The First Mitochondrial Control Region Dataset: Critically Endangered Freshwater Turtles in Malaysia, *Orlitia borneensis* and *Batagur borneoensis*

Mohd Hairul Mohd Salleh\textsuperscript{a,b}, Yuzine Esa\textsuperscript{a,c,*}, Nur Amalina Mohd Izam\textsuperscript{d}, Amir Asyraf Zainudin\textsuperscript{a}, Khairul Azhar Romli\textsuperscript{a}, Wan Nur Fatin Syafiqah Wan Nawang\textsuperscript{a}, Athirah Mohd Bakri\textsuperscript{a}, Nor Jasmin Hussein\textsuperscript{e}, Siti Amalia Aisyah Abdul Halim\textsuperscript{a}

\textsuperscript{a}Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia; \textsuperscript{b}Royal Malaysian Customs Department, Persiaran Perdana, Presint 2, 62596 Putrajaya, Malaysia; \textsuperscript{c}International Institute of Aquaculture and Aquatic Sciences, Universiti Putra Malaysia, Lot 960 Jalan Kemang 6, 71050 Port Dickson, Negeri Sembilan, Malaysia; \textsuperscript{d}Universiti Teknologi Mara Pahang Branch, Jengka Campus, 26400 Bandar Tun Abdul Razak Jengka, Pahang; \textsuperscript{e}Faculty of Applied Sciences, Universiti Teknologi MARA (UiTM) 40540, Shah Alam, Selangor, Malaysia

Abstract The Malaysian Giant Turtles and Painted Terrapin mitochondrial control region has the first data deposited to the GenBank database. This dataset describes phylogenetic tree relationships between the genera *Orlitia* and *Batagur*. *Orlitia borneensis* and *Batagur borneoensis* are freshwater turtles listed as critically endangered on the International Union for Conservation of Nature (IUCN) Red List and among 24 species of turtles in Malaysia. Since both species are critically endangered, the data provided here can be used for future conservation genetics studies in order to protect the species from being driven to extinction. A sample of each species was collected aseptically from adult *O. borneensis* (male) and *B. borneoensis* (female) in captivity at Bukit Paloh, Kuala Berang, Terengganu. The data for this study can be found in the GenBank database with accession numbers OQ571740 and OQ571740.

Keywords: Malaysian Giant Turtle, Painted Terrapin, phylogenetic, genetics, GenBank.

Introduction

The control region (CR), also called the displacement loop (D-loop), is a non-coding region of mitochondrial DNA (mtDNA) that plays a vital role in regulating the replication and transcription of the mitochondrial genome [1]. The CR is usually found in the mtDNA molecule between the tRNA-Pro and tRNA-Phe genes, and the sequence of the CR is very different between different individuals and species [2]. This variability makes the CR a valuable region for studying population genetics, phylogenetics, and evolutionary biology [3].

With the Asian Turtle Crisis [4] still going on, the number of many turtle species is decreasing, especially in the family Geoemydidae [5]. This is because of their habitat destruction and targeted exploitation [6].

*Orlitia borneensis* [7] (Figure 1a) and *Batagur borneoensis* [8] (Figure 1b), two large river turtles that live in the Malaysian estuary [9], were the two species of freshwater turtles that were studied. The former is locally known as 'Juku-juku Besar' and the latter as 'Tuntung Laut', based on their nesting preferences [9]. Both species are listed as critically endangered species on the IUCN Red List and listed among 24...
species of turtles in Malaysia [6]. Therefore, this study aims to provide an mtDNA CR dataset for *Oritia borneensis* and *Batagur borneoensis* for future conservation genetics studies geared towards sustainability.

![Photo of Oritia borneensis](image1)

![Photo of Batagur borneoensis](image2)

**Figure 1.** (a) The male Malaysian Giant Turtle, *Oritia borneensis*; (b) The female Painted Terrapin, *Batagur borneoensis*. Both were collected from captive breeding at Bukit Paloh, Kuala Berang, Terengganu
Materials and Methods

The study site is in Bukit Paloh, Kuala Berang (KB), Terengganu (5.0939° N, 102.7821° E) (Figure 2). Following the protocols in [10], blood was drawn using two different venipuncture techniques: the subcarapacial venous plexus (SVP) and the jugular vein. Around 1.5 mL of blood was collected into a 2 mL microcentrifuge tube and mixed with 0.5 mL of EDTA (1:3 ratio) for preservation before being stored at -20 °C. The Department of Wildlife and National Parks, Peninsular Malaysia (DWNP), approved B-00335-16-20 for a research and field permit.

Figure 2. The study site is at Bukit Paloh, Kuala Berang, Terengganu, Malaysia
DNA Extraction and PCR Amplification

Nucleic acids were isolated from 200 µL of each EDTA whole blood sample. Following cell lysis and protein denaturation, extractions were carried out using an automated system, the ReliaPrep™ Blood gDNA Miniprep System (Promega, Madison, WI, USA), with Binding Column technology. The input volume of the EDTA whole blood sample yielded a final extraction volume of 200 µL. The amounts of extracted DNA were determined using a Thermo Scientific™ NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The automated system’s competence to extract nucleic acids of great purity was verified by direct gel loading. Following NanoDrop quantification of the isolated nucleic acids, the results were loaded directly into the 1% (w/v) agarose gel with molecular markers. A primer set of 5′-TTTTCCCTAGCATACCA-3′ (forward) and 5′-AGTTGCTCCTCGGATTAGG-3′ (reverse) designed previously by [11] was used. PCR amplification for gene CR fragments was performed in a Go Taq Flexi PCR (Promega, Madison, USA) reaction mixture containing 2 µL DNA template, 0.5 µL primer, 5 µL 5x PCR buffer, 2 µL x 25 mM MgCl2, 0.5 µL dNTP, 0.2 µL Taq DNA polymerase, and 14.3 µL double-distilled water (ddH2O). Denaturation at 94 °C for 3 min was followed by 25 cycles of denaturation at 94 °C for 35 s. Meanwhile, the primer annealing stage was completed at 60 °C for 1 min and 30 s, followed by a 2 min extension at 72 °C and a final 2 min extension at 72 °C. The purified PCR products were shipped to a private company (First BASE Laboratories Sdn Bhd, Seri Kembangan, Selangor, Malaysia) for CR sequencing with only the forward primer.

Data Analysis

An *Orlitia borneensis* (male) and a *Batagur borneoensis* (female) mtDNA CR sequences were aligned using ClustalW in MEGA11 (Global SaaS Software Company, Paris, France) [12], and the Neighbour Joining (NJ) phylogenetic tree was made using NJ [13], with a confidence level estimated from 1000 bootstrap replicates.

Registration and Availability of Data

The first data of mitochondrial CR sequences from male *Orlitia borneensis* (sample ID: KBJB) and female *Batagur borneoensis* (sample ID: KBTL) were submitted to GenBank and assigned the accession numbers OQ571740 and OQ571741 (Table 1). Both aligned CR lengths were 658 bp. Available online at GenBank, https://www.ncbi.nlm.nih.gov/

<table>
<thead>
<tr>
<th>Scientific Name</th>
<th>English Name</th>
<th>Accession No.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Orlitia borneensis</em></td>
<td>Malaysian Giant Turtle</td>
<td>OQ571740</td>
<td>This Study</td>
</tr>
<tr>
<td><em>Batagur borneoensis</em></td>
<td>Painted Terrapin</td>
<td>OQ571741</td>
<td>This Study</td>
</tr>
<tr>
<td><em>Batagur affinis</em></td>
<td>Southern River Terrapin</td>
<td>MZ555651</td>
<td>[14]</td>
</tr>
<tr>
<td><em>Batagur kachuga</em></td>
<td>Red-crowned Roofed Turtle</td>
<td>MZ156025</td>
<td>[15]</td>
</tr>
</tbody>
</table>

Value of the Data

This is the first data release on the mitochondrial DNA CR of the critically endangered species *Orlitia borneensis* and *Batagur borneoensis*. Establishing a dataset for conservation genetics can be viewed as the first step towards sustainability. The data supplied here could be used to conduct future studies on both species. Additionally, the data presented here can be used to understand and compare other turtles, which are almost all threatened with extinction globally. Moreover, the data can provide valuable insights and information that can drive informed decision-making. By analysing data, one can identify patterns, trends, and correlations, which can be used to optimise processes and make strategic decisions.

Data Interpretation

Based on 658 bp mtDNA CR sequences to analyse and visualise the relationships between mtDNA CR sequences. Surprisingly, only two *Batagur sp.* mtDNA CRs (*B. affinis* and *B. kachuga*) have been reported in GenBank (https://www.ncbi.nlm.nih.gov/genbank) to date, despite the vast number of species in this genus. Therefore, with a significant bootstrap confidence level (>50%), Figure 3 depicts the construction of an NJ phylogenetic tree [13] using the dataset and Mega-11 [12]. According to the cluster in the phylogenetic tree, *B. affinis* and *B. kachuga* have the closest evolutionary ties. Following the subject matter, *B. borneoensis*, all three are in the same genus. Meanwhile, *Orlitia borneensis* is the only
species in the genus [16]. Thus, it will help delineate the future evolutionary relationship between both turtles.

![Tree Diagram]

**Figure 3.** The optimal tree is shown between *Orlitia borneensis* and *Batagur borneoensis* and the other species from Genbank. The Diamondback terrapin, *Malaclemys terrapin* [17], is an outgroup in the phylogeny. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is displayed next to the branches [18]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method [19] and are in units of the number of base substitutions per site. This analysis involved four nucleotide sequences. Codon positions included were 1st, 2nd, 3rd, and non-coding. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There was a total of 658 positions in the final dataset. Evolutionary analyses were conducted in MEGA11 [12]

**Conclusion**

In conclusion, our population genetics dataset provides valuable insights into the genetic diversity and structure of the studied population. Through the analysis of genetic markers, we have uncovered significant patterns and relationships that shed light on the evolutionary history and demographic processes at play. Further research with larger sample sizes and genomic-scale approaches would be beneficial to validate and expand upon these findings.

**Conflicts of Interest**

The author(s) declare that there is no conflict of interest regarding the publication of this paper.
Acknowledgement

We want to thank the Breeding Genetics Laboratory, Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, for helping us with this research project and giving us space to work. We also want to thank the Department of Wildlife and National Parks, Peninsular Malaysia, for helping us with this study. Lastly, we’d like to thank our anonymous reviewers for the ideas they gave us to make this paper better. This study was funded by the Geran Dalaman UiTM (DDN), Universiti Teknologi Mara Pahang Branch, under allocation 600-TNCP1 5/3/DDN (06) (015/2021).

References


