Mitochondrial DNA Diversity of *Tor douronensis* Valenciennes (Cyprinidae) in Malaysian Borneo

1,2YUZINE ESA, 3SITI SHAPOR SIRAJ, 3SITI KHALIJAH DAUD, 1KHAIRUL ADHA A. RAHIM, 1MOHD TAJUDDIN ABDULLAH, 3JEFFRINE ROVIE RYAN JAPNING & 2SOON GUAN, TAN

1Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia 
2Biology Department, Universiti Putra Malaysia 43400 Serdang, Selangor, Malaysia 
3Institute of Biodiversity, Bukit Rengit, 28500 Lanchang, Pahang, Malaysia 

E-mail: kelahzine@yahoo.com

Keywords: *Tor douronensis*, freshwater fish, taxonomy, population structure, COI sequence

ABSTRAK


ABSTRACT

This study examines the population structure and taxonomy of *Tor douronensis*, an important indigenous freshwater fish species in Malaysian Borneo, by using sequence analysis of 466 base pairs of the mitochondrial cytochrome c oxidase I (COI) gene. A total of 62 fish samples were collected from five locations in Sarawak (N=54) and Sabah (N=8). The phylogenetic analysis using the Neighbour-Joining (NJ) method supported the monophyletic status between *T. douronensis* and *Tor tambroides*, which further reinforced their taxonomic status as distinct species. The *T. douronensis* haplotypes were further divided into three major groups, with the Pelian fish from Sabah forming its own group (Cluster III) with strong bootstrap support. The large genetic differences separating the Sabah haplotypes from its Sarawak congeners suggested that the Pelian fish might represent a cryptic species. The current study showed high levels of intra and inter-population variations in *T. douronensis*. Within all population variations, *T. douronensis* populations were found, except in Bario. The presence of fixed haplotype differences along with high *F*<sub>ST</sub> values among the populations of *T. douronensis*, support the conclusion that little or no migration occurred among the extant
populations separated by large geographic distances or river systems. However, the sharing of haplotypes between some such populations, for example between Batang Ai and Bario (HS6), and between Batang Ai and Ulu Limbang/Ba Kelalan (HS2) provided support that *T. douronensis* had a historically widespread natural distribution in the region probably during the Quaternary period. Overall, the present study was able to shed light on the taxonomy and population structure of *T. douronensis* in Malaysian Borneo.

**INTRODUCTION**

Freshwater fishes of the genus *Tor* Gray, commonly known as mahseers, belong to the family Cyprinidae (subfamily Cyprininae) (Mohsin and Ambak, 1983; Roberts, 1989; Kotellat *et al.*, 1993). They are distributed throughout the Indian subcontinent, Southeast Asia and Southern China and inhabit the upper streams and headwaters of most major river systems (Kottelat *et al.*, 1993; Rainboth, 1996). Environmental degradation such as river pollution, deforestation and watershed erosion had led to the rapid destruction of *Tor* natural habitat. Uncontrolled fish harvest (overfishing) has also greatly reduced their population size (Ng, 2004). Their distributions in Malaysian Borneo are now limited to the upper streams and protected areas (natural parks) of Sarawak and Sabah (Litis *et al.*, 1997; Nyanti *et al.*, 1999; Ng, 2004). There are currently three described *Tor* species in Malaysia: *Tor tambroides* Bleeker, *Tor tambra* Valenciennes, and *Tor douronensis* Valenciennes (Kottelat and Whitten, 1996; Roberts, 1989; Rainboth 1996; Ng, 2004).

*T. douronensis* is presumably the most widespread mahseer species recorded in East Malaysia, apart from the less abundant *T. tambroides* found in Sarawak, and it is the only mahseer currently described from Sabah (Inger and Chin, 2002). *T. douronensis*, locally known as “ikan semah” in Sarawak and “ikan pelian” in Sabah has been named “the state fish of Sarawak” due to its importance as a high value food fish as well as for eco-tourism and recreational fishing (Litis *et al.*, 1997; Ng, 2004).

Nevertheless, the taxonomic differentiation of *T. douronensis* and its related species, *T. tambroides* is still unresolved, with many conflicting descriptions among different researchers (Roberts, 1989; Kottelat *et al.*, 1993; Rainboth, 1996; Zhou and Chu, 1996; Ng, 2004). Roberts (1999) classified them to be a single species, and a junior synonym to *T. tambra*.

Thus, the application of molecular techniques (such as DNA sequencing) should offer better insights into the unresolved taxonomy and population size of *T. douronensis* (Nguyen *et al.*, 2006). Molecular markers can give more reliable and consistent results for rapid species identification (Ryan and Esa, 2006), levels of genetic variability, levels of gene flow and population subdivisions and for understanding factors contributing to fitness in freshwater fishes (Vrijenhoek, 1998). Analysis of DNA sequence polymorphism utilizing the existing “universal primers” for mitochondrial DNA (mtDNA) (Palumbi *et al.*, 1991) provides the highest resolution of genetic variation which had been widely applied in molecular systematic studies (Arnason *et al.*, 2002; Liu and Chen, 2003; Nguyen *et al.*, 2006). Thus, this study was conducted to clarify aspects of the systematics and population structure of *T. douronensis* in various populations in Malaysian Borneo by analysing the cytochrome oxidase I (COI) nucleotide sequences of the mitochondrial DNA.

**MATERIALS AND METHODS**

**Sample Description and Collection Location**

Samples of *T. douronensis* were collected from five locations in Sarawak and three locations in Sabah (Fig. 1). However, samples from Sabah (Keningau, Liwagu and Tawau) were pooled together into a single population (Sabah population) due to the small number of individuals obtained for this study. The fish were sampled using a variety of fishing methods, including seine, gill and cast nets and fishing rod. Some samples from Sarawak were
provided by the Indigenous Fish Research and Production Center (IFRPC), Tarat, Sarawak. The weight and standard length of the fish samples used in the study ranged from 50-500 g and 5-100 cm, respectively. Whole samples were morphologically identified by using the keys provided by Inger and Chin (2002), Mohsin and Ambak (1983), and Kottelat et al. (1993). Fresh samples in the form of full specimens, muscle tissues, scales or fin clips were placed in a -80°C freezer for long-term storage. However, in most cases, samples collected in the field were preserved in 95% ethanol and stored at -20°C prior to genetic analyses.

**DNA Extraction and Polymerase Chain Reaction (PCR)**

Total DNA was isolated using the modified CTAB method (Grewe et al., 1993) in the presence of Proteinase K. The pelleted DNA was redissolved in 100μL of sterilized distilled water. The DNA quality and approximate yield were determined by electrophoresis in a 1% agarose gel containing ethidium bromide at 90 V for 30 min. The isolated genomic DNA was used for the mtDNA analysis.

A 500 bp segment of the *cytochrome c oxidase* I gene was amplified with the oligonucleotide primers COIf (5' CCTGCAGGAGGAGGAYCC 3', forward) and COIe (5' CCAGAGATTAGAGGAATCAGTG 3', reverse) (Palumbi et al., 1991). Approximately, 50-100 ng of the template DNA was amplified in a 25 μl reaction mixture containing 50 mM 10X buffer, 2 mM MgCl₂, 0.2 mM of each dNTP (Promega), 0.1 mM of each primer, and 0.5 units of *Taq* DNA Polymerase (Promega). The cycle parameters consisted of 35 cycles of denaturation (95°C, 30 seconds), annealing (45°C, 30 seconds), and extension (72°C, 60 seconds. The amplified products were visualized on a 1% agarose gel containing ethidium bromide for approximately 30 min at 90 V and photographed under UV light. A digested lambda DNA ladder (GeneRuler™ 1 kb DNA Ladder) was used as a size standard marker (Promega). The PCR products were further purified using a DNA purification kit (Invitrogen) according to the manufacturer’s instructions. The purified PCR products were then directly sequenced using the ‘BigDye® Terminator v3.0 Cycle Sequencing kit on an
ABI 377 automated DNA sequencer (PE Applied Biosystem) using only the forward primer (COIf).

Statistical Analysis

The CHROMAS (version 1.45) program was used to display the fluorescence-based DNA sequencing results. The multiple sequence alignment for the forward reactions was done using the CLUSTAL X program (version 1.81; Thompson et al., 1997), and subsequently aligned by eye. The pairwise genetic distance between populations was calculated using the Tamura-Nei distance (Tamura and Nei, 1993), based on unequal base frequencies and unequal ratios of transition to transversion (Ti:Tv) implemented in MEGA (version 3.1; Kumar et al., 2004). The MEGA program was also used to construct a neighbour-joining (NJ) tree (Saitou and Nei, 1987) using two indigenous cyprinids, (Barbonymus gonionotus (Genbank accession number: DQ532806) and Barbonymus schwanenfeldii (Genbank accession number: DQ532805)) obtained from the Jempol River, Negeri Sembilan as outgroup species. Four haplotypes of T. tambroides (Genbank accession number: DQ532827, EF192458, EF192460, EF192461) were also included in the analysis to demonstrate the reciprocally monophyletic status between the two mahseers. The phylogenetic confidence was estimated by bootstrapping with 1000 replicate data sets (Felsenstein, 1985).

The levels of mtDNA COI variation within the T. douronensis population were examined by computing the nucleotide (with the Jukes-Cantor correction; Jukes and Cantor, 1969) and haplotype diversity indices implemented in the DnaSP (version 4.0) program (Rozas et al., 2003). The level of population subdivision (FST) (Hudson et al., 1992) between populations and the Chi-square probability test for population differentiation using 1000 permutations of the data sets were also estimated using the DnaSP program.

RESULTS

Sixty-two partial sequences of 466 base pairs (bp) each of the mtDNA COI gene were obtained, representing the six populations of T. douronensis. We observed 38 (8%) variable/parsimorophic sites including 33 (7%) parsimony-informative sites while 430 sites (92%) were conserved. A total of 14 haplotypes were distinguished in the nucleotide data set with 11 haplotypes being unique and three haplotypes being shared among the six populations (Table 1). In total, 41 substitutions were found among the haplotypes, of which there were 35 transitions and 6 transversions. The sequences of each of the haplotype have been deposited in the GenBank (GeneBank Reference Numbers: EF192444-EF192457).

The mean total nucleotide composition was A=25.7%, T=32.3%, C=22.7% and G=19.3%.

The T. douronensis samples from Sabah harboured four unique haplotypes (HS11P to HS14P) which were not shared with T. douronensis populations from Sarawak. On the other hand, HS6 was the only haplotype found in the Bario population although it was also found in the Batang Ai population. The Layar/Spak population also had four haplotypes, three being unique haplotypes while the fourth one (HS10L) was also found in the Batang Ai population (Table 1).

Overall, the nucleotide diversity was low with the Sabah population showing the highest value (0.016). The haplotype diversity varied, ranging from 0 to 0.900 (Layar/Spak) (Table 1). The pairwise FST (Hudson et al., 1992) and the results of the Chi-square tests for genetic differentiation among the populations are presented in Table 2. Significant levels of genetic differentiation were found in all comparisons among the T. douronensis populations except between the Ulu Limbang and the Ba Kelalan populations (FST= 0.075), and between the Layar/Spak and the Sabah populations, although their pairwise FST value was high (0.726) (Table 2).

Phylogenetic analysis of the haplotypes using the NJ method strongly supported the reciprocally monophyletic status between T.
MITOCHONDRIAL DNA DIVERSITY OF *TOR DOURONENSIS VALENCIENNES* (CYPRINIDAE)

**TABLE 1**
Distribution of 14 observed mtDNA COI haplotypes, nucleotide diversity, number of haplotypes and number of polymorphic sites among populations of *T. douronensis*

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>GenBank Accession Numbers</th>
<th>Population</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HS1BA</td>
<td>EF192444</td>
<td>Ulu Limbang</td>
<td>0.001</td>
<td>0.006</td>
<td>0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS2</td>
<td>EF192445</td>
<td>Ba Kelalan</td>
<td>0.004</td>
<td>0.804</td>
<td>0.865</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS3</td>
<td>EF192446</td>
<td>Bario</td>
<td>0.046</td>
<td>0.047</td>
<td>0.726</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS4</td>
<td>EF192447</td>
<td>Sabah</td>
<td>0.026</td>
<td>0.027</td>
<td>0.042</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS5</td>
<td>EF192448</td>
<td>Layar/Spak</td>
<td>0.005</td>
<td>0.006</td>
<td>0.046</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS6</td>
<td>EF192449</td>
<td>Batang Ai</td>
<td>1.000</td>
<td>0.440</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS7L</td>
<td>EF192450</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS8L</td>
<td>EF192451</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS9L</td>
<td>EF192452</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS10L</td>
<td>EF192453</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS11P</td>
<td>EF192454</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS12P</td>
<td>EF192455</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS13P</td>
<td>EF192456</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS14P</td>
<td>EF192457</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers under each population indicate the frequencies of individuals with that haplotype in each population.

**TABLE 2**
Lower diagonal: pairwise Tamura-Nei genetic distances among the six populations of *T. douronensis*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.075&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.947&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.811&lt;sup&gt;**&lt;/sup&gt;</td>
<td>0.880&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.147&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>2</td>
<td>0.001</td>
<td>0.947&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.804&lt;sup&gt;**&lt;/sup&gt;</td>
<td>0.865&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.175&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>3</td>
<td>0.046</td>
<td>0.047</td>
<td>0.726&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.724&lt;sup&gt;***&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.026</td>
<td>0.027</td>
<td>0.027</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.005</td>
<td>0.006</td>
<td>0.006</td>
<td>0.046</td>
<td></td>
</tr>
</tbody>
</table>

Upper diagonal: population subdivision (F<sub>ST</sub>) values and probability test (Chi-square) for population differentiation based on 1000 permutations of the sequence data set, significance levels (ns=not significant, P<0.05=*, P<0.01=**, P<0.001=***).

PERTANIKA J. TROP. AGRIC. SCI. VOL. 29 NOS. 1 & 2 (2006) 51
douronensis and T. tambroides (Fig. 2) and further divided the former mahseer into three major groups (Cluster I to III) with strong bootstrap supports. Cluster I grouped haplotypes from the Ulu Limbang, Ba Kelalan and Bario populations (Northern Sarawak) with those from the Batang Ai population (Southern Sarawak). Cluster II grouped all the Southern Sarawak haplotypes consisting of three unique haplotypes from the Layar/Spak population and one shared haplotype (HS10L) with the Batang Ai population. Interestingly, Cluster III grouped all the four unique haplotypes from Sabah (North Borneo).

The pairwise genetic distances (number of nucleotide substitutions per site) calculated using the Tamura-Nei model (Tamura and Nei, 1993) among the T. douronensis populations are shown in Table 2. The highest genetic distance was observed between the Sabah population and both the Bario and Ba Kelalan populations of Sarawak (4.7%) while the lowest value was between the Ulu Limbang population and the Ba Kelalan population (0.1%). Within the Sarawak populations, the Layar/Spak population had genetic distances of 2.4% to 2.7% separating it from the other four Sarawak populations. Interestingly, the Batang Ai population had a closer genetic distance (0.5% to 0.6%) with the Northern Sarawak populations (Ulu Limbang, Ba Kelalan and Bario) than with the Southern Layar/Spak population (2.4%).

DISCUSSION
The results of the mtDNA analysis in this study enabled us to shed light on the taxonomic status of T. douronensis in the Malaysian part of Borneo Island. The phylogenetic analysis of the COI gene confirmed the reciprocally monophyletic status between T. douronensis and T. tambroides, thus further reinforcing their taxonomic status as distinct species (Roberts,
MITOCHONDRIAL DNA DIVERSITY OF *Toro douronensis valenciennes* (Cyprinidae)

1989; Kottelat *et al.*, 1993; Rainboth, 1996; Zhou and Chu, 1996; Ng, 2004). The current mtDNA results also did not show any mixing of haplotypes between *T. douronensis* and *T. tambroides* as was observed by Nguyen *et al.* (2006). The major finding of this study is the bifurcation of the *T. douronensis* haplotypes into three highly differentiated groups, with the Sabah (North Borneo) haplotypes forming its own subgroup (Cluster III). The bootstrap support among the three clusters was high although the consensus positioning of Cluster I and Cluster II with regards to Cluster III was moderately supported. A plausible explanation for this is that the Pelian fish from Sabah might represent a cryptic species.

This study found high levels of intra and inter-population variations in *T. douronensis*. Within population variations were found in all the *T. douronensis* populations except in the Bario population. The large mtDNA differences currently found among the *T. douronensis* populations could be explained by one or several factors including small population sizes, past bottleneck events, or the presence of physical barriers to gene flow among the populations (Nguyen *et al.*, 2006). The presence of fixed haplotype differences among the populations, along with high FST values among populations of *T. douronensis*, supported the conclusion that little or no migration occurred among the extant populations separated by large geographic distances, or river systems (Nguyen *et al.*, 2006). Nevertheless, the sharing of haplotypes between such populations does occur, for example between Batang Ai and Bario (HS6), and between Batang Ai and Ulu Limbang/Ba Kelalan (HS2) and this provided support that, in the past, *T. douronensis* had a widespread natural distribution in the region. Geological evidence suggested that the river systems of the Northern and the Southern parts of Sarawak were historically interconnected, most probably during the Tertiary and Quaternary periods (Inger and Chin, 2002).

The large genetic differences between the *T. douronensis* population from Sabah with its congeners from Sarawak, and the presence of fixed haplotypes supported the hypothesis that the North Borneo region was the most isolated region and probably had no connection with the other Borneo regions during the Pleistocene glaciation periods (Inger and Chin, 2002). Thus, the Pelian fish from Sabah could possibly have evolved through allopatric speciation and formed new or cryptic mtDNA lineages. The lack of a clear geographical structuring of haplotype distributions between the Semah fish from the Northern and the Southern parts of Sarawak is also demonstrated in other indigenous freshwater fish species with a widespread natural distribution such as in *Hampala macrolepidota* (Ryan and Esa, 2006).

This study demonstrated the usefulness of genetic studies in assessing the taxonomy and population structures of Malaysia’s indigenous freshwater fish taxa for appropriate conservation and management strategies. However, further studies are required using larger sample sizes per population, samples from other areas of their geographical distributions, sequence data from other mtDNA regions and information based on nuclear DNA (i.e. single locus microsatellite) markers.

**ACKNOWLEDGEMENTS**

We thank Dr Stephen Sungan and officers from the Indigenous Fish Production Research Centre (IFPRC), Department of Agriculture, Sarawak, for their assistance in sample collection. We also thank the Fisheries Department of Malaysia for their full support throughout the project, members of the Molecular Ecology Laboratory, Universiti Malaysia Sarawak, members of the Genetics Laboratory, Universiti Putra Malaysia, and all the people involved in the sample collection and laboratory work. This project was funded by Universiti Malaysia Sarawak through research grant 01(80)411/2003(148) and National Biotechnology Directorate (MOSTI) grant no: 01-02-04-008/BTK/ER/33.

PERTANIKA J. TROP. AGRIC. SCI. VOL. 29 NOS. 1 & 2 (2006) 53
REFERENCES


MITOCHONDRIAL DNA DIVERSITY OF *TOR DOUROENESIS VALENCIENNES* (CYPRINIDAE)


