DNA barcoding of fish diversity from Batanghari River, Jambi, Indonesia

Huria Marnis[1,#,*], Khairul Syahputra[1,#], Jadmiko Darmawan[1], Dwi Febrianti[2], Evi Tahapari[1], Sekar Larashati[2], Bambang Iswanto[1], Erma Primanita Hayuningtyas Primanita[1], Mochamad Syaifudin[3], Arsad Tirta Subangkit[1]

[1] Research Center for Fishery, National Research and Innovation Agency (BRIN), Cibinong 16911, Indonesia
[2] Research Center for Limnology and Water Resources, National Research and Innovation Agency (BRIN), Cibinong 16911, Indonesia
[3] Program Study of Aquaculture, Faculty of Agriculture, Sriwijaya University, South Sumatra 30662, Indonesia

Abstract

Global climate change, followed by an increase in anthropogenic activities in aquatic ecosystems, and species invasions, has resulted in a decline in aquatic organism biodiversity. The Batanghari River, Sumatra's longest river, is polluted by mercury-containing illegal gold mining waste (PETI), industrial pollution, and domestic waste. Several studies have provided evidence suggesting a decline in fish biodiversity within the Batanghari River. However, a comprehensive evaluation of the present status of biodiversity in this river is currently lacking. The species under investigation were identified through various molecular-based identification methods, as well as morphological identification, which involved the use of neighbor-joining (NJ) trees. All collected specimens were initially identified using morphological techniques and subsequently confirmed with molecular barcoding analysis. Morphological and DNA barcoding identification categorized all specimens (1,692) into 36 species, 30 genera and 16 families, representing five orders. A total of 36 DNA barcodes were generated from 30 genera using a 650-bp-long fragment of the mitochondrial cytochrome oxidase subunit I (COI) gene. Based on the Kimura two-parameter model (K2P), The minimum and maximum genetic divergences based on K2P distance were 0.003 and 0.331, respectively, and the average genetic divergence within genera, families, and orders was 0.05, 0.12, 0.16 respectively. In addition, the average interspecific distance was approximately 2.17 times higher than the mean intraspecific distance. Our results showed that the COI barcode enabled accurate fish species identification in the Batanghari River. Furthermore, the present work will establish a comprehensive DNA barcode library for freshwater fishes along Batanghari River and be significantly useful in future efforts to monitor, conserve, and manage fisheries in Indonesia.

Keywords: Batanghari River, Biodiversity, DNA barcoding, Fish biodiversity

Introduction

New challenges arise in the face of global climate change that affects every ecosystem on earth, including the aquatic environment, which causes accelerated loss of biodiversity in several aquatic ecosystems (Worm & Lotze, 2021). Climate change is expected to cause major changes in water scarcity in some areas and an increased risk of flooding in most areas. In addition, climate change has increased the temperature of the waters, thus affecting water quality and freshwater ecosystems (Mann et al., 2022). Several studies have shown clear evidence that biodiversity is already responding to climate change (Gillette et al., 2022). Apart from climate change, other anthropogenic impacts also affect freshwater ecosystems (Häder et al., 2020).

Fish comprise 50% of the total number of existing vertebrate species and are widely recognized for their notable variation in species diversity. Hence, fish comprise a significant portion of the taxonomic diversity within the animal kingdom. Moreover, these organisms hold significant significance as primary sources of animal protein for human consumption, thereby possessing considerable economic worth. Nevertheless, the presence of human disturbances renders the richness and abundance of fish biodiversity in aquatic ecosystems increasingly susceptible. The study reported that the extinction rate of fish in fresh waters was higher (74.2%) compared to 47.4% for marine fish (Wiens, 2016). The negative impacts caused by human activities, including deforestation, river damming, overfishing, water contamination from plastic and domestic waste, illegal hunting, and degradation of habitats, have led to a significant decline in the diversity of fish species (Giannetto & Innal, 2021).

The identification and classification of fish species commonly rely on morphometric and meristic characteristics. Nonetheless, it is important to note that morphological characteristics may not exhibit consistent stability across different developmental stages. Additionally, the assessment of incomplete samples or the evaluation of rare and cryptic species often presents challenges. Moreover, the process of fish identification can present challenges owing to the resemblance in physical characteristics among closely related species during their early developmental stages, as well as inconsistencies found in the available literature and taxonomic records. These difficulties persist even when examining complete, fully-grown individuals. Furthermore, it is important to acknowledge that taxonomists may possess diverse levels of proficiency and expertise in identification. As a result, there is a possibility of inconsistent identification of the same specimen by different taxonomists, leading to potential confusion when attempting to summarize and compare data (Hulley et al., 2019). Nevertheless, environmental and conservation studies require a high degree of precision, necessitating the identification of specimens at the species level (Krishna Krishnamurthy & Francis, 2012). Because of the fundamental constraints of alpha taxonomy and the decreasing number of taxonomic experts, genetic techniques for fish species identification are required (Hopkins & Freckleton, 2002).

Molecular identification, which uses molecular markers to identify species, is commonly employed today. DNA barcoding, which relies on the analysis of mitochondrial DNA (mtDNA), is considered a highly appropriate method for species-level identification when compared to other molecular approaches commonly employed for species-level molecular identification (Panprommin et al., 2020). As opposed to morphological techniques, mtDNA-based molecular identification has various benefits. In the context of species identification, it is not imperative to possess complete specimens. Instead, a minute portion of tissue is appropriate for DNA extraction (Tapilatu et al., 2021). Furthermore, DNA is more durable and resistant to deterioration than physical characteristics (Moran et al., 2016), and some species with identical physical traits, such as cryptic or sister species, are difficult to recognize (Melo et al., 2016). Molecular identification can aid in the proper differentiation of such species (Feng et al., 2019). In addition, DNA remains constant throughout an organism's developmental phases. Morphological characteristics, on the other hand, might alter throughout a life cycle, leading to species misidentification (Tillett et al., 2012). Consequently, molecular techniques are able to be used to identify from the embryonic stage of fish to adult forms (Rathnasuriya et al., 2021). Being a professional conventional taxonomist takes a significant amount of patience, effort, and materials (Godfray, 2002). In addition, technological advances have made it relatively easy to analyze DNA sequences, coupled with the ability of bioinformatic applications can autonomously compare them, the level of training necessary for molecular identification is considerably lower compared to that required for morphological classification. Molecular identification is used extensively in many fields apart from species identification, such as the fight against wildlife trafficking and food fraud, the detection of biological invasions, and the tracking of biodiversity (van Der Heyde et al., 2020). The cytochrome oxidase subunit I (COI) can function as a DNA barcode for biological verification across invertebrates and vertebrates (Francisco et al., 2022). Research-

ers from various scientific institutions worldwide have collaboratively established the Fish Barcode of Life Initiative (FISBOL) in order to create a comprehensive repository of DNA barcodes for every fish species, thereby establishing a standardized reference library (Becker et al., 2011).

The complexity of problems in Indonesia's freshwater biodiversity, including global warming, increasing human population, riverbank damming, overconsumption/exploitation, pollutants, and invasive alien species, has led to a decline in freshwater biodiversity. Furthermore, it is estimated that 1,275 fish species in Indonesia have been listed as endangered (IUCN, 2022).

The Batanghari River is recognized as one of the most extensive rivers situated on the island of Sumatra. It flows through the provinces of Jambi and West Sumatra and has seven major river branches that contain tributaries and lakes that serve as fish juvenile breeding grounds. Geographically, it is located between 1°10'50.6"S 101°59'24.3"E and 1°23'12.4"S 103°59'08.1"E (http://www.geomapapp.org). The current state of the Batanghari River is alarming, as it is not only polluted by mercury-containing illegal gold mining waste (PETI) and industrial waste but also by household waste (Kaban et al., 2020). Consequently, several fish species in the Batanghari River have been declared extinct, with populations declining to rare or endangered levels (IUCN, 2022). A prior investigation revealed that assessments of fish populations were carried out in the Batanghari River, Sumatra, spanning the years 1994 to 2003. A total of 297 fish species in the Batanghari River, with 48 of them being newly documented records (45 of which were previously unrecorded in Sumatra). A total of six novel species have been documented under the taxonomic groups Cyprinidae, Nemacheilidae, and Cobitidae. The species *Diplocheilichthys*, *Diplocheilichthys jentinkii*, *Crossocheilus pseudobagroides*, *Osteochilus scapularis*, *Osteochilus vittatoides*, *Rasbora hosii*, and *Leptobarbus rubripinna* have been revalidated (Hui & Kottelat, 2009). To date, research pertaining to fish biodiversity in the Batanghari River has predominantly relied on traditional morphological criteria, with limited utilization of comprehensive barcoding techniques. The primary objective of this study is to compile the first DNA barcoding-based inventory of freshwater fish species in the Batanghari River, Jambi, Indonesia. Future studies can use this inventory as a reference for DNA sequence screening. In addition, the genetic diversity of various species of freshwater fish was evaluated. Researchers will be able to use the DNA barcode data produced by this study to monitor and preserve fish di-

versity in the Batanghari River in particular and Indonesia in general.

Materials and Methods

Samples collection

The sampling was conducted from October to November 2022. The study involved the collection of samples from five prominent watersheds along the Batanghari River, specifically in Muara Teluk Kayu Putih, Batanghari Tebo (Muara Tebo), Batanghari Tabir (Muara Tabir), Batanghari Tembesi (Muara Tembesi), and Batanghari Hilir. These locations were strategically chosen to encompass the entire stretch of the Batanghari River, spanning from its upstream to downstream regions (Fig. 1). Fish were caught using commercial fish gear (gill nets, scoop netting, trapping nets, ground cage nets, trammel nets, fishing nets, and fishing rods) by fishermen. Fish collected were measured in length and weighed, and photos were taken with a camera for further morphological identification. Subsequently, segments of the tail fin were carefully excised from the gathered specimens employing aseptic scissors and tweezers, put in tubes that included 70% ethanol solution for preservation, then transported to the Genomic Laboratory of Nasional Research and Innovation Institute (BRIN, Indonesia), and stored at room temperature until further processing.

Morphological identification

During the surveys, the fish specimens were collected either as live or relatively fresh specimens. The specimens obtained were first identified based on their morphological aspects. For that purpose, each fish specimen was characterized based on the measurement of morphometric characters (head, and body parts). Additional morphological characters, including the snout, mouth, opercle, occiput, fontanel, and eye, morphology of pectoral, dorsal, pelvic, anal, caudal, and adipose (if present) fins, and body shape profile, were also noted and photographed. The measurement of the morphological characters in identifying the species were conducted based on the species identification key and descriptions given by previous ichthyologists for Indonesian fish species (Kottelat, 1993). Additional references from other regions for the same fish species under examination were also used (Rainboth, 1996). When necessary, mainly when the fish species under examination has been taxonomically revised, recent related references were also used. The scientific names of the identified species were referred to the established taxonomy
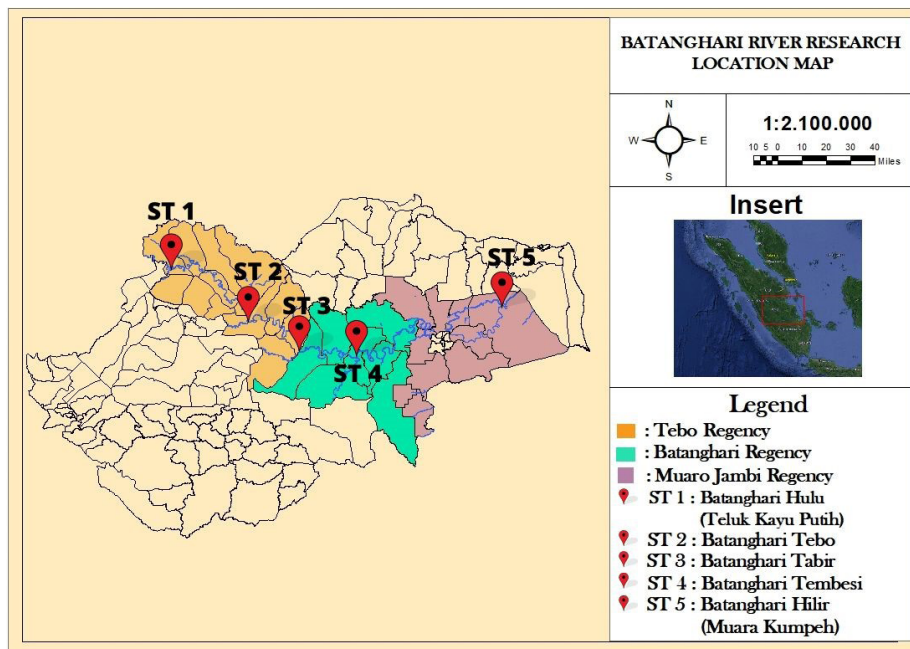
Fig. 1. Location of fish sampling (red pins) in Batanghari River, Jambi, Indonesia. The map was generated with ArcGIS 10.8.

reference (Hui & Kottelat, 2009) species following a catalogue of the fish species in Southeast Asia (Kottelat, 2013) and available information from www.fishbase.org (Froese & Pauly, 2023).

DNA extraction

The DNA was extracted from the caudal fin using a DNA extraction kit in accordance with the manufacturer's protocols (GeneJet Genomic DNA Purification, Thermo Scientific, Waltham, MA, USA). In order to evaluate the efficacy of the genomic DNA extraction procedure, the samples were subjected to mini-horizontal gel electrophoresis. The specimens were introduced into the agarose gel with a concentration of 1.5% (w/v), subjected to an electrical potential of 95 volts (equivalent to 7 volts per centimeter), and allowed to migrate for a duration of 50 minutes. Subsequently, the gel was subjected to staining using 2 μL of 1st Base FloroSafe DNA Stain (First Base Axil Scientific, Singapore). The concentration and purity of DNA were assessed using a NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer (Thermo Scientific) and subsequently stored at a temperature of –20℃ for subsequent analysis.

Cytochrome oxidase subunit I (COI) amplification and sequencing

The PCR standard was conducted in a final volume of 50 μL using the KOD One ® PCR Master Mix - Blue (Toyobo, Osaka, Japan) commercial kit, following the instructions provided by the manufacturer. A 655-bp fragment of the COI gene was amplified for species barcoding. The amplification was done using the primer pairs COI Fish-F (5'-TTC TCA ACTAAC-CAYAAAGAYATY GG-3') and COI-FishR (5'-TAGACT TCT GGG TGG CCR AAR AAY CA-3') as described by (Ward et al., 2005). The PCR mixture consisted of 25 μl of master mix, DNA template with a final concentration of 20 ng, and 1.5 μl of each primer with a final concentration of 0.3 μM. The reactions were conducted using a Veriti Thermal Cycler from Thermo Fisher Scientific. The thermal conditions for the experiment were as follows: an initial denaturation at 98℃ for 3 minutes, followed by 35 cycles. Each cycle consisted of a denaturation step at 98℃ for 10 seconds, an annealing temperature at 60℃ for 5 seconds, and an extension at 68℃ for 18 seconds. The experiment concluded with a final extension of 5 minutes at 72℃. The PCR products were separated on a 1.5% (w/v) agarose gel. The gel was stained using 1st Base FloroSafe DNA Stain from First Base Axil Scientific and the visualization was done using ultraviolet transillumination. The PCR products that showed satisfactory banding at approximately 655 bp were purified using the QI-Aquick Gel Extraction Kit from Qiagen. The purified products were then sequenced using the Applied Biosystems 3500 Ge-

netic Analyzer Sequencer, which is manufactured by Thermo Fisher Scientific. The sequencing services were provided by 1st BASE Axil Scientific Pte Ltd in Singapore. The resulting sequences were deposited in the NCBI GenBank genetic database and assigned the following accession numbers: OR357888, OR357890, OR357891, OR357892, OR288567, OR288570, OR288572, OR288574, OR288575, OR288576, OR288578, OQ151816, OQ151817, OQ151819, OQ151820, OQ151821, OQ151822, OQ151826, OQ151827, and OQ151828.

Molecular and bioinformatic data analysis

All sequences were aligned and edited in MEGA 11 (Tamura et al., 2021), and manual proofreading. All barcodes acquired were uploaded to GenBank databases. The identification of the COI barcode sequence for each sample was conducted using the BLAST and BOLD databases, with reference to the scientific name or species. The classification of specimens was conducted according to the fish taxonomic systems (Fricke et al., 2022). To determine their conservation status, their information was cross-referenced with the International Union for Conservation of Nature (IUCN) Red List of Threatened Species v. 2022-1 (Table 1). The Kimura two-parameter (K2P) distance (Kimura, 1980) was used to calculate sequence divergences. MEGA software 11 (Tamura et al., 2021) was used to generate an unrooted a neighbor-joining (NJ) tree based on K2P distances. K2P distances were calculated in the following categories: intraspecific distances, interspecific values within the same genus, and interspecific values between different genera within the same family. Using the K2P model and 1,000 bootstrap replications (Felsenstein, 2001), NJ tree was constructed for all COI sequences. MEGA 11 was responsible for both model testing and tree building.

Results

Morphology-based species identification

All collected specimens were initially identified using morphological techniques. Morphological identification categorized all specimens (1,692) into 36 species, 30 genera, and 16 families, representing five orders and all fish were identified to the species level (Table 1). In addition, of the identified sample, 58.69% belonged to the order Cypriniformes, Siluriformes (23.46%), and Perciformes (14.78%), respectively (Table 1). As per the 2022 edition of the Red List of Threatened Species published by the IUCN, *Chitala lopis* has been classified as an extinct species.

Furthermore, *Tor tambroides* has been designated as a species with a declining population. On the other hand, *Cyclocheilichthys apogon*, *Macrochirichthys macrochirus*, and *Trichopodus trichopterus* have been classified as species of least concern (LC), although their populations are also showing a decreasing trend. Furthermore, the other species have been classified as LC and data deficient (DD) (Table 1). The Muara Tebo had the highest percentage of fish abundance and diversity at 48.11%, followed by The Muara Teluk Kayu Putih at 45.52%. The Batanghari Merangin-Tembesi region had a lower percentage of fish abundance and diversity at 14.07%, while Batanghari Hilir (Muara Kumpeh) and Batanghari Tabir had even lower percentages at 12.64% and 6.57%, respectively (Table 2). Morphological characters were conducted by quantifying the dimensions of different anatomical regions (Table S1), and were also noted and photographed, (Fig. S1).

Identification of fish species using DNA barcodes

Of 1,692 specimens, we amplified 36 fish samples using PCR. All the sequences formed 650 bp after editing, and no stop codons, insertions, or deletions were found in any of the sequences. The minimum pairwise distance among sequences of observed fish samples based on the K2P was 0.003, and the maximum was 0.331. The average genetic divergence within genera, families, and orders, according to K2P, was 0.05, 0.12, and 0.16, respectively. Genetic distance among families showed a wide range, with the lowest genetic distance of 0.191 obtained between Botiidae and Cyprinidae and between Osphronemidae and Prestolepididae, while Polynemidae and Helostomatidae showed a genetic distance of 0.331 (Fig. 2). Interspecific distance was approximately 2.17 times greater than intraspecific distance.

The NJ tree revealed a clear division of the 16 fish families, with all species from the same family forming a monophyletic group (Fig. 3). Also, the phylogenetic tree indicates the presence of three main clades that represent the 16 fish families. The initial clade encompasses species from three families: Cyprinidae, Botidae, and Cobitidae. The second group included species from the families Synbanchidae, Siluridae, and Pangasiidae, while the third clade covered species from the families Osphronemidae, Prestolepididae, Helostomatidae, Anabantidae, Nototeridae, Eleotridae, Channidae, and Polynemidae. The accuracy of all species identifications based on GenBank ranged from 90.24% to 100.00% (Table 3).

Table 1. Classification, sample size and International Union for Conservation of Nature (IUCN) status of the fish species of Batanghari River, Jambi, Indonesia

Species No.	Species	Genus	Order	Family	Sample size	IUCN status
1	*Osteochilus vittatus*	*Osteochilus*	Cypriniformes	Cyprinidae	152	Least corcern
2	*Osteochilus hasseltii*	*Osteochilus*	Cypriniformes	Cyprinidae	20	Least corcern
3	*Osteochilus kappenii*	Osteochilus	Cypriniformes	Cyprinidae	5	Data deficient
4	*Leptobarbus hoevenii*	Leptobarbus	Cypriniformes	Cyprinidae	28	Least corcern
5	*Labiobarbus leptocheilus*	Labiobarbus	Cypriniformes	Cyprinidae	154	Least corcern
6	*Barbonymus schwanefeldii*	Barbonymus	Cypriniformes	Cyprinidae	61	Least corcern
7	*Barbonymus gonionotus*	Barbonymus	Cypriniformes	Cyprinidae	20	Least corcern
8	*Puntioplites waandersi*	Puntioplites	Cypriniformes	Cyprinidae	123	Data deficient
9	*Labeo chrysophekadion*	Labeo	Cypriniformes	Cyprinidae	1	Least corcern
10	*Rasbora lateristriata*	Rasbora	Cypriniformes	Cyprinidae	130	Least corcern
11	*Thynnichthys thynnoides*	Thynnichthys	Cypriniformes	Cyprinidae	264	Least corcern
12	*Macrochirichthys macrochirus*	Macrochirichthys	Cypriniformes	Cyprinidae	13	Least corcern
13	*Cyclocheilichthys apogon*	Cyclocheilichthys	Cypriniformes	Cyprinidae	17	Least corcern
14	*Tor tambroides*	Tor	Cypriniformes	Cyprinidae	1	Least corcern
15	*Tor douronensis*	Tor	Cypriniformes	Cyprinidae	1	Least corcern
16	*Syncrossus hymenophysa*	Syncrossus	Cypriniformes	Botiidae	1	Least corcern
17	*Pangio semicincta kuhlii*	Pangio	Cypriniformes	Cobitidae	2	Least corcern
Total cypriniformes					993 (58.69%)	
18	*Channa striata*	Channa	Perciformes	Channidae	35	Least corcern
19	*Channa micropeltes*	Channa	Perciformes	Channidae	11	Least corcern
20	*Polynemus dubius*	Polynemus	Perciformes	Polynemidae	26	Least corcern
21	*Trichogaster trichopterus*	Trichogaster	Perciformes	Osphronemidae	10	Least corcern
22	*Anabas testudineus*	Anabas	Perciformes	Anabantidae	67	Least corcern
23	*Pristolepis fasciata*	Pristolepis	Perciformes	Prestolepididae	65	Least corcern
24	*Oxyeleotris marmorata*	Oxyeleotris	Perciformes	Eleotrididae	11	Least corcern
25	*Helostoma temminckii*	Helostoma	Perciformes	Helostomatidae	25	Least corcern
Total perciformes					250 (14.78%)	
26	*Pseudeutropis brachypopterus*	Pseudeutropis	Siluriformes	Pangasiidae	16	Least corcern
27	*Pangasius nasutus*	Pangasius	Siluriformes	Pangasiidae	1	Least corcern
28	*Belodontichthys dinema*	Belodontichthys	Siluriformes	Siluridae	26	Least corcern
29	*Belodontichthys* sp	Belodontichthys	Siluriformes	Siluridae	1	Least corcern
30	*Kryptopterus bicirrhis*	Kryptopterus	Siluriformes	Siluridae	34	Least corcern
31	*Hemibagrus nemurus*	Hemibagrus	Siluriformes	Bagridae	161	Least corcern
32	*Mystus singaringan*	Mystus	Siluriformes	Bagridae	126	Least corcern
33	*Clarias batrachus*	Clarias	Siluriformes	Clariidae	32	Least corcern
Total siluriformes					397 (23.46%)	
34	*Monopterus albus*	Monopterus	Synbranchiformes	Synbranchidae	49	Least corcern
	Total synbranchiformes				49 (2.90%)	
35	*Chitala lopis*	Chitala	Osteoglossiformes	Notopteridae	1	Extinct
36	*Notopterus notopterus*	Notopterus	Osteoglossiformes	Notopteridae	2	Least corcern
Total osteoglossiformes					3 (0.18%)	
Total sample size					1,692	

Table 2. Classification, sample size and number of fish was found the sampling areas within the Batanghari River in Jambi, Indonesia

No.	Species	Sampling locations					Total
		Muara Teluk Kayu Putih	Batanghari Tebo	Batanghari Tabir	Batanghari Merangin-Tembesi	Batanghari Hilir (Muara Kumpeh)	
1	*Osteochilus vittatus*	25	97	8	6	16	
2	*Osteochilus hasseltii*		20				
3	*Osteochilus kappenii*		5				
4	*Leptobarbus hoevenii*	10	18				
5	*Labiobarbus leptocheilus*	33	37	25	32	27	
6	*Barbonymus schwanefeldii*	10	18	3	14	16	
7	*Barbonymus gonionotus*	7	10	0	2	1	
8	*Puntioplites waandersi*	11	77	5	16	14	
9	*Labeo chrysophekadion*		1				
10	*Rasbora lateristriata*	26	30	24	23	27	
11	*Thynnichthys thynnoides*	63	92	23	56	30	
13	*Macrochirichthys macrochirus*		13				
14	*Cyclocheilichthys apogon*		17				
15	*Tor tambroides*	1					
16	*Tor douronensis*		1				
17	*Syncrossus hymenophysa*				1		
18	*Pangio semicincta kuhlii*				1		
19	*Channa striata*	9	11	2	4	9	
20	*Channa micropeltes*		11				
21	*Polynemus dubius*	7	19				
22	*Trichopodus trichopterus*		10				
23	*Anabas testudineus*		67				
24	*Pristolepis fasciata*	18	21		14	13	
26	*Oxyeleotris marmorata*	2	9				
27	*Helostoma temminckii*		25				
28	*Pseudeutropis brachypopterus*		10		6		
29	*Pangasius nasutus*				1		
30	*Belodontichthys dinema*	5	15		6		
31	*Belodontichthys* sp		1				
32	*Kryptopterus bicirrhis*	10	20		4		
33	*Hemibagrus nemurus*	40	55	15	25	26	
34	*Mystus singaringan*	32	40	4	23	27	
35	*Clarias batrachus*	6	13	2	3	8	
36	*Monopterus albus*		49				
37	*Chitala lopis*		1				
38	*Notopterus notopterus*		1		1		
	Total	315	814	111	238	214	1,692

	Anabantidae	Cyprinidae	Siluridae	Channidae	Clariidae	Helostomatidae	Bagridae	Synbranchidae	Nototeridae	Eleotrididae	Pangasiidae	Cobitidae	Polynemidae	Prestolepididae	Botiidae	Osphronemidae
Anabantidae																
Cyprinidae	0.270															
Siluridae	0.249	0.223														
Channidae	0.206	0.211	0.238													
Clariidae	0.278	0.206	0.215	0.236												
Helostomatidae	0.284	0.273	0.294	0.275	0.295											
Bagridae	0.250	0.233	0.245	0.238	0.225	0.291										
Synbranchidae	0.258	0.235	0.259	0.256	0.226	0.288	0.255									
Nototeridae	0.264	0.248	0.271	0.224	0.256	0.275	0.255	0.228								
Eleotrididae	0.272	0.244	0.281	0.210	0.259	0.273	0.271	0.219	0.229							
Pangasiidae	0.260	0.240	0.246	0.220	0.250	0.276	0.190	0.252	0.233	0.235						
Cobitidae	0.282	0.220	0.259	0.216	0.265	0.261	0.248	0.262	0.269	0.238	0.241					
Polynemidae	0.279	0.255	0.287	0.255	0.292	0.331	0.279	0.274	0.195	0.264	0.241	0.288				
Prestolepididae	0.264	0.262	0.269	0.240	0.304	0.251	0.305	0.263	0.276	0.281	0.283	0.250	0.285			
Botiidae	0.270	0.191	0.259	0.210	0.223	0.251	0.246	0.245	0.236	0.195	0.238	0.246	0.247	0.240		
Osphronemidae	0.244	0.258	0.289	0.244	0.281	0.243	0.281	0.278	0.266	0.290	0.278	0.257	0.296	0.191	0.242	

The average genetic divergence

Genera	Families	Order
0.05	0.12	0.16

Fig. 2. Distribution pairwise genetic distance (Kimura two-parameter model, K2P) among 16 fish families from Batanghari River.

Species on tree (with bootstrap support and distances):

- *Trichogaster trichopterus* 0.010732 — 73, 43
- *Pristolepis fasciata* 0.08849
- *Helostoma temminkii* 0.10391 — 7
- *Anabas testudineus* 0.10984 — 93
- Chitala lopis -0.00479 — 100
- *Notopterus notopterus* 0.00479 — 64
- *Oxyeleotris marmorata* 0.10297 — 63
- *Channa micropletes* 0.08592 — 93
- *Channa striata* 0.07634
- *Polynemus dubius* 0.10716
- *Monopterus albus* 0.10424 — 82
- *Belodontichthys dinema* 0.04366 — 100
- *Belodontichthys sp* 0.08104 — 74
- *Kriptoterus bicirrhis* 0.07415 — 21
- *Clarias batrachus* 0.09012 — 95
- *Mystus singaringan* 0.08136 — 95
- *Pseudeutropius brachypopterus* 0.07235 — 85
- *Pangasius nasutus* 0.05931 — 94
- *Hemibagrus nemurus* 0.07286
- *Labiobarbus leptocheilus* 0.0692 — 23
- *Osteochilus hasselti* 0.00488 — 100
- *Osteochilus vittatus* -0.00212 — 93
- *Osteochilus sp* 0.00212 — 84
- *Labeo chrysophekadion* 0.09499 — 92
- *Thynnichthys thynnoides* 0.0539 — 95
- *Tor douronensis* 0.01169 — 100
- *Tor tambroides* 0.01806 — 69
- *Cyclocheilichthys apogon* 0.0497 — 95
- *Puntioplites waandersi* 0.06242 — 36
- *Barbonymus schwanenfeldii* 0.00034 — 100
- *Barbonymus gonionotus* 0.00501 — 77
- *Leptobarbus hoevenii* 0.06893 — 34
- *Macrochirichthys macrochirus* 0.08394 — 76
- *Rabora lateristriata* 0.08406 — 99
- *Syncrossus hymenophysa* 0.09378 — 95
- *Pangio kuhlii* 0.09378

Family labels (right side): Osphronemidae, Prestolepididae, Helostomatidae, Anabantidae, Nototeridae, Eleotrididae, Channidae, Polynemidae, Synbranchidae, Siluridae, Bangridae, Pangasiidae, Bangridae, Cyprinidae, Botiidae, Cobitidae
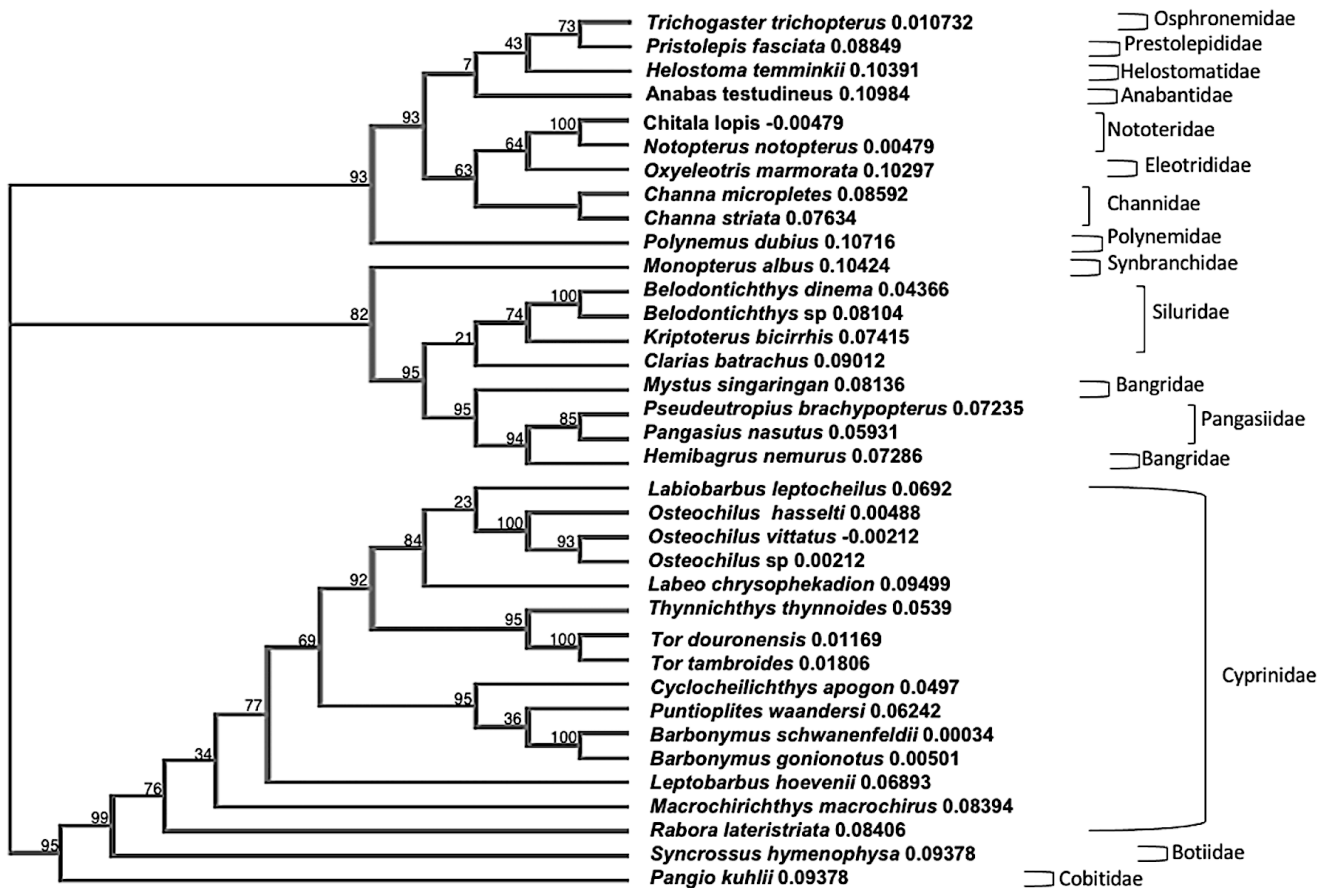
Fig. 3. Neighbor-joining (NJ) tree based on cytochrome oxidase subunit I (COI) barcodes from fishes of the Batanghari River.
The numbers next to the species names represent bootstrap support in 1,000 replications.

Table 3. Comparison of the Batanghari River fish cytochrome oxidase subunit I (COI) sequence with published fish species COI sequences using NCBI blastn

Scientific name	Per. Ident	Closely related to	Accession
Anabas testudineus	99.66%	*Anabas testudineus*	MG407353.1
Barbonymus gonionotus	100%	*Barbonymus gonionotus*	OR357888.1
Barbonymus schwanenfeldii	100%	*Barbonymus schwanenfeldii*	OR288576.1
Belodontichthys dinema	100%	*Belodontichthys dinema*	OR288567.1
Belodontichthys sp	90.24%	*Belodontichthys truncatus*	KY607139.1
Channa micropeltes	95.62%	*Channa micropeltes*	MF496867.1
Channa stiata	100%	*Channa stiata*	OR288578.1
Clarias batrachus	100%	*Clarias batrachus*	OQ151820.1
Cyclocheilichthys apogon	100%	*Cyclocheilichthys apogon*	OR288570.1
Helostoma temmimkii	100%	*Helostoma temmimkii*	MK120521.1
Hemibagrus nemurus	100%	*Hemibagrus nemurus*	OR357890.1
Kriptoterus bicirrhis	100%	*Kriptoterus bicirrhis*	OQ151816.1
Labeo chrysophekadion	98.65%	*Labeo chrysophekadion*	MK049387.1
Labiobarbus leptocheilus	100%	*Labiobarbus leptocheilus*	MN342618.1
Leptobarbus hoevenii	100%	*Leptobarbus hoevenii*	OR288575.1
Macrochirichthys macrochirus	100%	*Macrochirichthys macrochirus*	OQ151821.1
Monopterus albus	100%	*Monopterus albus*	OQ151825.1
Mystus singaringan	100%	*Mystus singaringan*	MN992971.1
Chilata lopis	100%	*Chilata lopis*	OR288572.1
Notopterus notopterus	100%	*Notopterus notopterus*	KU692675.1
Osteochilus hasselti	100%	*Osteochilus hasselti*	OR357892.1
Osteochilus sp	99.66%	*Osteochilus* sp	JX074151.1
Osteochilus vittatus	100%	*Osteochilus vittatus*	OQ151828.1
Oxyeleotris marmorata	100%	*Oxyeleotris marmorata*	MK448189.1
Pangasius nasutus	100%	*Pangasius nasutus*	OQ151819.1
Pangio semicincta kuhlii	100%	*Pangio kuhlii*	OQ151817.1
Polynemus dubius	100%	*Polynemus dubius*	OQ151822.1
Pristolepis fasciata	100%	*Pristolepis fasciata*	OQ151826.1
Pseudeutropius brachypopterus	100%	*Pseudeutropius brachypopterus*	OQ151827.1
Puntioplites waandersi	100%	*Puntioplites waandersi*	HM536928.1
Rasbora lateristriata	100%	*Rasbora lateristriata*	LC130652.1
Syncrossus hymenophysa	100%	*Syncrossus hymenophysa*	NC033951.1
Thynnichthys thynnoides	96.28%	*Thynnichthys thynnoides*	MK448126.1
Tor douronensis	100%	*Tor douronensis*	OR288574.1
Tor tambroides	100%	*Tor tambroides*	KT001033.1
Trichogaster trichopterus	99.66%	*Trichogaster trichopterus*	KC789557.1

Discussion

A 650-bp-long fragment of the mitochondrial COI gene was used to generate 37 DNA barcodes from 36 species and 16 families. These 16 families represented 25% of the region's reported fish biodiversity in the Batanghari River (Hui & Kottelat, 2009). This might be because the study was carried out at a number of sampling points, including not only the five major watersheds of the Batanghari River but also all the small rivers and lakes that flow into it. Furthermore, the duration of time required

for research to catch fish influences the number of specimens obtained. Moreover, fishing was more challenging when waters were still relatively high in October-November (the wet season), resulting in fewer fish species being caught. It was also indicated that the Batanghari River has been polluted by PETI and industrial waste, as well as by domestic waste (Kaban et al., 2020).

This study clarified the taxonomic status of several taxa, including *Osteochilus vittatus* and *Osteochilus hasseltii*, both of which are members of the *Osteochilus vittatus* group, which is consistent with the findings (Hui & Kottelat, 2009). Another *Osteochilus* sp. found in Muara Tebo, Batanghari River, is an unnamed species. We identified *Pangio semicincta* based on morphological analysis, but when we compared it to Pangio kuhli (99.68%) based on COI analysis, *Pangio semicincta* and Pangio kuhli are both members of the *Pangio semicincta* group (Kottelat, 2013).

When BLAST searches were used to compare the COI sequences from the 36 species that were gathered, it was found that a lot of them (71.05%) were very similar to GenBank references. However, it is worth noting that a number of the acquired sequences exhibited a lower degree of sequence similarity (ranging from 90% to 96%) when compared to reference sequences sourced from the GenBank database. This includes *Channa micropeltes*, which displayed a sequence similarity of 95.62%, and *Belodontichthys truncatus*, which exhibited a sequence similarity of 90.24% (Table 3). The identification of these species may have been limited to the level of genera, as has been demonstrated in several prior investigations (Panprommin et al., 2019). The absence of reference sequences that exhibit complete similarity in the GenBank database leads to their designation as "sp" (Hinchliff & Smith, 2014).

Furthermore, it is worth noting that the mean K2P distance exhibited a higher value among families (0.274) compared to genera (0.256). The observed augmentation in genetic variation corresponded with the escalation in taxonomic hierarchy, suggesting a noticeable divergence in genetic differentiation across genera boundaries. In accordance with the findings of (Sheraliev & Peng, 2021), it was observed that the pairwise genetic distance K2P exhibited an upward trend in relation to the taxonomic hierarchy. Specifically, the values of 0.05, 0.12, and 0.16 were recorded for genera, families, and orders, respectively.

Phylogenetic tree analyses showed a clear relationship between species, by grouping closely related species together in sub-clusters and spacing out distantly related ones. Although one of the goals of barcoding is to identify species boundaries, there is in fact a phylogenetic signal in the data contained in the COI gene. Species of the same genus continued to group together and, in most cases, surpassed the confamilial species cluster (Bariche et al., 2015). In addition, the topology of the NJ tree illustrates species' relationships (Fig. 3). NJ tree test revealed a clear clustering pattern, providing and informative phylogenetic relationship between taxonomic levels. All fish samples were grouped and separated into distinct clades in accordance with morphological identification and the BLAST algorithm.

The analysis of fish distribution revealed that the upstream segment of the river (Muara Teluk Kayu Putih and Batahnghari Tebo) had greater diversity and species richness, as evidenced by the high values of diversity and richness indices (Tabel 2). The upstream section of the river (Batanghari Tembesi, Batanghari Merangin-Tabir, and Batanghari Hilir) exhibited a comparatively lower degree of human disturbances in contrast to the downstream region, which is characterised by urbanisation and the presence of various pollutants such as industrial waste, domestic waste, and illegal gold mining waste containing mercury. Consequently, the downstream river has become a site of accumulation for these pollutants (Febrianti et al., 2023; Kaban et al., 2020). Moreover, the Batanghari Tebo site demonstrated a significantly greater proportion of fish abundance and diversity in comparison to other locations, which could likely be attributed to the heightened prevalence of macrozoobenthos in this specific area (Febrianti et al., 2023). The preservation of global fish diversity is currently facing significant challenges. In addition to inherent factors that naturally limit species populations, there is an increasing recognition of the detrimental effects caused by introduced species (Radočaj et al., 2021). Concurrently, there is a growing escalation in the detrimental effects of anthropogenic factors on the biodiversity of freshwater rivers (Tickner et al., 2020). The annual decline in the number of biological species has led to the increasing utilization of DNA barcoding as a versatile method for assessing fish biodiversity, observing fish conservation efforts, and managing fisheries (Sheraliev & Peng, 2021). The application of DNA barcoding in our study provides significant benefits to the taxonomic classification of fish species in the Batanghari River. Additionally, it is crucial to address the taxonomic ambiguity surrounding misidentified invasive species that have successfully acclimated to this particular river ecosystem.

Conclusion

The present work revealed the current status of fish diversity

in Batanghari River, consisting of 1,692 specimens, 16 families, and 5 orders based on morphological analysis. DNA barcoding COI confirmed that a total of 36 species and 16 families were identified from 36 fish specimens. Our results showed that DNA barcodes using the COI gene sequence was an efficient approach to identify these fish species. This information may be useful for future studies to verify the declining fish diversity issue and managing the fisheries resource for sustainability in Batanghari River.

Supplementary materials

Competing interests

Funding sources

Acknowledgements

Availability of data and materials

Upon reasonable request, the datasets of this study can be available from the corresponding author.

Ethics approval and consent to participate

This study conformed to the guidance of animal ethical treatment for the care and use of experimental animals.

ORCID

Huria Marnis	https://orcid.org/0000-0002-2152-9953
Khairul Syahputra	https://orcid.org/0000-0002-1429-8421
Jadmiko Darmawan	https://orcid.org/0000-0002-8467-471X
Dwi Febrianti	https://orcid.org/0000-0002-1616-5550
Evi Tahapari	https://orcid.org/0000-0002-7681-5762
Sekar Larashati	https://orcid.org/0000-0001-6967-3172
Bambang Iswanto	https://orcid.org/0000-0003-2212-3565
Erma Primanita Hayuningtyas Primanita	
	https://orcid.org/0000-0002-7350-669X
Mochamad Syaifudin	https://orcid.org/0000-0002-2586-9672
Arsad Tirta Subangkit	https://orcid.org/0009-0001-0519-6925

References

Bariche M, Torres M, Smith C, Sayar N, Azzurro E, Baker R, et al. Red Sea fishes in the Mediterranean Sea: a preliminary investigation of a biological invasion using DNA barcoding. J Biogeogr. 2015;42:2363-73.

Becker S, Hanner R, Steinke D. Five years of FISH-BOL: brief status report. Mitochondrial DNA. 2011;22:3-9.

Febrianti D, Darmawan J, Marnis H, Syahputra K, Tahapari E, Larashati S, et al. Assessment of spatial variations in water quality, plankton and macrozoobenthos diversity of Batanghari River, Indonesia. Aquac Aquar Conserv Legis. 2023;16:1519-30.

Felsenstein J. PHYLIP: Phylogeny Inference Package, version 36. Seattle, WA: University of Washington; 2001.

Feng C, Tang Y, Liu S, Tian F, Zhang C, Zhao K. Multiple convergent events created a nominal widespread species: *Triplophysa stoliczkae* (Steindachner, 1866) (Cobitoidea: Nemacheilidae). BMC Evol Biol. 2019;19:1-12.

Francisco SM, Lima CS, Moreira I, Shahin AAB, Faleh AB. DNA barcoding of commercially relevant marine fish species in Tunisian waters. J Mar Biol Assoc UK. 2022;102:178-85.

Fricke R, Eschmeyer WN, Van der Laan R. ECoF. Eschmeyer's catalog of fishes: genera, species, references. San Francisco: California Academy of Sciences; 2022.

Froese R, Pauly D. FishBase [Internet]. 2023 [cited 2023 Aug 1]. http://www.fishbase.org

Giannetto D, Innal D. Status of endemic freshwater fish fauna inhabiting major lakes of Turkey under the threats of climate change and anthropogenic disturbances: a review. Water. 2021;13:1534.

Gillette DP, Edds DR, Jha BR, Mishra B. Thirty years of environmental change reduces local, but not regional, diversity of riverine fish assemblages in a Himalayan biodiversity

hotspot. Biol Conserv. 2022;265:109427.

Godfray HCJ. Challenges for taxonomy. Nature. 2002;417:17-9.

Häder DP, Banaszak AT, Villafañe VE, Narvarte MA, González RA, Walter Helbling E. Anthropogenic pollution of aquatic ecosystems: emerging problems with global implications. Sci Total Environ. 2020;713:136586.

Hinchliff CE, Smith SA. Some limitations of public sequence data for phylogenetic inference (in plants). PLOS ONE. 2014;9:e98986.

Hopkins GW, Freckleton RP. Declines in the numbers of amateur and professional taxonomists: implications for conservation. Anim Conserv Forum. 2002;5:245-249.

Hui TH, Kottelat M. The fishes of the Batang Hari drainage, Sumatra, with description of six new species. Ichthyol Explor Freshw. 2009;20:13-69.

Hulley EN, Tharmalingam S, Zarnke A, Boreham DR. Development and validation of probe-based multiplex real-time PCR assays for the rapid and accurate detection of freshwater fish species. PLOS ONE. 2019;14:e0210165.

International Union for Conservation of Nature [IUCN]. The IUCN red list of threatened species. Version 2022-1 [Internet]. IUCN Red List of Threatened Species. 2022 [cited 2023 Jan 1]. https://www.iucnredlist.org/

Kaban S, Armanto ME, Ridho MR, Hariani PL. Heavy metal (Mercury and Plumbum) accumulation of two fish species in Sipin and Teluk Lake, Jambi Province. Ecol Environ Conserv. 2020;26:1119-23.

Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 1980;16:111-20.

Kottelat M. A review of the eeloaches of genus Pangio (Teleostei: Cobitidae) from the Malay Peninsula, with description of six new species. Raffles Bull Zool. 1993;41:203-49.

Kottelat M. The fishes of the inland waters of Southeast Asia: a catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. Raffles Bull Zool. 2013.

Krishna Krishnamurthy P, Francis RA. A critical review on the utility of DNA barcoding in biodiversity conservation. Biodivers Conserv. 2012;21:1901-19.

Mann PJ, Strauss J, Palmtag J, Dowdy K, Ogneva O, Fuchs M, et al. Degrading permafrost river catchments and their impact on Arctic Ocean nearshore processes. Ambio. 2022;51:439-55.

Melo BF, Ochoa LE, Vari RP, Oliveira C. Cryptic species in the

neotropical fish genus *Curimatopsis* (Teleostei, Characiformes). Zool Scr. 2016;45:650-8.

Moran Z, Orth DJ, Schmitt JD, Hallerman EM, Aguilar R. Effectiveness of DNA barcoding for identifying piscine prey items in stomach contents of piscivorous catfishes. Environ Biol Fishes. 2016;99:161-7.

Panprommin D, Soontornprasit K, Tuncharoen S, Iamchuen N. The utility of DNA barcoding for the species identification of larval fish in the lower Ing river, Thailand. Turk J Fish Aquat Sci. 2020;20:671-9.

Panprommin D, Soontornprasit K, Tuncharoen S, Pithakpol S, Keereelang J. DNA barcodes for the identification of species diversity in fish from Kwan Phayao, Thailand. J Asia-Pac Biodivers. 2019;12:382-9.

Radočaj T, Špelić I, Vilizzi L, Povž M, Piria M. Identifying threats from introduced and translocated non-native freshwater fishes in Croatia and Slovenia under current and future climatic conditions. Glob Ecol Conserv. 2021;27:e01520.

Rainboth WJ. Fishes of the Cambodian Mekong. FAO Species Identification Field Guide for Fishery Purpose. Rome: Food and Agriculture Organization (FAO); 1996. p. 310.

Rathnasuriya MIG, Mateos-Rivera A, Skern-Mauritzen R, Wimalasiri HBU, Jayasinghe RPPK, Krakstad JO, et al. Composition and diversity of larval fish in the Indian Ocean using morphological and molecular methods. Mar Biodivers. 2021;51:1-15.

Sheraliev B, Peng Z. Molecular diversity of Uzbekistan's fishes assessed with DNA barcoding. Sci Rep. 2021;11:16894.

Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Bio Evol. 2021;38(7):3022-7.

Tapilatu RF, Tururaja TS, Siriyadi, Kusuma AB. Molecular phylogeny reconstruction of grouper (Serranidae: Epinephelinae) at Northern part of Bird's Head Seascape - Papua inferred from COI gene. Fish Aquat Sci. 2021;24:181-90.

Tickner D, Opperman JJ, Abell R, Acreman M, Arthington AH, Bunn SE, et al. Bending the curve of global freshwater biodiversity loss: an emergency recovery plan. BioScience. 2020;70:330-42.

Tillett BJ, Field IC, Bradshaw CJA, Johnson G, Buckworth RC, Meekan MG, et al. Accuracy of species identification by fisheries observers in a north Australian shark fishery. Fish Res. 2012;127–128:109-15.

van der Heyde M, Bunce M, Wardell-Johnson G, Fernandes K,

White NE, Nevill P. Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. Mol Ecol Resour. 2020;20:732-45.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. DNA barcoding Australia's fish species. Philos Trans R Soc B Biol Sci. 2005;360:1847-57.

Wiens JJ. Climate-related local extinctions are already widespread among plant and animal species. PLOS Biol. 2016;14:e2001104.

Worm B, Lotze HK. Marine biodiversity and climate change. In: Letcher T, editor. Climate change: observed impacts on planet Earth. Amsterdam: Elsevier; 2021. p. 445-64.