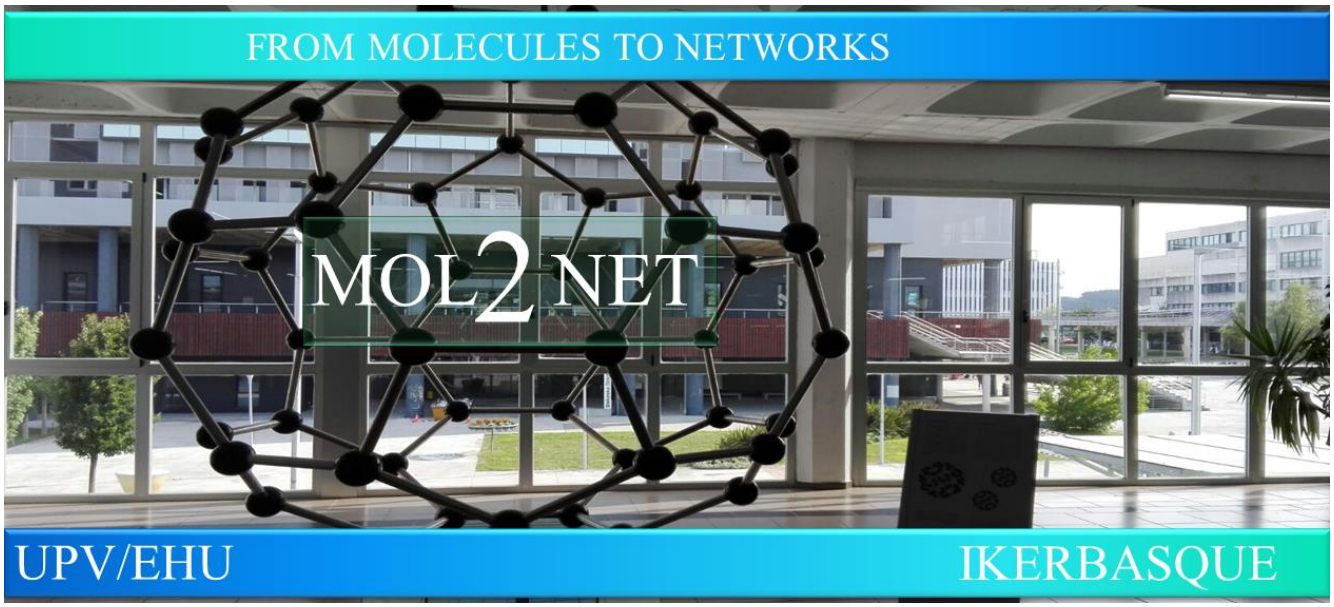
MOL2NET'22, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 8th ed.

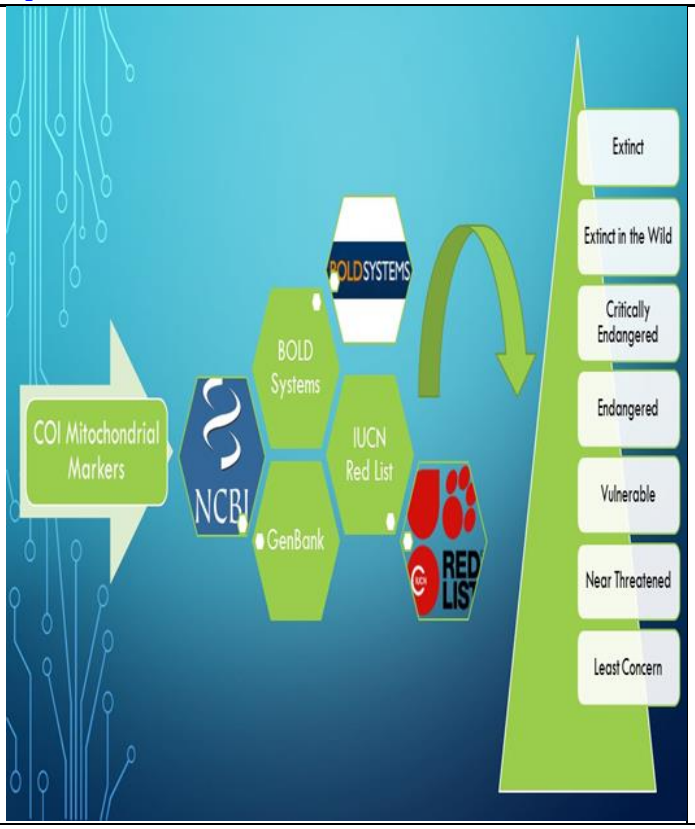


Conservation Status of Globally Testudines Terrapins Based on COI Mitochondrial Markers

Mohd Hairul Mohd Salleh [a,b], Yuzine Esa [a,c]

[a] Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400, UPM, 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.

Graphical Abstract	**Abstract.**
	Terrapins inhabit brackish water or coastal salt marshes. Terrapins are adapted to intermediate salinities but frequently face saltwater-inundated marsh habitats. To date, 12 species of terrapin have been reported worldwide. The present study aims to determine the global utility of terrapin



DNA barcoding using novel COI sequences and compare them to other COI sequences previously published in BOLD systems and GenBank. A total of 26 COI sequences of worldwide terrapins were assessed in this study, including four COI sequences generated from field sampling on the east and west coasts of Peninsular Malaysia. The assessment of the COI sequences with the UPGMA tree identified three families, with 33% of terrapins being classified as least concern (LC) and 25% of them being classified as critically endangered (CR). In this study, we looked at the genome and bioinformatics of terrapins, which could be used as a starting point for future research on terrapin species all over the world.

Key words: *DNA barcoding, COI sequences, BOLD Systems, GenBank, Peninsular Malaysia, UPGMA tree, genome.*

1. Introduction

DNA barcoding is a complementary technology in classical taxonomy and systematics study that has enabled a more precise understanding of the current flora and fauna worldwide [1]. This technology utilises the use of a specific part of the DNA, often from the mitochondrion genome that has a simple genetic structure and relatively rapid rate of evolution, to distinguish closely related taxa. Research attempts have been made at the global and local levels to identify freshwater turtles using DNA barcoding [2]. The DNA barcoding procedure is summarised in Figure 1.

In this study, we collected four samples of *Batagur affinis ssp.* from Peninsular Malaysia. The genetic profiles of these samples were evaluated with the cytochrome oxidase subunit I (COI) marker gene and compared to the DNA sequences of other 11 species of terrapins retrieved from the public database portal. Furthermore, *B. affinis* sequences from this study were submitted to the GenBank database portal for the first time, adding valuable genetic resources for global conservation research. All in all, this study aimed to define the global utility of terrapin DNA barcoding via novel COI sequences and compare them to other COI sequences available in BOLD systems and GenBank.

Figure 1. Summary of DNA barcoding procedure.

2. Materials and Methods

2.1 Study sites

In this study, four *Batagur affinis ssp.* individuals from two different population regions on the east and west coasts of the Malaysia Peninsula were selected, and sampling was conducted in 2020 (Figure 2). The blood samples of *B. affinis affinis* (N = 1) were collected from a captive hatchling population at the Bota Kanan head-starting facility (BK; GPS coordinate: 4.3489° N and 100.8802° E) located at Perak, Malaysia. The blood samples of wild hatchlings *B. affinis edwardmolli* (N = 3) were collected from a population in the Bukit Paloh area of Kuala Berang (KB; GPS coordinate: 5.0939° N, 102.7821° E) at Terengganu, Malaysia. Venipuncture techniques were used to draw blood from the species through the subcarapacial venous plexus (SVP) and internal jugular vein. Before storing at -20 °C, a total of 1.5 mL of blood was preserved with 0.5 mL of EDTA in a 2 mL microcentrifuge tube at a 1:3 ratio. The research and field permit approval number is B-00335-16-20, issued by the Department of Wildlife and Parks, Peninsular Malaysia.

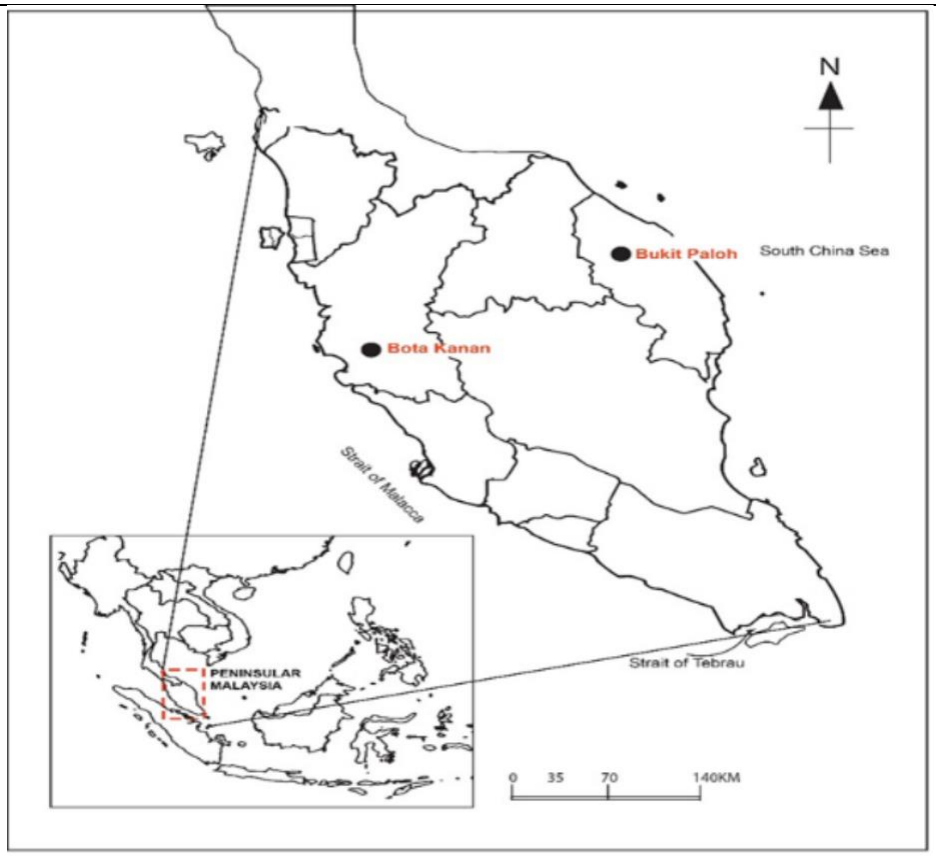


Figure 2. Sampling sites of *Batagur affinis* in Peninsular Malaysia.

2.2 DNA isolation, PCR, and sequencing

Each 200 µL of EDTA whole blood sample was used for nucleic acids extraction. Following cell lysis and protein denaturation, DNA extraction was performed using ReliaPrep^{TM} Blood gDNA Miniprep System with Binding Column technology (Promega, Madison, USA) by following the manufacturer's instructions. According to the input volume of the EDTA whole blood sample, the final extracted volume was adjusted to 200 µL. The quantity and purity of extracted DNA samples were assessed using Thermo Scientific^{TM} NanoDrop 2000c spectrophotometer model ND-2000 (Thermo Fisher Scientific, Waltham, USA). Following the quantification of the isolated nucleic acids, the DNA samples were loaded into 1% (w/v) agarose gel with molecular markers and electrophoresis was performed to evaluate the integrity and intactness of the high molecular weight DNA band.

For PCR, Tuntong set primers 5'-CGCGGAATTAAGCCAACCAG-3' (forward sequence) and 5'-TTGGTACAG-GATTGGGTCGC-3' (reverse sequence) targeting the COI marker gene was used [3]. The PCR amplification of COI gene fragment was performed in a Go Taq Flexi PCR (Promega, Madison, USA) reaction mixture containing 2 µL of DNA template, 0.4 µL of primers, 4 µL of 5x PCR buffer, 1.6 µL of 25 mM MgCl2, 0.4 µL of dNTPs, 0.2 µL of Taq DNA polymerase, and 11 µL of distilled water (ddH_2O). After an initial denaturation at 94°C for 4 minutes, 35 cycles consisting of denaturation at 94°C for 45 seconds, annealing at 55°C for 35 seconds, and extension at 72°C for 1 minute were performed, and lastly a final extension step for 10 minutes at 72°C. Finally, the PCR products were purified and sent to a local company (First BASE Sdn Bhd) to sequence the COI gene of mitochondrial DNA (mtDNA-COI) using Sanger sequencing technology. Additionally, 17 COI sequences of *B. affinis* were data mined and downloaded from GenBank, while five COI sequences of *B. affinis* were mined from BOLD Systems, yielding a total of 26 sequences for this work (Tables 1).

Table 1. List of terrapins species studied through DNA barcoding with the BOLD IDs of their respective COI sequences and the GenBank accession of each species.

Scientific Name	English Name	GenBank Accession	BOLD ID	IUCN Red List	Locallity	References
Batagur baska	Northern river terrapin	KF894752	GBGCR2852-19	CR	India	[4]
Batagur baska	Northern river terrapin	HQ329671	GBGCR2716-19	CR	India	[5]
Batagur borneoensis	Painted terrapin	HQ329672	GBGCR2717-19	CR	Indonesia	[5]
Batagur borneoensis	Painted terrapin	None	BENT109-08	CR	Indonesia	[6]
Morenia ocellata	Bengal eyed terrapin	HQ329690	GBGCR2724-19	EN	Myanmar	[5]
Morenia ocellata	Bengal eyed terrapin	None	BENT264-09	EN	Myanmar	[6]
Malaclemys terrapin	Diamondback terrapin	HQ329654	GBGC11262-13	VU	America	[5]
Malaclemys terrapin	Diamondback terrapin	KX559038	GBGCR2938-19	VU	America	[7]
Emys orbicularis	European pond terrapin	HQ329643	GBGC11273-13	NT	Unknown	[5]
Emys orbicularis	European pond terrapin	KP697925	None	NT	Germany	[8]
Melanochelys trijuga	Indian pond terrapin	KC354725	GBGC11418-13	LC	India	[4]
Melanochelys trijuga	Indian pond terrapin	KC354724	GBGC11419-13	LC	India	[4]
Rhinoclemmys rubida	Mexican spotted terrapin	HQ329701	GBGCR2766-19	NT	Mexico	[5]
Trachemys scripta elegans	Red-eared terrapin	KX559044	GBGCR1038-18	LC	America	[7]
Trachemys scripta elegans	Red-eared terrapin	KM216748	GBGCR1008-15	LC	America	[9]
Pelusios sinuatus	Serrated hinged terrapin	None	BENT174-08	LC	Southern Africa	[6]
Pelusios sinuatus	Serrated hinged terrapin	HQ329735	GBGC11221-13	LC	Southern Africa	[5]
Siebenrockiella crassicollis	Smiling terrapin	HQ329704	GBGCR2769-19	EN	Unknown	[5]
Siebenrockiella crassicollis	Smiling terrapin	None	BENT190-08	EN	Unknown	[6]
Mauremys caspica	Striped-neck terrapin	AY337348	GBGC0806-06	LC	Iran	[10]
Mauremys caspica	Striped-neck terrapin	AY337347	GBGC0805-06	LC	Bahrain	[10]
Batagur affinis	Southern river terrapin	None	MTD042-21	CR	Malaysia	[6]
Batagur affinis affinis	Southern river terrapin	OL658844	HYT001-21	CR	Malaysia	This study
Batagur affinis edwardmolli	Southern river terrapin	OL658845	HYT002-21	CR	Malaysia	This study
Batagur affinis edwardmolli	Southern river terrapin	OL658846	HYT003-21	CR	Malaysia	This study
Batagur affinis edwardmolli	Southern river terrapin	OL658847	HYT004-21	CR	Malaysia	This study

2.3 DNA barcode sequence quality control measures and analysis

For each sample, chromatograms that show the nucleotide sequences of both DNA strands were generated. Some low-quality noisy sequences on both ends and chromatograms with more than 2% ambiguous bases were trimmed. The bidirectional reads were removed by benchmarking against a quality value of more than 40. A computer programme that is known as SeqScape, version 2.7 (Applied Biosystems) was used to view and combine the forward and reverse chromatograms to get the consensus sequences. The sequences were checked against the GenBank and the BOLD Systems database for accession numbers and BOLD sequence identifiers (Table 1). MEGAX [11] was used to align all the sequences and generate multiple sequence alignments of all sequences with the same length and starting point. Phylogenetic analysis was performed using MEGAX with 1000 bootstrap replicates and the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) [12]. Microsoft Excel 2016 was used to make the pie charts that show all terrapins' different family groups and conservation statuses.

3. Results and Discussion

Twenty-six COI sequences from *B. affinis* were analysed in this study, in which all of them were from the order Testudines. The species are classified into three families: Geoemydidae (18%), Emydidae (6%) and Pelomedusidae (2%) (Figure 3). All species studied in this study are worldwide terrapins that live in freshwater and brackish water. Notably, [13] identified 13 species of terrapins worldwide but disregarded a Seychelles black terrapin species (*Pelusios seychellensis*) that was believed to be extinct. However, genetic analysis of the lectotype revealed that this terrapin has not been extinct and has been referred to as *Pelusios castaneus*. Specimens were possibly mislabeled or mixed up in a private collection before being acquired by the Zoological Museum Hamburg in 1901 [14].

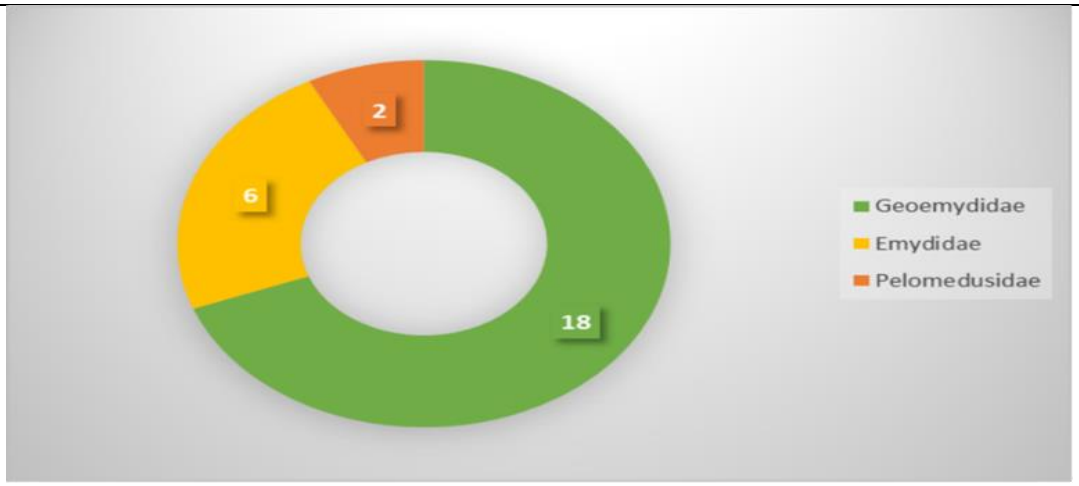


Figure 3. The family of sampled terrapins.

With 100 bootstrap supports, the UPGMA phylogeny of the studied dataset demonstrated cohesive monophyletic clustering of all studied species (Figure 4). Cohesive clustering was also observed between the database reference sequences for the representative species and the produced sequences in the phylogenetic tree. The species were grouped according to their family, with Geoemydidae the most numerous. According to the evolutionary tree, *B. baska* was originated in India and was closely linked to *B. affinis* from Malaysia. Additionally, *Melanochelys trijuga* is closely related to *Mauremys caspica* from the Persian Gulf, whilst the *Malaclemys terrapin* is strongly related to *Trachemys scripta elegans* is native to North America.

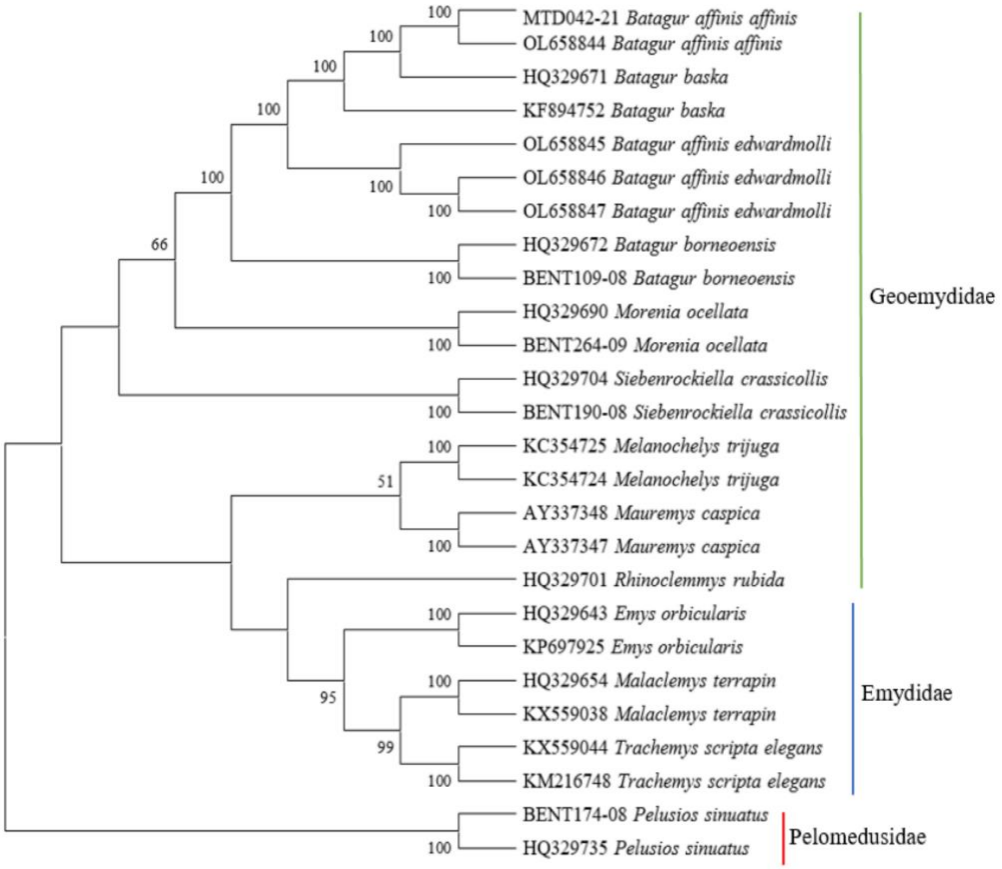


Figure 4. UPGMA tree constructed with MEGAX based on COI sequences belonging to order Testudines.

3.1 Conservation Status

The International Union for the Conservation of Nature (IUCN) Red List is a critical health indicator for biodiversity. The IUCN is far more than a list of species and their states; it is a crucial instrument for informing and catalysing conservation and policy change activity. These steps are critical to preserving the natural resources humans rely on for survival [15, 16]. The IUCN Red List Categories and Criteria are intended to provide a simple framework for identifying species on the verge of extinction on a global scale. It assigns species to one of the following categories: Not Evaluated (NE), Data Deficient (DD), Least Concern (LC), Near Threatened (NT), Vulnerable (VU), Endangered (EN), Critically Endangered (CR), Extinct in the Wild (EW), or Extinct (EX) [15]. Accordingly, 33% of the 12 species of terrapins in this study were classified under the LC category, followed by 25% of the samples belonging to the CR category, 8 % of the samples belonging to the VU category and the remaining samples were classified under EN and NT categories at 17% respectively (Figure 5).

Almost every country with indigenous species has a captive breeding programme. Based on Table 1, three *Batagur sp.* of terrapins are in CR status, including *B. affinis*, *B. baska*, and *B. borneoensis*. However, in Southeast Asia countries such as Indonesia, Singapore, Thailand, and Vietnam, *B. affinis* is assumed to be classified under EW [17, 18]. Furthermore, *B. baska* could possibly be regionally extinct in Myanmar and Thailand [19]. Moreover, *B. borneoensis* was distributed in Indonesia, Malaysia, and Brunei but it was virtually extinct in Thailand [20].

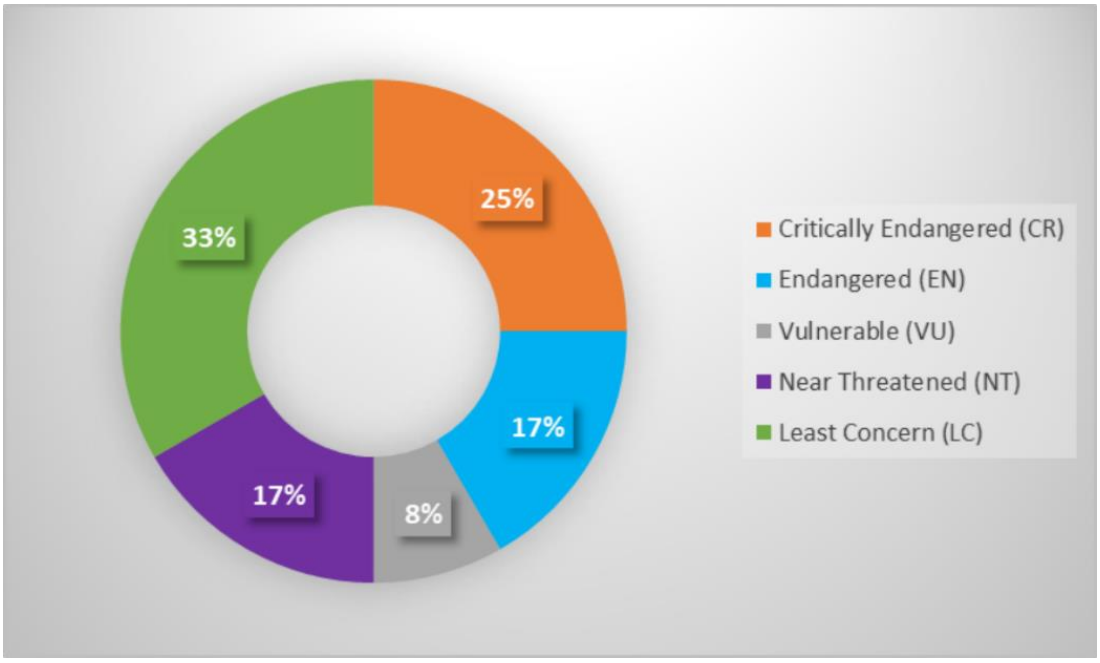


Figure 5. The conservation status of the terrapins is based on the IUCN Red List.

4. Conclusions

In conclusion, the COI marker is an effective DNA barcode marker for terrapin species, providing vital evidence that it may be utilised to distinguish and recognise genera and species of these Testudines organisms. Nonetheless, in the future study, other molecular markers and additional samples from new sampling sites should be included to assess terrapin populations extensively. The genomic and bioinformatics analyses of terrapins taxonomy reported here may serve as a foundation for future research on this species throughout the world, allowing for more practical conservation work for this threatened species.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines and approved by the Department of Wildlife and National Parks, Peninsular Malaysia, with the research permit (B-00335-16-20).

Data Availability Statement: The data presented is authentic and was not improperly selected, manipulated, enhanced, or fabricated. The data can be found in GeneBank (accession numbers OL658844-OL658847) and BOLD Systems (Sequence IDs HYT001-21 to HYT004-21).

Acknowledgements: We want to acknowledge the Universiti Putra Malaysia for funding my study through a Graduate Research Fellowship. In addition, we would also like to express our gratitude to the Genetics and Breeding Laboratory, Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia for the facilities and support of this research project. Furthermore, we thank the Turtle Conservation Society of Malaysia and the Department of Wildlife and National Parks, Peninsular Malaysia, for their collaboration. Last but not least, we thank the anonymous reviewers for their insights and suggestions to improve this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Philos Trans R Soc Lond B Biol Sci*. **2003**, 270:313–322. https://doi.org/10.1098/rspb.2002.2218.
2. Salleh, M. H. M., & Esa, Y. The mtDNA D-loop Marker Identifies the Genetic Variability of Indochina's *Batagur affinis*. In *1ST Postgraduate Seminar on Agriculture and Forestry 2021 (PSAF 2021)* (p. 82), 6 October 2021.
3. Guntoro, J., & Riyanto, A. The very low genetic variability on Aceh Tamiang's (Indonesia) population of Painted Terrapin (*Batagur borneoensis*) inferred by cytochrome oxidase I (CO I) and D-loop (control region). *Biodiversitas Journal of Biological Diversity* **2020**, *21*(6). https://doi.org/10.13057/biodiv/d210624.
4. Kundu S, Das KC, Ghosh SK. Taxonomic rank of Indian tortoise: revisit with DNA barcoding perspective. *DNA Barcodes* **2013**; 1:39-45. DOI: 10.2478/dna-2013-0003.
5. Reid BN, LE M, McCord WP, Iverson JB, Georges A, Bergmann T, Amato G, Desalle R, Naro-Maciel E. Comparing and combining distance-based and character-based approaches for barcoding turtles. *Molecular Ecology Resources* **2011**, 11(6), 956-967. doi: 10.1111/j.1755-0998.2011.03032.x.
6. Ratnasingham, S., & Hebert, P. D. BOLD: The Barcode of Life Data System (http://www.barcodinglife. org). *Molecular ecology notes* **2007**, *7*(3), 355-364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.
7. Spinks, P. Q., Thomson, R. C., McCartney-Melstad, E., & Shaffer, H. B. Phylogeny and temporal diversification of the New World pond turtles (Emydidae). *Molecular Phylogenetics and Evolution* **2016**, *103*, 85-97. https://doi.org/10.1016/j.ympev.2016.07.007.

8. Hawlitschek O, Morinière J, Dunz A, Franzen M, Rödder D, Glaw F, Haszprunar G. Comprehensive DNA barcoding of the herpetofauna of Germany. *Mol Ecol Resour*. **2016**, 16(1):242-53. doi: 10.1111/1755-0998.12416.
9. Yu, D., Fang, X., Storey, K. B., Zhang, Y., & Zhang, J. Complete mitochondrial genomes of the yellow-bellied slider turtle *Trachemys scripta scripta* and anoxia tolerant red-eared slider *Trachemys scripta elegans*. *Mitochondrial DNA Part A* **2016**, *27*(3), 2276-2277. https://doi.org/10.3109/19401736.2014.984178.
10. Feldman, C. R., & Parham, J. F. Molecular systematics of Old World stripe-necked turtles (Testudines: Mauremys). *Asiatic Herpetological Research* **2004**, *10*(LBNL-54656).
11. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution* **2018**, *35*(6), 1547. https://doi.org/10.1093/molbev/msy096.
12. Sneath, P.H.A. and Sokal, R.R. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. WF Freeman & Co., San Francisco, 1973; 573 p.
13. Turtle Owner. What Are Terrapins? (General Info, Pictures and Care Guide). https://turtleowner.com/what-are-terrapins-general-info-pictures-and-care-guide/ (accessed on 1 January 2022).
14. Stuckas H, Gemel R, Fritz U. One Extinct Turtle Species Less: *Pelusios seychellensis* Is Not Extinct, It Never Existed. *PLoS ONE* **2013**, 8(4): e57116. https://doi.org/10.1371/journal.pone.0057116.
15. IUCN. *IUCN Red List Categories and Criteria: Versian 3.1*. IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK., 2001; Ii + 30 pp.
16. Vié, J. C., Hilton-Taylor, C., Pollock, C., Ragle, J., Smart, J., Stuart, S. N., & Tong, R. The IUCN Red List: a key conservation tool. *Wildlife in a changing world–An analysis of the 2008 IUCN Red List of Threatened Species*, 2009; 1.
17. Mistar, Siregar, A. J., & Singleton, I. Presence and distribution of the Southern River Terrapin *Batagur affinis* and Painted Terrapin *Batagur borneoensis* in eastern coast of Sumatra. Unpublished Report to Auckland Zoo, Auckland, New Zealand, 2012; 25 pp.
18. Moll, E.O., Platt, S.G., Chan, E.H., Horne, B.D., Platt, K., Praschag, P., Chen, P.N. and Van Dijk, P.P., *Batagur affinis* (Cantor 1847) Southern river terrapin, tuntong. *Chelonian Research Monographs* **2015**, 5, 090.1–090.17. https://doi.org/10.3854/crm.5.090.affinis.v1.2015
19. Praschag, P., & Singh, S. *Batagur baska*. *The IUCN Red List of Threatened Species*. 2019. http://dx.doi.org/10.2305/IUCN.UK.2019-1.RLTS.T97358453A2788691.en.
20. Jenkins M.D. *Tortoises and Freshwater Turtles: The trade in Southeast Asia*. TRAFFIC International, United Kingdom, 1995; 48 pp.