DATASETS ARTICLE

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

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

[a]Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia; [b]Royal Malaysian Customs Department, Persiaran Perdana, Presint 2, 62596 Putrajaya, Malaysia; [c]International Institute of Aquaculture and Aquatic Sciences, Universiti Putra Malaysia, Lot 960 Jalan Kemang 6, 71050 Port Dickson, Negeri Sembilan, Malaysia; [d]Universiti Teknologi Mara Pahang Branch, Jengka Campus, 26400 Bandar Tun Abdul Razak Jengka, Pahang; [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.

Introduction

*For correspondence:
yuzine@upm.edu.my

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 *Orlitia borneensis* and *Batagur borneoensis* for future conservation genetics studies geared towards sustainability.

(a)

(b)

Figure 1. (a) The male Malaysian Giant Turtle, *Orlitia 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′-TTTTTCCCCTAGCATATCACCA-3′ (forward) and 5′-AGTTGCTCTCGGATTTAGGG-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 5× PCR buffer, 2 µL × 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 1. List of freshwater turtle species studied through CR sequences and the GenBank accession number of each species

Scientific Name	English Name	Accession No.	Reference
Orlitia borneensis	Malaysian Giant Turtle	OQ571740	This Study
Batagur borneoensis	Painted Terrapin	OQ571741	This Study
Batagur affinis	Southern River Terrapin	MZ555651	[14]
Batagur kachuga	Red-crowned Roofed Turtle	MZ156025	[15]

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.
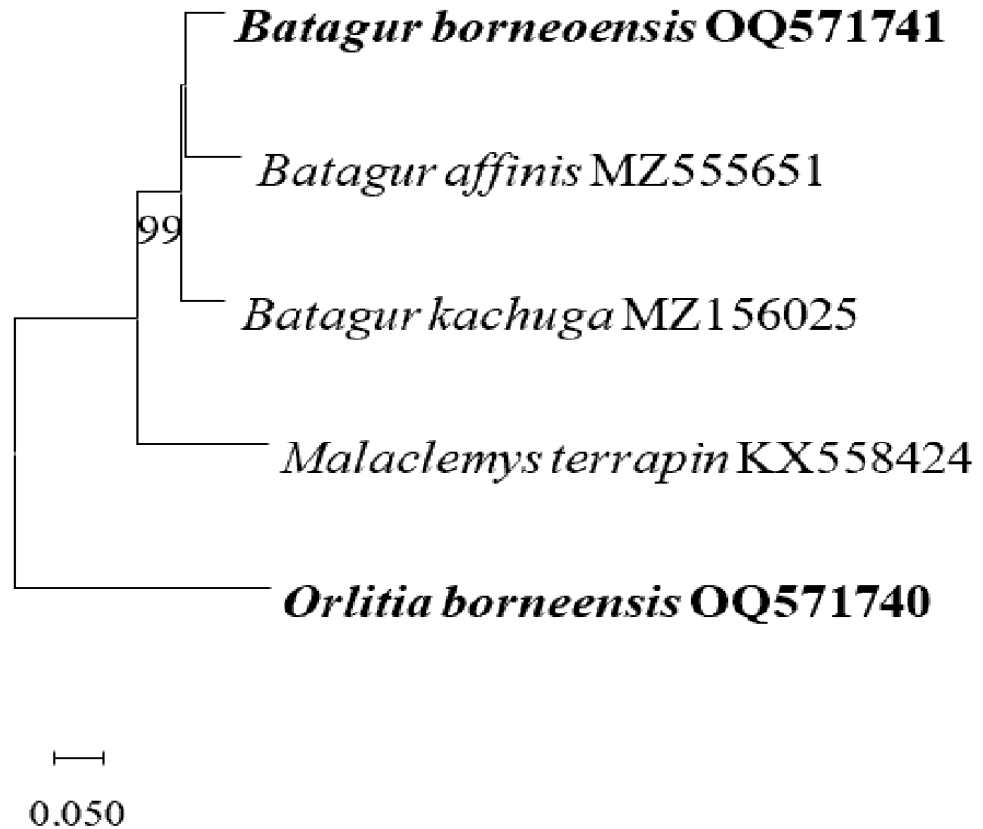

Figure 3. The optimal tree is shown between *Orlitia borneensis* and *Batagur borneoensis* and the other species from Genbank. The Diamondback terrapin, *Malaclemys terrapin* [17], is an outgroup in the phylogeny. The percentage of replicate trees in which the associated taxa clustered together in the 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

References

[1] Avise, J. C. (2004). *Molecular markers, natural history and evolution (2nd ed.).* Sunderland, Mass: Sinauer Associates.

[2] Moritz, C., Dowling, T. E., & Brown, W. M. (1987). Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annual Review of Ecology and Systematics*, *18*(1), 269-292.

[3] Nachman, M. W. (1998). Mitochondrial DNA and the evolution of vertebrates. *Current Opinion in Genetics & Development*, *8*(6), 668-672.

[4] Cheung, S. M., & Dudgeon, D. (2006). Quantifying the Asian turtle crisis: Market surveys in southern China, 2000–2003. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *16*(7), 751-770.

[5] Palupčíková, K., Somerova, B., Protiva, T., Rehak, I., Velenský, P., Hulva, P., Gunalen, D. & Frynta, D. (2012). Genetic and shell-shape analyses of *Orlitia borneensis* (Testudines: Geoemydidae) reveal limited divergence among founders of the European zoo population. *Zootaxa, 3280*(1), 56-66.

[6] Mohd Salleh, M. H., Esa, Y., Salleh, S. M., & Mohd Sah, S. A. (2022). Turtles in Malaysia: A review of conservation status and a call for research. *Animals*, *12*(17), 2184.

[7] Gray, J. E. (1873). On a new freshwater tortoise from Borneo (*Orlitia borneensis*). *Journal of Natural History*, *11*(62), 156-157.

[8] Schlegel, H. & S. Müller. (1845). *Over de schildpadden van den Indische Archipel, en beschrijving eener nieuwe soort van Sumatra. In: Temminck, C.J.* (ed.), Verhandelingen over de Natuurlijke Geschiedenis der Nederlandsche Overzeesche bezittingen. 1839-1844. Zoologie. *2*, 29-36. Leiden

[9] Duli, N. B. (2009). Morphometric and genetic variability of river terrapin (*Batagur baska*) and painted Terrapin (*Batagur borneoensis*). Unpublished M. Sc. Thesis. University of Science Malaysia, George Town, Malaysia, 45.

[10] Salleh, M. H. M., & Esa, Y. (2022). Minimally invasive blood collection techniques as a source of gDNA for genetic studies on turtles and tortoises. *International Journal of Aquatic Biology*, *10*(2), 145-150.

[11] Çilingir, F. G., Seah, A., Horne, B. D., Som, S., Bickford, D. P., & Rheindt, F. E. (2019). Last exit before the brink: Conservation genomics of the Cambodian population of the critically endangered Southern River Terrapin. *Ecology and Evolution*, *9*(17), 9500-9510.

[12] Tamura K., Stecher G., & Kumar S. (2021). MEGA 11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution.* https://doi.org/10.1093/molbev/msab120.

[13] Saitou, N. & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution 4*, 406-425.

[14] Salleh, M. H. M., Esa, Y. & Pau, N. S-S. (2023). Conservation genetics of the critically endangered Southern River Terrapin (*Batagur affinis*) in Malaysia: Genetic diversity and novel subspecies distribution Ranges. *Biology, 12*.

[15] Krishnan, R., Sunil, M., Majhi, A., Tejwan, Y., Srivastava, A. & Das, S. P. (2021). *GenBank: MZ156025, Batagur kachuga isolate BK1 D-loop, complete sequence; mitochondrial.* National Center for Biotechnology Information (NCBI). https://www.ncbi.nlm.nih.gov/nuccore/MZ156025.

[16] Murray, R. A. (2004). Endohelminths from six rare species of turtles (Bataguridae) from Southeast Asia confiscated by international authorities in Hong Kong, China (Doctoral dissertation, Texas A&M University).

[17] Spinks, P. Q., Thomson, R. C., McCartney-Melstad, E., & Shaffer, H. B. (2016). Phylogeny and temporal diversification of the New World pond turtles (Emydidae). *Molecular Phylogenetics and Evolution*, *103*, 85-97.

[18] Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, *39*, 783-791.

[19] Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution, 16*, 111-120.