Molecular characterization of *Channa* species from Bangladesh based on Cytochrome c Oxidase Subunit I (COI) gene

Md. Sagir AHMED\(^1\), Sabrina Rahman DINA\(^1\), Luthfun NAHAR\(^1\), Nafisa Nawal ISLAM\(^2\), Hasan Al REZA\(^3\)

\(^1\)Department of Zoology, University of Dhaka, Dhaka 1000, Bangladesh.
\(^2\)Department of Biotechnology and Genetic Engineering, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh.
\(^3\)Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh.

**Corresponding author:** *E-mail: sagir@du.ac.bd*

**Abstract**

The Snakehead fishes have significant economic importance in Bangladesh as food due to their exquisite taste. Some species of snakeheads are considered in threatened categories due to their population decline and loss of habitats. DNA Barcoding has been considered as a global bio-identification system for freshwater and marine fishes in recent years. In the present study, four snakehead species: *Channa gachua*, *C. marulius*, *C. punctata* and *C. striata* were investigated using both morphological methods and molecular identification with mitochondrial Cytochrome c Oxidase subunit I (COI) gene. Our study confirms that no *C. orientalis* existing in Bangladesh and all the older reports are pertaining to *C. gachua*. Phylogenetic relationships, the divergence of sequences between species was determined using Kimura 2-parameter distance model, and a Maximum Likelihood tree was generated in which the COI sequences for each of the four species could clearly be discriminated. The average interspecies genetic divergence was 26.90%. The study strongly validated the efficacy of COI as an ideal marker for DNA barcoding of Bangladeshi freshwater fishes and identification of specific species that necessitate further taxonomic study.

**Keywords:** Snakeheads, COI gene, Molecular identification, Phylogeny, Bangladesh.

**Zoobank:** urn:lsid:zoobank.org:pub:79A84D65-7745-44FE-B4FF-B3673D448605

**Introduction**

The members of family Channidae commonly known as the snakeheads are native to Asia. Snakehead fishes are important candidates for aquaculture in Bangladesh. The annual snakehead fish catch of all rivers of Bangladesh was 0.21% (311 metric tons) (BBS 2016). There are five known species of *Channa* in Bangladesh, viz. *C. punctata* (Bloch, 1793) (Spotted snakehead), *Channa striata* (Bloch, 1793) (Striped Snakehead), *C. marulius* (Hamilton, 1822), (Great snakehead), *C. gachua* (Hamilton, 1822), and *C. barca* (IUCN 2015; Rahman 2005). Among them, *C. barca* has been reported as critically endangered (IUCN 2015). Previously, snakehead fishes of Bangladesh were identified by morphometric and meristic methods (Rahman 2005). However, it is difficult to identify snakeheads on the basis of morphological and meristic analysis only, especially during their juvenile stages. Due to the lack of taxonomic expertise molecular analysis like DNA barcoding can offer a rapid and precise method of identification of snakeheads (Serrao et al. 2014). Genetic calibration of snakehead fish diversity can significantly help taxonomic resolution within the group and interspecies differences. Moreover, proper identification can assist in managing snakehead fishes for long-term sustainable conservation.

The rapid and accurate identification of closely related fish species is essential for scientific research and artificial breeding. DNA barcoding offers the opportunity for a standardized system of species identification based on the analysis of small fragments of DNA (Hebert et al. 2004). DNA barcoding is a promising technique for species identification that uses a short mitochondrial DNA sequence of Cytochrome c Oxidase I (COI) gene. Mitochondrial COI region is appropriate for discriminating between closely related species across diverse animal phyla, and this technique has been used for marine and freshwater fishes (Ward et al. 2005). A comparison of mitochondrial sequences from 2238 species in 11 animal phyla showed that 98% of closely related species pairs had more than 2% sequence difference, which is enough for successful identification of most species (Stoeckle and Hebert 2008). There have been various studies investigating the mitochondrial sequence variations in the snakehead fishes occurring within a particular region of the world (Adamson et al. 2010; Aquilino et al. 2011;
Four snakehead species of Bangladesh viz. *C. marulius*, *C. gachua*, *C. punctata*, *C. striata* were investigated using both morphological methods and molecular identification with mitochondrial Cytochrome C Oxidase subunit I (COI) gene. The primary aim of this study is to generate comprehensive mtDNA sequences to distinguish the aforementioned freshwater fish species, compare conspecific populations and demonstrate its value for conservation research.

**Material and Methods**

**Sample collection:** The fish specimens were collected from the Tanguar Haor, Sunamganj (25.07N, 19.3E) from the local fishermen’s catch. These specimens were immediately frozen and brought to the Advanced Fisheries and DNA Barcoding Laboratory, Department of Zoology, University of Dhaka for further studies.

**Morphological Analysis:** Morphometric study was done by measuring the length and width of various parts of the body. Meristic characteristics were analyzed by studying lateral lines count, scale count and fin count (Talwar 1991; Barman et al. 2014).

**DNA extraction and amplification:** For DNA extraction, approximately 20 mg (5 mm) tissue was cut and chopped and homogenized gently in 500 μl TES buffer (200 mM Tris, 100 mM EDTA, 250 mM NaCl). Then 10 μl of 20 mg/ml of Proteinase K was added and incubated at 56°C overnight. An equal volume of Phenol: Chloroform: Isoamyl alcohol (25:24:1) was added and centrifuged for 10 minutes at 12000 rpm. Then the upper phase was transferred to a new tube and added to an equal volume of Chloroform: Isoamyl alcohol and centrifuged at 12000 rpm for 10 minutes. The upper aqueous layer was transferred to a fresh, sterilized micro-centrifuge tube, and double volume of chilled absolute Ethanol was added. After overnight precipitation at -20°C, the above sample was centrifuged at 10,000 rpm for 10 minutes, and the pellet was retained. 500 μl of 70% ethanol was added and again centrifuged at 7000 rpm for 10 minutes. The pellet was air-dried and resuspended in 50 μl TE buffer (Chowdhury et al. 2016).

The COI gene of the snakehead fishes was successfully amplified using the universal fish primer named as FishF1 and FishR1 (Ward et al. 2005). The sequences of the primers are: FishF1 (5’TCAACCAACCACAAAGACATTGGCAC3’) and FishR1 (5’TAGACTTCTGGGTGGCAGACATC3’). Each 25 μL PCR reaction consisted of 2 μL of extracted DNA template, 12.5 μl of GoTaq® G2 Hot Start Colorless Master mix (Ref. M743A, Promega, Madison, WI USA), 1 μl (0.01 mM) of reverse primer, 1 μl (0.01 mM) of reverse primer, and 8.5 μl of nuclease free water (Ref. P119A, Promega, Madison, WI USA). The PCR thermal cycling conditions involved an initial denaturation of 95°C for 2 min, followed by 30 cycles of [denaturation at 95°C for 30 s, annealing at 54°C for 40 s and extension at 72°C for 1 min], with a final extension at 72°C for 7 min. However, an annealing temperature of 52°C was used for *C. marulius*.

**Sequencing, phylogenetic and statistical analyses:** The purified PCR products were sent to First BASE Laboratories SdnBhd, Malaysia for sequencing. The sequencing data from the chromatogram was curated and converted into FASTA format. A BLAST of the COI sequences was done at the nucleotide database of National Center for Biotechnology Information (NCBI) to determine the best match homology. The barcodes determined in this study were deposited in the GenBank/EMBL database under the Accession No. KT762386.1, KT762387.1, KX808573.1 and KT364793.1 (*C. punctate, C. striata, C. marulius, C. gachua*). Evolutionary analyses of the aligned sequences were conducted in the program MEGA7 (Kumar et al. 2016). The Phylogenetic trees were rooted using *Parachanna obscura* (Accession No.: HM882959.1, HM880234.1, KJ937453.1) and *P. insignis* (Accession No.: KJ937444.1, KJ937415.1, KJ937413.1) as outgroups. A total of 35 sequences from the *Channa* genus (Accession No.: MF496899.1, MF496935.1, KJ936643.1, KY425545.1 etc.) were also included in the study as a reference. The evolutionary history was inferred using the Maximum Likelihood method (Nei and
Ahmed et al. - Molecular characterization of *Channa* species from Bangladesh

Table 1. Mean morphometric and meristic characters of *Channa* species.

<table>
<thead>
<tr>
<th>Parameters (cm)</th>
<th><em>C. punctata</em></th>
<th><em>C. striata</em></th>
<th><em>C. marulius</em></th>
<th><em>C. gachua</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total length (TL)</td>
<td>30</td>
<td>25.5</td>
<td>33.4</td>
<td>15.4</td>
</tr>
<tr>
<td>Standard length (SL)</td>
<td>26.4</td>
<td>21.4</td>
<td>29</td>
<td>12.7</td>
</tr>
<tr>
<td>Head length (HL)</td>
<td>6</td>
<td>4.8</td>
<td>7.3</td>
<td>3.8</td>
</tr>
<tr>
<td>Body depth (BD)</td>
<td>2.3</td>
<td>1.9</td>
<td>2.6</td>
<td>1.6</td>
</tr>
<tr>
<td>Eye diameter (ED)</td>
<td>0.6</td>
<td>0.4</td>
<td>0.7</td>
<td>0.5</td>
</tr>
<tr>
<td>Dorsal fin base (DFB)</td>
<td>18</td>
<td>12.3</td>
<td>6.7</td>
<td>6</td>
</tr>
<tr>
<td>Pectoral fin base (PD)</td>
<td>1.3</td>
<td>0.9</td>
<td>1.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Pelvic fin base (PL)</td>
<td>0.9</td>
<td>0.4</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Anal fin base (AFB)</td>
<td>10.7</td>
<td>15.6</td>
<td>11.2</td>
<td>4.6</td>
</tr>
<tr>
<td>Caudal length (CL)</td>
<td>3.6</td>
<td>4.1</td>
<td>4.4</td>
<td>2.7</td>
</tr>
<tr>
<td>Scales between the orbit and the pre orbit</td>
<td>5</td>
<td>9</td>
<td>9-10</td>
<td>5</td>
</tr>
<tr>
<td>Scales between snout and base of the dorsal</td>
<td>11</td>
<td>15</td>
<td>17</td>
<td>11-12</td>
</tr>
<tr>
<td>Scales with lateral line</td>
<td>40-41</td>
<td>54-60</td>
<td>54-64</td>
<td>42-45</td>
</tr>
<tr>
<td>Number of scales lateral line first passes with</td>
<td>15</td>
<td>16-18</td>
<td>18-19</td>
<td>12-13</td>
</tr>
<tr>
<td>Number of scales lateral line descends to</td>
<td>1 row</td>
<td>2-3rows</td>
<td>2rows</td>
<td>1row</td>
</tr>
<tr>
<td>Dorsal fin ray</td>
<td>30</td>
<td>44</td>
<td>52</td>
<td>34</td>
</tr>
<tr>
<td>Pectoral fin ray</td>
<td>18</td>
<td>15</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>Pelvic fin ray</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 2. Accession No. and BLAST results of *Channa* species.

<table>
<thead>
<tr>
<th>Query sequence IDs</th>
<th>Accession no.</th>
<th>Percent similarity</th>
<th>E-value</th>
<th>Accession no. of the best match</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. punctata</em></td>
<td>KT762386.1</td>
<td>99%</td>
<td>0.0</td>
<td>KJ937392.1</td>
</tr>
<tr>
<td><em>C. striata</em></td>
<td>KT762387.1</td>
<td>99%</td>
<td>0.0</td>
<td>KJ936901.1</td>
</tr>
<tr>
<td><em>C. marulius</em></td>
<td>KX808573.1</td>
<td>99%</td>
<td>0.0</td>
<td>KX389281.1</td>
</tr>
<tr>
<td><em>C. gachua</em></td>
<td>KT364793.1</td>
<td>99%</td>
<td>0.0</td>
<td>MF496784.1</td>
</tr>
</tbody>
</table>

Kumar 2000). Pairwise genetic distance, intraspecies and interspecies distances were calculated using the Kimura 2-parameter distance model (Kimura 1980).

Results

**Morphological study:** The species were first identified by size, color, shape, fin ray count and other morphometric and meristic characters (Table 1).

**DNA extraction and purified PCR product analysis:** Molecular study was done through DNA barcoding utilizing COI gene of mtDNA. The extracted DNA from *C. marulius, C. gachua, C. punctata,* and *C. striata* were amplified by PCR using COI gene-specific primer which had product size of around 658bp. Amplified PCR products were resolved and detected in 1% (w/v) agarose gel.

**BLAST analysis:** BLAST analysis revealed that all the COI gene sequences of four species have similarity (99%) with the respective sequences in the mitochondrial region in the GenBank database (Table 2).

**Phylogenetic and genetic divergence analysis:** The sequence analysis revealed that the overall GC content was 48.6% (S.E.=0.196), with *C. argus* having the highest (51.5%) and *C. striata* having the lowest (46.3%). All the species exhibited unique barcodes, and all individuals within species and genus formed cohesive clusters. The mean interspecies genetic divergence was 27.3%, while the mean intraspecies divergence was only 1.2% indicating the discrimination power of the barcodes.

Aligned sequences were used to generate K2P-distances to infer a Maximum Likelihood phenogram to generate a pictorial representation of the barcode variation among and between species, with bootstrap analysis (based on 1000 replications) (Felsenstein 1985; Kimura 1980; Kumar et al. 2016; Nei and Kumar 2000). Phylogenetic
analysis showed the interspecies genetic divergence between *C. marulius*, *C. gachua*, *C. punctata* and *C. striata*, were clustered together in agreement with their taxonomic classification at the species level (Fig. 1). The Maximum Likelihood tree showed that *C. punctata* shares the same node with *C. marulius*, this means that they share variations obtained from a common ancestor. *Channa striata* shares a node with common ancestor of *C. punctata* and *C. marulius*, indicating it diverged at an earlier time. Finally, *C. gachua* and *C. barca* are the most distantly related compare to the aforementioned species indicated by the distinct cluster (Fig. 1). So, *C. punctata*, *C. marulius*, *C. striata*, *C. gachua* can be differentiated using DNA barcoding.

Figure 1. Phylogenetic tree of COI gene sequence of *Channa marulius*, *C. gachua*, *C. punctata*, and *C. striata* with outgroups. The Maximum Likelihood (ML) tree displaying inter-specific variation was constructed using K2P model.
Ahmed et al.- Molecular characterization of *Channa* species from Bangladesh

**Table 2.** The genetic divergence (K2P Distance %) within the species and between the species.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>No. of sequences</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>S.E.</th>
<th>Overall mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraspecies</td>
<td>47</td>
<td>1.2</td>
<td>0.00</td>
<td>14.1</td>
<td>0.648</td>
<td>26.9</td>
</tr>
<tr>
<td>Interspecies</td>
<td></td>
<td>27.3</td>
<td>7.4</td>
<td>52.8</td>
<td>0.448</td>
<td></td>
</tr>
</tbody>
</table>

*Standard Error

**Discussion**

In the past, morphometric and meristic characteristics were the only tools used for inferring fish phylogenetic relationship, and to understand speciation (Musikasinthorn 2000). But it is difficult to differentiate the fishes, especially Channidae species because of the similarity in their external morphology (Khan et al. 2013; Miyan et al. 2014). Therefore, the construction of phylogenetic trees based on only morphology is controversial due to the complex evolutionary changes in their morphological and physiological characters. A large variety of snakeheads are available in Bangladesh, and their identification based on morphological characteristics alone is highly ambiguous. There are many reported cases of mislabeling Channidae species all over the world (Galal-Khallaf et al. 2014; Nagalakshmi et al. 2016; Yan et al., 2016).

In our study, *Channa* species were initially identified based on morphological characteristics, and for resolving the taxonomic problems, a molecular study has also been done considering the efficacy of DNA barcoding in the identification of genus *Channa*. Therefore, this study shows that the result obtained from the morphometric, meristic and molecular analysis is similar, and thus *C. marulius*, *C. gachua*, *C. punctata* and *C. striata* are different species that can be differentiated under the same genus. Phylogenetic relationship among *Channa* species observed in this study was quite similar as described by the previous authors (Barman et al. 2018; Conte-Grand et al. 2017).

As *C. barca* is rare and not seen in the fishermen’s catch, and also ranked as critically endangered in the Red List of IUCN Bangladesh (IUCN 2015), this species could not be included in the study. *Channa orientalis* is an endemic species of Sri Lanka, and it is often confused with *C. gachua* (Tanomtong et al. 2014). The major morphological difference between the two species is that *C. gachua* has ventral fins and *C. orientalis* does not. They also differ in their breeding behavior (Ng et al. 1999). We could not find any specimen of *C. orientalis*, and this led us to believe that *C. orientalis* not existing in Bangladesh and all the older reports are pertaining to *C. gachua* (Conte-Grand et al. 2017).

The bootstrap values were not convincing, but this problem can be easily remedied by adding more sequences from the *Channa* species. The identification of four snakehead fishes *C. marulius*, *C. striata*, *C. punctata* and *C. gachua* was done through morphometric, meristic and molecular study. From this study, it is possible to produce the molecular database that assembles the molecular characteristics of these species distributed in Bangladesh. These sequences can also be used in further studies to determine the genetic divergence of *Channa* species distributed in different geographic locations. Bangladesh has to assist the management of their population in wild to prevent them from being extinct. Fishermen and local people should be made aware to conserve these species. Hence, molecular identification will help fish experts to understand adaptive radiation of these fishes and produce hybrid species in the near future.

**Acknowledgments**

We acknowledge the financial support from the Ministry of Education, Government of the People’s Republic of Bangladesh under the grant for Advanced Research in Sciences (2015-2016).
Literature cited


Ahmed et al.- Molecular characterization of *Channa* species from Bangladesh

investigation using DNA barcoding. Food Control 59: 196-200. (doi:10.1016/j.foodcont.2015.05.018)


