Efficiency of Universal Barcode Gene (Coxi) on Morphologically Cryptic Mugilidae Fishes Delineation

1C. Prasanna Kumar, 2B. Akbar John, 3S. Ajmal Khan, 1P.S. Lyla, 1S. Murugan, 4M. Rozihan and 5K.C.A. Jalal
1Centre of Advanced Study in Marine Biology, Annamalai University, Parangipettai, Tamil Nadu, India
2Institute of Oceanography and Maritime Studies, International Islamic University Malaysia, Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia
3Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

Corresponding Author: P.S. Lyla, Centre of Advanced Study in Marine Biology, Annamalai University, Parangipettai, Tamil Nadu, India

ABSTRACT

An effort was made to assess the utility of 650 bp partial Cytochrome C oxidase subunit I (DNA barcode) gene in delineating the members of taxonomically ambiguous marine fin fishes (Family: Mugilidae). To address the issue we used all the 95 barcode sequences of Mugilidae family available at NCBI (National Centre for Biotechnological Information) along with the barcode data generated from Mugilidae fishes of Parangipettai coastal waters. The average GC content of Mugilidae was found to be 46.46%. Crenimugil crenilabis showed less GC content (44.55%) whereas Liza macropepis showed high GC content (48.53%) among the mullet species studied. The phylogenetic and genetic distance data showed that Mugil platanus and M. liza represent the continuum of same species. Among the members of family Mugilidae, the genus Mugil might possibly contains more haplotype diversity as revealed by intra-species genetic distance data. Species within genera of Mugilidae family invariably clustered in single clade with high bootstrap value. We conclude that partial COI sequencing (barcoding) in identifying the members of the family and that way has resolved the taxonomic ambiguity among the members of the family Mugilidae.

Key words: Molecular phylogene, mugilidae taxonomy, intra-species variations, DNA barcoding, GC content

INTRODUCTION

The family Mugilidae includes marine fin fishes belonging to 17 genera and more than 60 species of grey mullets (Papasotiropoulos et al., 2007). This family was previously classified in the order Perciformes but is now considered the sole representative of the order Mugiliformes. They are most commonly found in the coastal waters and estuaries of the tropical and subtropical zones of the world. Mullets are generally confined to shallow inshore waters; some of them are stenohaline and others exhibit different degrees of euryhalinity (Papasotiropoulos et al., 2001). The grey mullets are of considerable importance in the capture and culture fisheries in many parts of the world but to ichthyologists mugilids are one of the most difficult taxonomic groups due to its
morphic conserving. Number of researchers has experienced difficulties with animal groups owing to highly overlapping morphological characters which are of taxonomic value and to explore the evolutionary relationship (Blasco-Costa et al., 2008; Li et al., 2011; Liu and Zhao, 2010; Ghajarieh et al., 2006). These characters undergo marked changes during growth. Due to its high degree of conservative morphology, classifying them using classical morphometry and morphology has proved to be complex and difficult (Gilbert, 1993; Thomson, 1997). Since the early 19th century adipose eyelid, thickness of the lips with the presence and absence of papillae or lamellae, scales, teeth on palate and tongue and lateral scale numbers are being used as key characters for distinguishing genera and species in this family (Heras et al., 2007). The results obtained were in most cases controversial, failing to provide any conclusive answers. The general morphological uniformity displayed by the members of Mugilidae family, restricts the number of suitable characters that can be used to answer phylogenetic questions unambiguously. As a result, the phylogenetic status of the family remains particularly obscure, especially at the interspecific level (Rossi et al., 1998). This difficulty is due to very few suitable characters key characters to establish unambiguously the phylogenetic relationships among species (Caldara et al., 1996). The distribution pattern of the species also compounds confusion among conventional taxonomists in classifying the species to the species level. The taxonomy of Mugil liza (Valenciennes, 1836), M. platanus (Günther, 1880) and M. cephalus (Linnaeus, 1758) is confounded in Western Atlantic. M. liza has been reported as a valid species with a distribution ranging from southern Florida to Rio de Janeiro (Thomson, 1997). A substitution of M. liza for M. platanus with more southerly latitudes represents a parapatric distribution involving separate but adjacent habitats. However, M. liza and M. cephalus are particularly difficult to distinguish morphologically (Heras et al., 2007).

The close relationship between M. liza and M. platanus based on their shared haplotypes was proven by the mtDNA data (Fraga et al., 2007). Their morphological similarities including gill rakers besides overlapping values of lateral series scales (LT) counts justifies combined consideration of these taxa (Eiras-Stofella et al., 2001; Cousseau et al., 2005). Despite the fact that the systematics of different Mugilid species has been revised many times, there are still substantial disagreements in foreign literature including Russian. Specifically, in Russian scientific literature Liza aurata is often named Mugil auratus (Papasotriopoulos et al., 2001). The situation with Liza haematocheila is even more complicated. In Russia this species was described as Mugil soiuy while in Korea and Japan waters this species was initially described as Mugil haematocheilus later it was renamed Liza haematocheila (Papasotriopoulos et al., 2001).

As regards the phylogeny of the Mugilidae family, it appears particularly obscure at both the intra- and inter-specific levels; it is extremely difficult to distinguish among species, especially in the juvenile stages because their morphological and physiological characters frequently do not exhibit significant differences (Papasotriopoulos et al., 2002). It was also observed that due to close conservative morphology displayed by all mullets, many investigations based on various morphological characters did not elucidate this problem (Gilbert, 1993) while the use of the pharyngobranchial organ as a key character to address the taxonomy and the phylogeny of grey mullets were also proven futile (Harrison and Howes, 1991). Moreover karyotypic studies on several mullet species did not result in a clear phylogenetic figure (Rossi et al., 2000, 1997). That was despite all these major revisions, the systematic status of some species and genera within the family Mugilidae is still confused (Rossi et al., 1998).

Recently the efficacy of 648 bp of Cytochrome C oxidase subunit I gene in mitochondrial DNA (barcode) has been found to be useful in delineating morphologically cryptic organisms including
fishes to their species level (Ward et al., 2005; Khan et al., 2010; Kamaruzzaman et al., 2011). However, comprehensive study on its utility towards mugilidae members is still scanty. Hence present study was aimed to explore (1) the phylogenetic status of mugilidae fishes, (2) species congruence within Mugilidae family and (3) efficiency of COXI gene in delineating the members of mugilidae fishes to its lowest possible taxon level beyond various geographical boundary.

MATERIALS AND METHODS
Sampling, DNA extraction, amplification and sequencing: Mullets from Parangipettai coastal waters (Mugil cephalus, Liza tade and Liza parsia) were collected alive and transported to the laboratory where the right side of the fishes was photographed and a cube of lateral muscle (5-7 mm) from left side of the fish was exercised for DNA isolation. The fishes were preserved in 95% ethanol for future references. Salting out protocol was adopted for precise and quick DNA isolation from the fish tissue.

The tissue was placed in 1.5 mL eppendorf tube and 500 μL of solution I (50 mM Tris-HCl pH 8, 20mM EDTA pH8 and 2% SDS) was added. The tissue was homogenized with sterile homogenizer and 5 μL of Proteinase K (20 mg mL⁻¹) was added and quick vortexed. The sample was incubated at 55°C in water bath for 2 h with occasional mixing. Following incubation the sample was chilled over ice for 10 min and 250 μL of solution II (6 M NaCl) was added and inverted several times for thorough mixing. The tube was chilled on ice for 5 min and centrifuged at 8000 rpm for 15 min. About 500 μL of supernatant was carefully collected in to new-labeled 1.5 mL tube and twice the volume (i.e., 1 mL) of 100% AR grade ethanol was added to precipitate the DNA. The precipitate was pellet down at 8000rpm for 5 min and the supernatant was removed without touching the pellet. The DNA pellet was rinsed with 500 μL of cold ethanol and centrifuged at 11000 rpm for 5 min. The supernatant was carefully removed and the excess liquid was drained using pipette. The pellet was partially dried (devoid of Ethanol) with lid off at 55°C on heating block. The pellet was re-suspended with 50-200 μL of fresh sterile H₂O depending on size of pellet (100 μL average) by gently pipetting sample with wide-bore filter tip until dissolved. This dissolved DNA acted as a template for Polymerase Chain Reaction (PCR).

The fragment of COI was amplified by GeneAmp PCR system 9700. PCR was carried out in 25 μL volumes [2.5 μL of 10X PCR buffer, 1.5 μL of MgCl₂ (2mM μL⁻¹), 1 μL of DNA template, 1 μL of each primer (10 pmoles μL⁻¹), 2 dNTPs (1 mM μL⁻¹), 10 U of 1 μL of Taq polymerase (Bioserve biotechnologies Pvt, Ltd, Hyderabad, India) and 15 μL of sterile Mill Q water]. Fish F1 [5′-TCAACCAAACCACCGAAGACATTGGGCAC-3′] and Fish R1 [5′-TAGACTTCTGGGTGCCAAAAGATCA-3′] primers were employed for COI amplification (Ward et al., 2005; Khan et al., 2010). The thermocyclic conditions for PCR included the initial denaturation at 94°C for 1 min, five cycles of 94°C for 30 sec, annealing at 45°C for 40 sec and extension at 72°C for 1 min, with a final extension at 72°C for 10 min, followed by indefinite hold at 4°C.

Following PCR, about 10 μL of PCR product with 2 μL of bromo thymol blue were added to 2% agarose gel, prepared with 2.5 μL of 1% Ethidium bromide and electrophorized at 90 V until the dye moved for 6 cm in the gel. The gel was moved to gel doc system for viewing the amplicons with the aid of UV trans-illuminator. Sequencing PCR was carried out using Dye terminator mix v3.1 and quantified in Euro bio-agarose gel. The samples were loaded onto MegaBace sequencer (MB 1000) at Bioserve Biotechnologies, Pvt. Ltd. Hyderabad, India.
Sequence data analysis: The electropherogram generated by automated DNA sequencer was read by Chromas Pro v1.42 and the sequences were carefully checked for mis-calls and base spacing. 95 barcode sequences of Mugilidae were extracted via FASTA format from NCBI. ClustalX 2.0.6 was used to align the nucleotide sequences (Thomson, 1997). The GC content of all 98 barcodes was estimated by BioEdit sequence alignment editor (Hall, 1999). MEGA 4.1 was used to construct phylogenetic trees via Neighbourhood joining method using Kimura 2-parameter and to calculate genetic distance of the given set of sequences (Tamura et al., 2007). Barcode sequence of Lates calcarifer (IOBLMA5) sampled from Parangipettai coastal waters was used as an out group in Phylogenetic tree construction.

RESULTS AND DISCUSSION
Sequence features: Ninety five sequences (collected from NCBI) belonging to various species of mullets (Table 1) disseminated around the world along with the barcode data (651 bp) generated from this work (constitutes 98 sequences belonging to 17 species representing 6 genera of mullets) was considered for phylogram construction and genetic distance analysis. Uniformity in GC content in the barcode region (5' cytochrome oxidase C subunit I) of the family Mugilidae was noted and it ranged between 44.55 and 48.59% while the maximum GC content was found in Liza macrolepis and the minimum was observed in Crenimugil creilabis. The average GC content of mullets was found to be 46.46%.

Phylogenetic analysis: Two phylogenetic trees were constructed to verify the following hypothesis. The first phylogenetic tree (Fig. 1) was constructed to test the efficacy of COI in delineating the members of mugilidae fishes to its species level. Barcode sequence of Lates calcarifer from Parangipettai coastal waters was used as an out group and this has been clearly distinguished as an out group in the phylogenetic tree. Crenimugil sp. shows more genetic relatedness to Mugil sp. as they were placed among the clusters of Mugil sp. (evident from Clade A). Interestingly, Liza parsia and Liza tate were placed among the clade A with members of genus Mugil sp. These

<table>
<thead>
<tr>
<th>Species</th>
<th>No. of sequences used</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aldrichetta forsteri</td>
<td>1</td>
<td>Australia</td>
</tr>
<tr>
<td>Chelon haematocheilus</td>
<td>3</td>
<td>Russia</td>
</tr>
<tr>
<td>Chelon labrosus</td>
<td>4</td>
<td>Spain</td>
</tr>
<tr>
<td>Crenimugil creilabis</td>
<td>2</td>
<td>Canada</td>
</tr>
<tr>
<td>Liza aurata</td>
<td>8</td>
<td>Spain, Russia, Greece</td>
</tr>
<tr>
<td>Liza parsia</td>
<td>2</td>
<td>India</td>
</tr>
<tr>
<td>Liza subsiridis</td>
<td>1</td>
<td>Russia</td>
</tr>
<tr>
<td>Liza macrolepis</td>
<td>7</td>
<td>India, Russia</td>
</tr>
<tr>
<td>Liza saliens</td>
<td>4</td>
<td>Spain, Russia, Greece</td>
</tr>
<tr>
<td>Liza tate</td>
<td>6</td>
<td>India</td>
</tr>
<tr>
<td>Liza ramada</td>
<td>4</td>
<td>Spain, Russia, Greece</td>
</tr>
<tr>
<td>Mugil liza</td>
<td>2</td>
<td>Spain</td>
</tr>
<tr>
<td>Mugil cephalus</td>
<td>17</td>
<td>Canada, Spain, Russia, Mexico, India, Greece</td>
</tr>
<tr>
<td>Mugil curema</td>
<td>15</td>
<td>Spain, USA.</td>
</tr>
<tr>
<td>Mugil platanus</td>
<td>10</td>
<td>Canada, Spain.</td>
</tr>
<tr>
<td>Valamugil seholi</td>
<td>1</td>
<td>Australia</td>
</tr>
<tr>
<td>Valamugil cunningius</td>
<td>10</td>
<td>Canada, India, China, Russia.</td>
</tr>
</tbody>
</table>
Fig. 1: Kimura 2-parameter distance Neighbour Joining (NJ) tree of 98 barcode sequences from 17 species belonging to family Mugilidae. COI sequence of *Lates calcarifer* collected from Parangipettai coastal waters was used as an out group. Specimen number denote the accession number of NCBI database and IOBML* represents the specimens collected from Parangipettai coastal waters and the barcode sequences submitted at Barcode of Life Database (BOLD, www.barcodinglife.org)
misplacements could be resolved by analyzing multiple mitochondrial gene sequencing techniques. Clade B contained the members of Valamugil sp., Chelon sp., Aldrichetta sp. and Liza sp. Within the family Mugilidae, Liza aurata was found to be the distant relative of Mugil cephalus as they were placed in two extremes of the clades of phylogenetic tree. However, each species of family Mugilidae strictly clustered with respective genus which explains the efficacy of COI in delineating the fishes to species level.

Occurrence of Mugil liza in the clades of M. platanus was an interesting observation which triggered the doubt of its close genetic relatedness and hence the second phylogram was constructed (Fig. 2). The phylogenetic tree was constructed using shuffled input sequences of M. liza, M. platanus, M. cephalus (different species within the same genus), Liza tade (species from another genus) and the same out group sequence of Lates calcarifer (a species from different family). The constructed phylogram proved the congeneric relatedness existing among the Mugil liza and Mugil platanus. The sequence of Lates calcarifer was clearly placed as an out group and the resolution of clade to genus and species level was apparent, as Liza tade was isolated from the members of Mugil sp., Mugil cephalus was transparently placed outside the clusters of M. platanus/M. liza complex in the clade.

Congeneric nature of M. platanus and M. liza was recently reported by multiple mitochondrial gene (16s rRNA, COI and Cyt b) sequencing technique (Fraga et al., 2007). The confusion might have originated due to wide geographic distribution of the species. Robins et al. (1986), in his field guide to Atlantic coastal fishes, first reported the occurrence of M. liza in the coastal waters of Boston (USA) while Scorvo Filho et al. (1982) reported the occurrence of same species in Brasilian
waters under the name *M. liza* which might have been the origin of the ambiguity regarding species nomenclature (Agapow et al., 2004). Present study was well corresponded with this finding and constructed phylogram proved the congeneric nature of both the species and thus present conventional morphological taxonomy of these species (*M. platanus* and *M. liza*) should be reinvestigated.

**Genetic distance:** The genetic distance within and between the species and genera of family Mugilidae was calculated using pair-wise distance analysis via Maximum likelihood method (Table 2). The intra species genetic distance within the members of genus *Chelon* sp. was found low (0.05). It may be due to its recent divergence from other members of the family. The intra species variation of genus *Mugil* was found (1.633) higher (Fig. 3). This might be due to possible haplotype diversity existing within the species of *Mugil cephalus*, as reported earlier (Fraga et al., 2007). The analysis showed that among the Mugilidae members, members of genus *Liza* were distantly related to the members of genus *Mugil* as they were grouped at both extreme ends of the phylogram. The overall genetic distance of Mugilidae family was found to be 1.078 which showed the richness of haplotype diversity existing within this family.

**CONCLUSION**

The present analysis showed the close congeneric relationship between *M. liza* and *M. platanus* indicating high degree of gene flow within them and they do not support differentiation at species level. Until new phylogenetic groups are fully identified and implemented, the present species’ status should be preserved to minimize risks of loss of important components of biodiversity. The COI sequence in the phylogram constructed clearly clustered the species of same genus in individual group, proving the efficacy of COI gene in delineating the members of Mugilidae to their
species level. Hence we conclude that COI sequence (DNA barcode) could be potentially used to identify the morphologically cryptic individual members of family Mugilidae.

ACKNOWLEDGMENT
We are thankful to Dr. T. Balasubramanian, Professor and Director of our centre for the encouragement and the authorities of our university for the facilities. We also extend our sincere thanks to Dr. Wafer, National Institute of Oceanography, Goa and Bioserve Biotechnologies (India) Pvt. Ltd. for their constant supporting throughout the research.

REFERENCES