank and analyzed them through genetic distances and cluster analysis to reevaluate the monophyletic criteria for species-level identification. In addition, the present effort will facilitate to screen the actual diversity of *M. prabini* in the northeast of India.
Material and Methods
A total of six specimens of *M. prabini* were collected from its type locality and different riverine systems of the Brahmaputra basin in the northeast of India. The collection localities were mentioned in the GenBank and also pointed in the map (Fig. 1). Morphological identification of the collected specimens was confirmed by the described taxonomic characters (Darshan et al. 2019). The voucher specimens were stored in National Zoological Collections of Zoological Survey of India, Kolkata. Total genomic DNA extraction, Polymerase chain reaction (PCR), and purification of the PCR products of each sample were done as per the published laboratory protocol (Sambrook et al. 1989; Ward et al. 2005). The Veriti® Thermal Cycler (Applied Bio systems, Foster City, CA) was used for PCR and ABI-3730 DNA analyzer was used for Sanger Sequencing as per standard protocol. Both forward and reverse chromatograms (>600 bp) of the partial mitochondrial cytochrome c oxidase subunit 1 (mtCOI) gene were obtained to build the consensus sequence for each studied sample. The online tool, nucleotide BLAST (https://blast.ncbi.nlm.nih.gov) and NCBI-ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/) were used to confirm the alignment and amino acid array. Total 46 mtCOI sequences of the same and congeners of *Mystus* were acquired from the public database (GenBank) to make a final dataset. The database sequence of *Apristurus ampliceps* (Carcharhiniformes) (Accession No. EU398546) was also acquired from GenBank and used as an out-group. The alignment of the studied dataset was performed by ClustalX (Thompson et al. 1997). The genetic distances and maximum-likelihood (ML) topology with 1000 bootstrap support was estimated by MEGAX (Kumar et al. 2018). The best fit model for this dataset ‘HKY+G’ was estimated using MEGAX with lowest BIC value=6915.456.

Results and Discussion
The generated DNA barcode sequences of *M. prabini* showed 99-100% similarity with the publicly available sequences of *Mystus* sp. in GenBank. The database sequences of these newly discovered species are nameless and needed to be updated as *M. prabini*. The overall mean K2P genetic distances of the studied dataset (22 species) was 16.2%. The intra-species genetic distance was ranging from 0% to 1.19%, however the inter-species genetic distance was ranging from 6.43% to 29.15% in the present dataset. All the studied species were clearly discriminated by sufficient K2P genetic distances and revealed monophyletic clustering in ML tree (Fig. 1). The northeast Indian species, *M. prabini* is closely clustered with *M. ngasep* and *M. cineraceus* with 11.80% genetic distances. Further, *M. prabini* also showed lowest genetic distance (9.67%) with *M. bleekeri* (Ganges River Basin, India) in comparison with other congeners. Nevertheless, *M. cineraceus* was discovered from

<table>
<thead>
<tr>
<th>GenBank Accession No</th>
<th>Collection localities</th>
<th>Coordinates</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MN520224</td>
<td>Dibang River, Arunachal Pradesh</td>
<td>28.15N 95.68E</td>
<td>This Study</td>
</tr>
<tr>
<td>MN520225</td>
<td>Siang River, Arunachal Pradesh</td>
<td>28.08N 95.34E</td>
<td>This Study</td>
</tr>
<tr>
<td>MN520226</td>
<td>Dibru River, Assam</td>
<td>27.58N 95.34E</td>
<td>This Study</td>
</tr>
<tr>
<td>MN520227</td>
<td>Dhansiri River, Assam</td>
<td>26.65N 93.71E</td>
<td>This Study</td>
</tr>
<tr>
<td>MN520228</td>
<td>Diphlu River, Assam</td>
<td>26.61N 93.43E</td>
<td>This Study</td>
</tr>
<tr>
<td>MN520229</td>
<td>Kameng River, Assam</td>
<td>26.72N 92.85E</td>
<td>This Study</td>
</tr>
<tr>
<td>KY290074</td>
<td>Diphlu River, Assam</td>
<td>-</td>
<td>Laskar et al. 2018a</td>
</tr>
<tr>
<td>MH175199</td>
<td>Sinkin River, Arunachal Pradesh</td>
<td>-</td>
<td>Darshan et al. 2019</td>
</tr>
<tr>
<td>MH175200</td>
<td>Sinkin River, Arunachal Pradesh</td>
<td>-</td>
<td>Darshan et al. 2019</td>
</tr>
<tr>
<td>MH175201</td>
<td>Sinkin River, Arunachal Pradesh</td>
<td>-</td>
<td>Darshan et al. 2019</td>
</tr>
<tr>
<td>MH175202</td>
<td>Sinkin River, Arunachal Pradesh</td>
<td>-</td>
<td>Darshan et al. 2019</td>
</tr>
<tr>
<td>MG582180</td>
<td>Sinkin River, Arunachal Pradesh</td>
<td>-</td>
<td>Darshan et al. 2019</td>
</tr>
</tbody>
</table>

Table 1. Collection details of the generated and GenBank sequences of *Mystus prabini* in the present study.
Irrawaddy River drainage of Myanmar (Ng and Kottelat 2009), and *M. ngasep* was discovered from Chindwin drainage from the Manipur state in northeast India (Darshan et al. 2011). Both the valid species showed negligible genetic distances (0%) with their named DNA data available in the GenBank, which need further investigation through molecular approaches with their type or actual topotype specimens.

The DNA based investigation is the most flourishing research in systematics and biodiversity studies (Hajibabaei et al. 2007). However, before aiming the DNA based investigation for addressing any kind of biological questions; the species identification should be reliable (Collins and Cruickshank 2013). The generated DNA sequences with other collateral data in the global database (GenBank and BOLD) provides a trustworthy platform for speedy and authentic species identification (Shen et al. 2013). Further, the taxon sampling from wide geographical regions and their DNA barcodes provides better scenario on their diversity (Bergsten et al. 2012).

So far, many small to large scale effort has been made to explore the ichthyofaunal diversity in India including northeastern region (Khedkar et al. 2014; Lakra et al. 2016; Barman et al. 2018; Kundu et al. 2019a, 2019b).
b, c). This present effort aids species-specific survey of other important fish species to substantiate their actual range distribution and status. In the original description of *M. prabini*, the authors were compared with eight species through molecular approaches (Darshan et al. 2019) and evidenced to be close with *M. bleekeri* (8.6%) and *M. albolineatus* (13.9%). However, the present analysis with 22 species illuminated more clear understanding of this species. Our new record from different riverine systems in Assam state corroborates the further range extension of this species and provides scope to revise the check-list of ichthyofauna locally. Although, the newly discovered species is spotted from the Brahmaputra River Basin, but the genetic intimacy with *M. bleekeri*, distributed in Ganges River Basin, India presumed gene flow within these two river basins.

**Acknowledgments**

We thank to the Director, Zoological Survey of India, Ministry of Environment, Forest and Climate Change (MoEF & CC) for providing necessary working facilities. The first author (S.K) acknowledges the fellowship grant received from Council of Scientific and Industrial Research (CSIR) Senior Research Associateship (Scientists’ Pool Scheme) Pool No. 9072-A.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This study is funded by the core funding of Zoological Survey of India, Kolkata, Ministry of Environment, Forest and Climate Change (MoEF&CC), New Delhi. The funders had no role in study design, data collection and analysis or preparation of the manuscript.

**Literature cited**


