Sains Malaysiana 54(6)(2025): 1451-1464
http://doi.org/10.17576/jsm-2025-5406-02

DNA Barcoding Shows Taxonomic Uncertainty in North Sumatran Mahseer, *Neolissochilus*: towards Comprehensive Revision in Indonesia

(Pengekodan Bar DNA Menunjukkan Ketidakpastian Taksonomi pada Mahseer Sumatera Utara, *Neolissochilus*: Ke Arah Semakan Komprehensif di Indonesia)

SEKAR LARASHATI[1,*], GEMA WAHYUDEWANTORO[2], MEY RISTANTI WIDORETNO[1], DWI FEBRIANTI[1], HURIA MARNIS[3], TERNALA ALEXANDER BARUS[4], YOLANDA SIMARETHA CHRISTIANY SINUHAJI[5], RUBY VIDIA KUSUMAH[3], OCTAVIANTO SAMIR[1] & TRI WIDIYANTO[1]

[1]*Research Center for Limnology and Water Resources, National Research and Innovation Agency, Cibinong, Jawa Barat, Indonesia*
[2]*Research Center for Biosystematics and Evolution, National Research and Innovation Agency, Cibinong, Jawa Barat, Indonesia*
[3]*Research Center for Fisheries, National Research and Innovation Agency, Cibinong, Jawa Barat, Indonesia*
[4]*Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Sumatera Utara, Sumatra Utara, Indonesia*
[5]*Department of Natural Resources and Environmental Management, Graduate School, Universitas Sumatera Utara, Sumatra Utara, Indonesia*

Received: 4 November 2024/Accepted: 7 April 2025

ABSTRACT

The mahseer species *Neolissochilus* holds significant ecological and socioeconomic value, but its wild population is declining. A major challenge in conservation efforts is taxonomy uncertainty. This study used DNA barcoding of three mitochondrial DNA (mtDNA) segments (COI, *Cyt b*, and 16S rRNA) to verify the morphological identification of nine specimens collected from Bonan Dolok River and eight *Neolissochilus sumatranus* from the Tulas River, as well as three *Neolissochilus soro* samples from the Tulas River, and eight from Boho Rivers, respectively, in North Sumatra. Morphological identification was based on body height, eye diameter/length, length of pectoral fin to the dorsal fin, anal fin to the caudal fin, and dorsal fin branch rays. Neighbour Joining analysis was used to construct the phylogenetic trees, showing that samples of *N. sumatranus* and *N. soro* clustered with supported bootstrap values of 63-71%, and no genetic distance between them. ASAP and ABGD species delimitation supported this clustering. This suggests both are the same species and closely related to *N. soroides* and *N. hendersoni* (genetic distance: 0.001-0.012 and 0.000-0.002, respectively). This challenges existing taxonomy and emphasizes the need to revisit *Neolissochilus* classification in Indonesia. Further study involving traditional taxonomy and DNA barcoding on *Neolissochilus* species in Indonesia is needed to clarify species distinction, validate taxonomy, and update the conservation status. This approach will enhance species identification, guide conservation efforts, and improve management of these vital freshwater fish species.

Keywords: Conservation; freshwater fish; mitochondrial DNA; phylogenetics; taxonomy validation

ABSTRAK

Spesies tengas *Neolissochilus* mempunyai nilai ekologi dan sosioekonomi yang penting, tetapi populasinya di habitat liar semakin berkurang. Salah satu cabaran utama dalam usaha pemuliharaannya adalah ketidakpastian dalam taksonomi. Kajian ini menggunakan pengekodan DNA bagi tiga segmen DNA mitokondria (mtDNA) (COI, *Cyt b* dan 16S rRNA) untuk mengesahkan pengecaman morfologi sembilan sampel yang dikumpulkan dari Sungai Bonan Dolok dan lapan *Neolissochilus sumatranus* dari Sungai Tulas, serta tiga sampel *Neolissochilus soro* dari Sungai Tulas dan lapan dari Sungai Boho di Sumatera Utara. Pengecaman morfologi dibuat berdasarkan ketinggian badan, diameter/panjang mata, panjang sirip pektoral ke sirip dorsal, sirip anal ke sirip kaudal dan jejari bercabang sirip dorsal. Analisis Jiran Menyambung digunakan untuk membina pokok filogenetik, yang menunjukkan bahawa sampel *N. sumatranus* dan *N. soro* berkelompok dengan nilai sokongan *bootstrap* 63-71% dan tiada jarak genetik antara mereka. Persempadanan spesies ASAP dan ABGD menyokong pengelompokan ini. Ini mencadangkan bahawa kedua-duanya adalah spesies yang sama dan mempunyai hubungan rapat dengan *N. soroides* (dan *N. hendersoni*) dengan jarak genetik masing-masing 0.001–0.012 dan 0.000–0.002. Penemuan ini mencabar taksonomi sedia ada dan menekankan keperluan untuk mengkaji semula pengelasan *Neolissochilus* di Indonesia. Kajian lanjut yang menggabungkan taksonomi tradisional dan pengekodan DNA terhadap

spesies *Neolissochilus* di Indonesia diperlukan untuk menjelaskan perbezaan spesies, mengesahkan taksonomi dan mengemaskini status pemuliharaan. Pendekatan ini akan meningkatkan pengecaman spesies, membimbing usaha pemuliharaan dan memperbaiki pengurusan spesies ikan air tawar yang penting ini.

Kata kunci: DNA mitokondria; filogenetik; ikan air tawar; pemuliharaan; pengesahan taksonomi

INTRODUCTION

Neolissochilus is a genus of mahseers residing across Asia including in Indonesian waters. Despite receiving less attention than the true mahseer, *Tor*, the presence of *Neolissochilus* is significant in terms of socioeconomic value and conservation. Local people in some areas of Indonesia, such as North Sumatra and West Java, consider the fish sacred and consumed during traditional ceremonies. The mahseers are valued for their texture and taste, as well as their low fat and high protein content (Chasanah et al. 2021; Khai et al. 2015). The price of mahseers, including *Neolissochilus* spp., in North Sumatra, is high, ranging from USD25 (Rp400,000) to USD31 (Rp500,000) per kg, while in Java, the price can be up to USD62 (Rp1,000,000) (Rumondang, Fuah & Aidil Huda 2023).

There are 33 species of *Neolissochilus* across their geographic distribution and four species are distributed in Indonesia (Froese & Pauly 2024; Kottelat 2013; Kottelat, Whitten & Kartikasari 1993). *Neolissochilus longipinnis*, *N. sumatranus*, *N. thienemanni*, and *N. soro* are the valid *Neolissochilus* living in Indonesian freshwater. There have been no publications reporting *N. longipinnis* and *N. thienemanni* in the past decade, apart from a few mentions of *N. thienemanni* being collected from the rivers in North Sumatra and Aceh (Nasir, Munira & Muchlisin 2018; Rachmad, Sihombing & Sabariyah 2019). However, lack of morphological identification information in the methods by Rachmad, Sihombing and Sabariyah (2019) raises doubt regarding the species identification, and the identified specimens by Nasir, Munira and Muchlisin (2018) could not be confirmed by personal communication.

Mahseers, including *Neolissochilus*, prefer to inhabit clear, rocky waters with slow to swift currents and are sensitive to environmental changes (Ali et al. 2014; Hoàng et al. 2015). The presence of mahseers in their natural habitats can indicate the overall health of freshwater ecosystems (Everard et al. 2021). Therefore, protecting mahseers is essential for preserving many others of the same habitat. In North Sumatra, several rivers are home to mahseers, *N. sumatranus* and *N. soro* or previously known as *Tor soro* (Barus et al. 2023; Chasanah et al. 2021; Larashati et al. 2022; Roesma, Chornelia & Mursyid 2019).

The mahseers population in the wild habitat is decreasing due to overexploitation, habitat degradation, water pollution, and invasive species (Kottelat, Whitten & Kartikasari 1993). Among the four *Neolissochilus* inhabiting Indonesian freshwater, *N. thienemanni* has been under fully protected by the Indonesian Ministry of Marine Affairs and Fisheries Decree No.1/2021 and categorized as Vulnerable by IUCN (World Conservation Monitoring Centre 1996). Some efforts to manage and protect the mahseers in Indonesia have been conducted through local wisdom, establishing protected areas by the local community, and conducting different studies related to the bioecology of the mahseers (Larashati et al. 2020; Safitri, Sulistiono & Hariyadi 2021).

Taxonomical challenges in identifying the mahseers have been documented resulting in difficulties in identifying them morphologically (Khaironizam, Zakaria-Ismail & Armbruster 2015; Pinder et al. 2019; Roberts & Khaironizam 2008; Walton et al. 2017), while taxonomy certainty is needed for determining their conservation status and performing effective management and conservation efforts. Variation in the oral morphology exists within *N. soroides* in the form of normal, truncated, and lobe types, with the most commonly found being normal and truncated types (Khaironizam, Zakaria-Ismail & Armbruster 2015). Four *Neolissochilus sumatranus* specimens collected from West Sumatra exhibits incipien *Tor* type morphology and have been considered a junior synonym to *N. soroides* (Robert & Khaironizam 2008), which needs further validation. *Neolissochilus soro* has long been placed under the genus *Tor* and its taxonomic status remains in question (Scharpf 2015), while its reclassification to genus *Neolissochilus* needs further clarification of the diagnostic characters used to distinguish between *Neolissochilus* species inhabiting Indonesian waters.

To resolve the taxonomic confusion, genetic-based identification of the mahseers using marker genes from mitochondrial DNA such as COI, *Cyt b*, and 16S rRNA, and nuclear DNA such as microsatellite has been employed (Esa et al. 2011; Hoàng et al. 2015; Khudamrongsawat et al. 2021; Nguyen et al. 2008; Sati et al. 2013; Walton et al. 2017). In fish DNA barcoding research, using multiple molecular markers can offer additional insights, helping to clarify uncertainties and ensuring reliable identification of different species at various developmental stages (Qi et al. 2013; Wu et al. 2024; Zhang & Hanner 2012).

To date, genetic-based study on *Neolissochilus* inhabiting Indonesian freshwater is limited to *N. sumatranus* and *N. soro* (Barus et al. 2023; Larashati et al. 2022). Barcoding using the COI gene validated the morphological identification showing that mahseers captured in Bonan Dolok River belong to *Neolissochilus* and *Tor*. Genetic differentiation of *N. sumatranus* and *N. soro* was shown in the study by Barus et al. (2023) using RAPD. However, the resolution of RAPD in distinguishing closely related species might be lower than other techniques like DNA barcoding or sequencing methods (Cermakova et al. 2023).

Identifying species morphologically can be challenging due to similarities and variations among species or genera. Furthermore, there is a lack of clarity in the diagnostic characteristics used to reclassify *T. soro* as *N. soro*. Therefore, it is necessary to incorporate molecular identification alongside morphological identification of *N. soro* and *N. sumatranus*. The objective of our study was to utilize three molecular markers COI, *Cyt b,* and 16S rRNA for DNA barcoding to confirm morphological identification of *N. soro* and *N. sumatranus* collected from rivers in North Sumatra. By using these three molecular markers, we aim to enhance the accuracy and reliability of fish identification. Furthermore, no barcode sequences are currently available for Indonesian *Neolissochilus* generated from *Cyt b* and 16S rRNA. As a result, the barcode sequences developed in this study will contribute to the existing database of Indonesian *Neolissochilus*.

MATERIALS AND METHODS

SAMPLE COLLECTION

The sampling sites were chosen based on information from local fishermen, who identified areas where mahseers were typically found. Nineteen fish specimens were collected using a gill net with a 2 cm mesh size by local fishermen from the rivers in Samosir Regency-North Sumatra in 2016. Eleven and eight specimens from the Tulas and Boho Rivers were collected, respectively (Figure 1 & supplementary Table 1). All individuals were euthanized by rapid cooling in water mixed with ice (Wilson, Bunte & Carty 2009). Muscle tissues were cut from the specimen's right side of the body and preserved in 96% ethanol. Afterwards, the specimens were fixated in 4% formaldehyde for a night, washed and preserved in 70% ethanol and deposited at the Research Center for Limnology and Water Resources BRIN. The fishes were morphologically identified based on Weber and de Beaufort (1916), Kottelat, Whitten and Kartikasari (1993), and Kottelat (2013). Nine specimens of *N. sumatranus* collected from Bonan Dolok River were described in Larashati et al. (2020). No specific permit was required for this study because the species are not protected, sampling locations are outside the protected area, and no cross-border sample movement is involved.

DNA EXTRACTION, AMPLIFICATION AND SEQUENCING, AND DATA ANALYSIS

DNA was extracted from the muscle part of the fish samples collected in 2016 following the protocol from gSYNC™DNA Extraction Kit (Geneaid, Taiwan) with 100 µL elution buffer added and genomic DNA was stored in -20 °C for further use. The DNA extraction of 2019 collection was described in Larashati et al. (2022). The concentration and purity of extracted DNA were measured using NanoDrop™ One Microvolume UV-Vis Spectrophotometer, resulting in concentration of 12-200 ng/µL and purity of 1.8-2.

DNA samples from the 2016 collection were amplified and sequenced using primer pairs of COI, *Cyt b,* and 16s rRNA, while DNA samples from 2019 collection were amplified and sequenced using primer pairs of *Cyt b* and 16s rRNA. The COI segments of 2019 collection have been barcoded and explained in Larashati et al. (2022). 650 bp of the COI segment was amplified using FishF1 and FishR1 primers (Steinke & Hanner 2011; Ward et al. 2005). The *Cyt b* segment was amplified using LA-danio and HA-danio primers (Yang et al. 2010), producing a 1,140 bp amplicon. The 16S rRNA segment of 583 bp was amplified using16Sar L and 16Sbr H based on the work of Palumbi et al. (1991).

Amplification was conducted in 25 µL containing 10-100 ng DNA template, 0.1 µL Taq DNA polymerase (5 U/µL), 0.2 mM of dNTPs, 10 pM of each primer, and 2.5 µL PCR buffer (10x DreamTaq buffer, 20 mM $MgCl_2$) using PCR Thermal Cycler (Biorad) with PCR conditions following Steinke and Hanner (2011) for COI, Yang et al. (2010) for *Cyt b,* and Palumbi et al. (1991) for 16S rRNA. A reaction without DNA template was used as a negative control for each of the mtDNA segments amplification. The PCR products were visualized in 1% agarose gel containing GelRed (Biotium, US). The purification and bidirectional sequencing were conducted from the COI, *Cyt b*, and 16S rRNA segments with the same primers, performed by 1stBASE DNA Sequencing Services (Malaysia).

The raw amplified sequences were trimmed at both ends according to the forward and reverse primers to remove low quality regions and non-target sequences using Mega version X (Kumar et al. 2018). The assembled sequences were aligned using ClustalW in Mega version X. The nucleotide sequences were converted into amino acids, making sure to select the correct reading frame based on the primers used for amplification. The alignment was visually inspected for gaps or shifts in the reading frame, which were indicative of indels.

The aligned sequences were subjected to species identification using the Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The multiple sequences obtained from samples and GenBank-NCBI were aligned by ClustalW. A phylogenetic tree was reconstructed using Neighbour Joining analysis estimated by bootstrapping with 1000 replicate data sets for each of the mtDNA segments (COI, *Cyt b*, and 16s rRNA). A cyprinid *Barbodes binotatus* (accession number: MG699688.1 (COI), MT483247.1 (*Cyt b*), and MZ708839.1 (16S rRNA), was included in the analysis as an outgroup species. Reference sequences of *Neolissochilus* from the NCBI database were used as the ingroup (supplementary Table 2) together with the sequences obtained from this study for phylogenetic tree and genetic distance analysis. The pairwise genetic distance among sequences was calculated using the Kimura 2-parameter model implemented in MEGA version X.

Two methods for species delimitation using Assemble Species by Automatic Partitioning (ASAP) (https://bioinfo.mnhn.fr/abi/public/asap/) and Automated Barcode Gap Discovery (ABGD) (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html) were employed in this study. The aligned sequences for each mtDNA segment without the outgroup were used as an input file, the Kimura (K80) TS/TV was selected for distance mode, and other parameters were kept at default. These results were compared with the morphological identification and phylogenetic tree results.

RESULTS

MORPHOLOGICAL IDENTIFICATION

The specimens described in this study were those collected from Tulas and Boho Rivers. Specimens collected from the Tulas River coded I6, I8, I9, I11–13, I18, I19, I22 are identified as *N. sumatranus* and coded I10 and I15 belong to *N. soro*. All specimens from the Boho River (BH2, BH3, BH4, BH6, BH7, BH8, BH9, and BH10) are identified as *N. soro*.

In this study, both *N. sumatranus* and *N. soro* specimens displayed an elongated shape, with a standard length (SL) ranging from 109.4 mm to 170 mm and from 74.2 mm to 146 mm, respectively. All specimens of *N. sumatranus* and *N. soro* showed no median lobe of the lower lip. Four morphological characters are shown to differ between *N. sumatranus* and *N. soro*, which are body height, eye diameter/length, the length of pectoral to the dorsal fin, the length of anal fin to the caudal fin, and number of dorsal fin rays (Table 1 & Supplementary Figure 1).

Neolissochilus sumatranus from Tulas and Bonan Dolok Rivers exhibits several 'longer' or proportionally

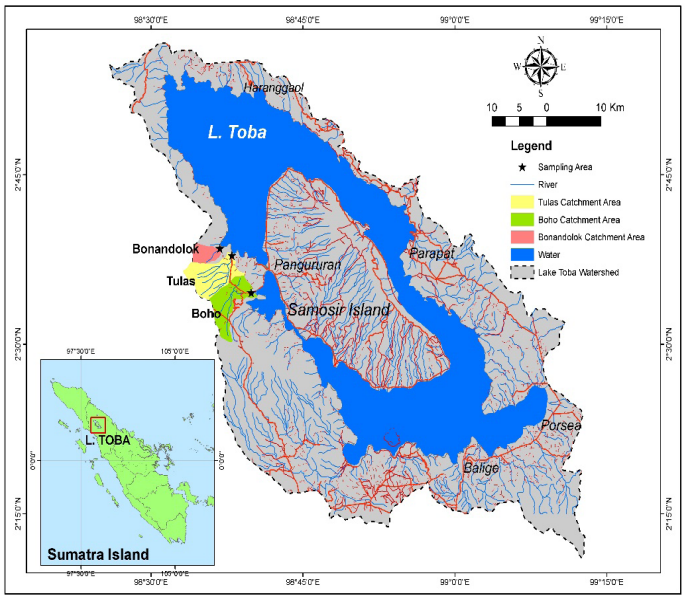


FIGURE 1. Sampling locations of *Neolissochilus* samples in this study

TABLE 1. Key characteristics used distinguishing *N. sumatranus* and *N. soro*

Species	Body depth	Eye diameter/length	Length of pectoral fin to the dorsal fin	Length of anal fin to the caudal fin	Number of dorsal fin rays
N. sumatranus	3.6–4.3 times shorter than the standard length	1.5 times shorter than the snout length, with four irregular rows of pores on each side of the snout under the eye	The pectoral fin tend to be almost the same or longer than the dorsal fin	The anal fin reaches the caudal fin	II, III, 8–11 weak dorsal rays
N. soro	3.6–3.9 times shorter than the standard length		The pectoral fin is shorter than the dorsal fin and tend to be further away from the pelvic fin	The anal fin does not reach the caudal fin	II, III, 7-9 weak dorsal rays

larger characteristics than *N. soro* from Tulas and Boho Rivers. These include the pectoral fin, anal fin, and the weak rays on the dorsal fin. Additionally, *N. sumatranus* has a greater body height proportion than *N. soro*. Notably, the presence of pores (tubercles) beneath the eyes in *N. sumatranus* is another distinguishing feature. The morphological characteristics of *N. sumatranus* from Bonan Dolok River were described in Larashati et al. (2022).

MOLECULAR DATA

Nineteen COI sequences, 28 *Cyt b* sequences, and 28 16S rRNA sequences were successfully amplified. No insertions, deletions, or stop codons were detected in the samples. The read lengths analyzed for the COI, *Cyt b,* and 16S rRNA were 602 bp, 1091 bp, and 565 bp, respectively. The analysis showed that all COI sequences, *Cyt b* sequences, and 16S rRNA sequences were identical within their respective datasets, with no observed variation across the aligned regions.

A similarity search in the Genebank database showed that sequences of *N. sumatranus* from Tulas River and *N. soro* are 100% similar to *N. sumatranus* (PQ237057.1) based on the COI segment. The reference *N. sumatranus* are from our previous study (Larashati et al. 2022). From the three rivers, *N. sumatranus* and *N. soro* samples are 99.74% similar to *N. soroides* (AP013114) based on the *Cyt b* segment, and 100% similar to *N. soroides* (OM203154.1) and *N. hendersoni* (OM202514.1) based on the 16S rRNA segment. The *Cyt b* and 16s rRNA sequences of morphologically identified *N. sumatranus* from Bonan Dolok and Tulas Rivers were submitted to Genbank-NCBI. The accession numbers for Cyt b sequences from these rivers are PQ623176-PQ623181. The accession numbers for 16S rRNA sequences from Bonan Dolok and Tulas Rivers are PQ620133-PQ620138. The COI sequences of samples collected from the Tulas River were not submitted to Genbank-NCBI due to their similarity with sequences of Bonan Dolok River samples. Only sequences from samples confidently identified as *N. sumatranus* were submitted, owing to uncertainty regarding samples identified as *N. soro*.

The phylogenetic tree analysis of the three segments showed a similar pattern, showing samples of *N. sumatranus* and *N. soro* are clustered together (bootstrap values 63-71%) with no genetic distance, suggesting that they belong to the same species and are closely related to *N. soroides* and *N. hendersoni* (Figure 2, Table 2-4) with genetic distance between the samples and *N. soroides/N. cf. soroides* ranged from 0.001 to 0.012, while the genetic distance between the samples and *N. hendersoni* ranged from 0.000 to 0.002. Based on the COI and *Cyt b* gene, the farthest genetic distance is to *N. hexastichus* and *N. hexagonolepis* (0.043–0.065), based on the 16S rRNA gene, it is to *N. hexastichus*, *N. hexagonolepis* and *N. pnar* (0.024).

The lower score from the ASAP method from COI resulted in 3 partitions (ASAP-score: 2.00; p-value: 1.33e-03; W: 1.64e-03; threshold dist.: 0.017618), from *Cyt b* resulted in 5 partitions (ASAP-score: 4.5; p-value: 3.00e-05; W: 7.59e-05; threshold dist.: 0.017585), from 16S rRNA resulted in 3 partitions (ASAP-score: 1; p-value: 2.06e-03; W: 1.76e-03; threshold dist.: 0.006224). The ABGD method from COI resulted in 8 partitions ranging from 2 to 9 groups, with partition 4 (prior maximal distance P=0.004642) indicating 4 groups, from *Cyt b* resulted in 7 partitions ranging from 2 to 9 groups, with partition 4 (prior maximal distance P=0.004642) indicating 6 groups, from 16S rRNA resulted in 5 partitions consisted of 3 and 4 groups, with partition 4 (prior maximal distance P=0.004642) indicating 3 groups. ASAP and ABGD methods recognized the *N. sumatranus* and *N. soro* samples, along with *N. soroides* and *N. hendersoni* references, as belonging to the same group or species.

DISCUSSION

Our study indicates that the samples morphologically identified as *N. sumatranus* and *N. soro* belong to a single lineage based on COI, *Cyt b,* and 16S rRNA markers. Both species delimitation analyses support this clustering, suggesting they represent the same species. However, although the sample size was small, we consistently observed distinct morphological differences between *N. sumatranus* and *N. soro* samples. Key characteristics that differentiate *N. sumatranus* include proportionally longer pectoral and anal fins, a higher number of weak rays on the dorsal fin, greater body height, and the presence of pores beneath the eyes.

The apparent contradiction between genetic homogeneity and morphological differences in our study aligns with broader taxonomic challenges observed in some freshwater fish species, as evidenced by different studies (Almeida et al. 2024; Hubert et al. 2008; Wu et al. 2024). Recent research on closely related mahseer species by Wu et al. (2024) has identified distinct species among genetically similar mahseers. They proposed that the species *T. laterivittatus* and *T. sinensis* should be considered synonymous. They also suggested that *N. compressus*, *N. stracheyi*, and *N. baoshanensis* may be the same species, differing only at the population level. The study also showed intraspecific variations in *N. qiaojiensis* and *N. hemispinus* samples. Furthermore, Chen et al. (2023) demonstrated that *N. soroides* and *N. hendersoni* from the Daying River in China show close genetic clustering, while the two species from Peninsular Malaysia are morphologically distinct (Khaironizam, Zakaria-Ismail & Armbruster 2015).

The discrepancies observed between genetic and morphological analyses may be due to recent divergence, suggesting that these species have not yet developed significant genetic differences in the analyzed mitochondrial genes (Wu et al. 2024). Alternatively, the

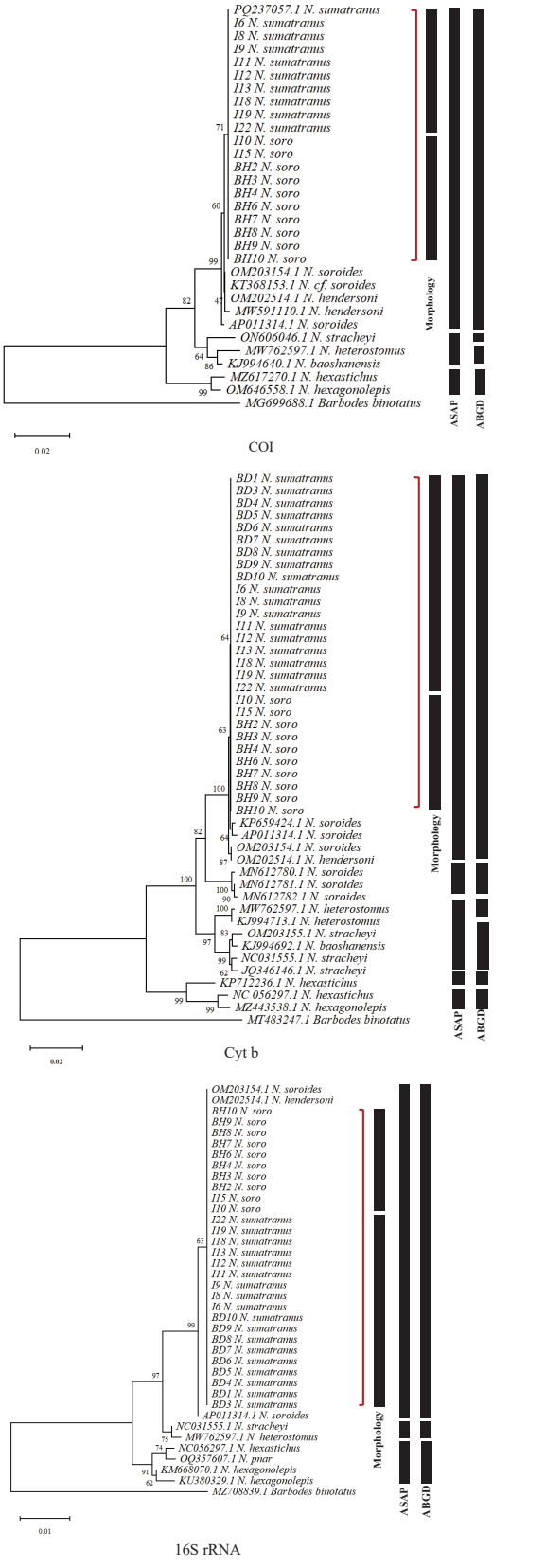


FIGURE 2. Neighbour Joining tree reconstructed from the COI, *Cyt b*, and 16S rRNA sequences for species of *Neolissochilus* with bootstrap value of 1,000 replicates. The red line indicates *Neolissochilus* examined in this study. Vertical black bars correspond to the morphological classification of the specimens, and the molecular units using ASAP and ABGD species delimitation methods

TABLE 2. Genetic distances between samples and Genbank database based on the COI sequences

No	Species	1	2	3	4	5	6	7	8	9	10
1	*Barbodes binotatus*										
2	*N. sumatranus*	0.168									
3	*N. soro*	0.168	0.000								
4	*N. soroides*	0.170	0.002	0.002							
5	*N. hendersoni*	0.171	0.002	0.002	0.001						
6	*N. stracheyi*	0.163	0.027	0.027	0.025	0.026					
7	*N. heterostomus*	0.179	0.029	0.029	0.027	0.028	0.020				
8	*N. baoshanensis*	0.183	0.022	0.022	0.020	0.021	0.017	0.010			
9	*N. hexastichus*	0.165	0.045	0.045	0.043	0.044	0.043	0.043	0.043		
10	*N. hexagonolepis*	0.168	0.043	0.043	0.041	0.042	0.045	0.045	0.041	0.008	

TABLE 3. Genetic distances between samples and Genbank database based on the *Cyt b* sequences

No	Species	1	2	3	4	5	6	7	8	9	10
1	*Barbodes binotatus*										
2	*N. sumatranus*	0.166									
3	*N. soro*	0.166	0.000								
4	*N. soroides*	0.168	0.012	0.012							
5	*N. hendersoni*	0.165	0.002	0.002	0.012						
6	*N. stracheyi*	0.171	0.030	0.030	0.030	0.031					
7	*N. heterostomus*	0.169	0.028	0.028	0.028	0.030	0.016				
8	*N. baoshanensis*	0.173	0.029	0.029	0.030	0.031	0.006	0.015			
9	*N. hexastichus*	0.164	0.062	0.062	0.064	0.062	0.064	0.058	0.065		
10	*N. hexagonolepis*	0.161	0.065	0.065	0.067	0.065	0.065	0.061	0.066	0.019	

TABLE 4. Genetic distances between samples and Genbank database based on the 16s rRNA sequences

No.	Species	1	2	3	4	5	6	7	8	9	10
1	*B. binotatus*										
2	*N. sumatranus*	0.079									
3	*N. soro*	0.079	0.000								
4	*N. soroides*	0.078	0.001	0.001							
5	*N. hendersoni*	0.079	0.000	0.000	0.001						
6	*N. stracheyi*	0.073	0.011	0.011	0.010	0.011					
7	*N. heterostomus*	0.073	0.013	0.013	0.012	0.013	0.002				
8	*N. pnar*	0.073	0.024	0.024	0.023	0.024	0.016	0.018			
9	*N. hexastichus*	0.073	0.024	0.024	0.023	0.024	0.016	0.018	0.004		
10	*N. hexagonolepis*	0.072	0.024	0.024	0.023	0.024	0.017	0.019	0.010	0.007	

genetic similarities might arise from incomplete lineage sorting, where ancestral polymorphisms are shared among the groups, or hybridization, which can occur when different species coexist in the same habitat, resulting in a mix of their genetic characteristics (Hubert et al. 2008).

While recent divergence, incomplete lineage sorting, and hybridization offer possible explanations, another factor to consider is the historical connectivity of river systems. The shared genetic background may reflect past gene flow between populations, with local adaptation driving morphological divergence. The dynamic nature of riverine systems supports this evolutionary context, as environmental changes can encourage genetic mixing and ecological specialization, leading to genetic similarity and morphological divergence.

While the genetic evidence supports combining *N. sumatranus* and *N. soro*, the morphological differences indicate that further investigation is needed to assess their taxonomic validity. The morphological differences may still reflect phenotypic plasticity or population-level variation, which complicates the determination of species boundaries. To address this uncertainty, an integrative approach that combines additional genetic markers, environmental analyses, and broader population sampling are recommended. Furthermore, the reliance on mitochondrial markers may confuse finer-scale genetic differences that can be detected using nuclear or whole-genome data. Future studies should incorporate high-resolution genomic analyses to confirm the presence or absence of gene flow and to identify potential genomic regions associated with morphological divergence.

Distinguishing species of Indonesian *Neolissochilus* is challenging due to inconsistent comparisons. Kottelat, Whitten and Kartikasari (1993) highlight key characteristics of Indonesian *Neolissochilus* differentiation such as body depth, the number of pre-dorsal scales, the number of pores on each side of the snout, and whether the groove behind the lower lip is interrupted. Kottelat (2013) reclassified *T. soro* as *N. soro*, but still referenced earlier classifications. Haryono and Tjakrawidjaja (2005) noted that *T. soro* specimens in BRIN lack a median lobe, while Walton et al. (2017) argued that those from the type location may have one (Supplementary Figure 2). More samples from varied locations are needed to confirm *N. soro*'s morphological traits and establish its diagnostic characteristics.

The analysis of the three markers showed a close relationship between our samples with the species *N. soroides* and *N. hendersoni*, except from the 16S rRNA result that our samples were found to be identical to *N. soroides* (OM203154.1) and *N. hendersoni* (OM202514.1) references. Although the 16S rRNA sequence is widely regarded as a universal marker, it may not provide enough resolution to differentiate closely related species due to its highly conserved regions (Jaafar et al. 2021; Mohanty et al. 2015; Xia et al. 2012). DNA barcoding of cultivable carp from India and Asiatic salamanders has shown that COI is more effective for accurate species identification (Mohanty et al. 2015; Xia et al. 2012).

The two species delimitation methods from COI, *Cyt b*, and 16S rRNA confirmed the clustering of *Neolissochilus* samples in this study with the references of *N. soroides* and *N. hendersoni*. However, we cannot conclude the data as there are no available specimens of *N. soroides* and *N. hendersoni* to compare. Morphologically, samples of *N. sumatranus* collected from West Sumatra bear a superficial resemblance to *N. soroides*, which is found in Peninsular Malaysia and Thailand (Khaironizam, Zakaria-Ismail & Armbruster 2015; Khudamrongsawat et al. 2021). Roberts and Khaironizam (2008) tentatively proposed that *N. sumatranus* is a junior synonym of *N. soroides*. We supported this conclusion based on findings from our earlier study (Larashati et al. 2022). However, it is premature to definitively state that *N. sumatranus* is a junior synonym of *N. soroides*, given the limited sample size of *N. sumatranus* in both our previous and current studies. Thus, it is essential to compare more samples of *N. sumatranus* from various regions of Sumatra with *N. soroides*, both morphologically and genetically, to determine whether these two species are indeed synonyms, especially considering the low average genetic distance observed between them.

Taxonomic ambiguity within the genus *Neolissochilus* may have arisen from phenotypic variations during the development of *Neolissochilus*. Such variations could be mistakenly interpreted as differences between species, adding to the complexity of the genus (Zhou et al. 2024). DNA barcoding using mitochondrial DNA has become a powerful tool for addressing the morphological limitations of complex genera that exhibit high morphological variation or share similarities among species. It aids in clarifying the taxonomy of certain mahseer species. However, DNA barcoding can only effectively identify and delineate species if accurate morphological identification is performed before submission to databases like GenBank or BOLD (Lalramliana et al. 2018). Misidentification may result in sequence inconsistencies, as seen in the COI sequences of *Neolissochilus* and *Tor* deposited in GenBank (Lalramliana et al. 2018; Walton et al. 2017).

Due to the limited information on the *Neolissochilus* genus in Indonesia, a comprehensive study incorporating traditional taxonomy and genetic analysis across their distribution range is necessary. This research would help clarify and establish the distinguishing characteristics within the genus, validate the taxonomic status, and update the conservation status. Additionally, ecological and behavioural studies are needed to investigate potential niche differentiation between *N. sumatranus* and *N. soro*. Taxonomic certainty is essential for effective fish management, ensuring accurate identification and understanding of fish species. This understanding supports effective conservation, regulation, and ecosystem management strategies. If *N. sumatranus* and *N. soro* are treated as a single species, conservation efforts could focus

on protecting habitats and maintaining genetic diversity across their combined range rather than managing them separately. In this case, preserving populations from various locations becomes critical, as they may possess unique genetic traits important for long-term adaptability.

CONCLUSION

DNA barcoding using three mitochondrial markers (COI, *Cyt b,* and 16S rRNA) showed that morphologically identified *N. sumatranus* and *N. soro* from North Sumatran rivers represent a single species, with close genetic relationships to *N. soroides* and *N. hendersoni* (genetic distances: 0.001–0.002). These findings challenge current taxonomic classifications and highlight the need to revise *Neolissochilus* taxonomy in Indonesia comprehensively. While DNA barcoding proved valuable in resolving species identities, our results emphasize the importance of integrating molecular and morphological approaches. We recommend: (1) extensive sampling across Indonesia; (2) standardized morphological analysis protocols; (3) application of multiple molecular markers, including the nuclear and whole genome approaches; and (4) development of clear diagnostic characters. This integrated approach will improve species identification accuracy, inform conservation strategies, and effectively support managing these economically important freshwater fishes.

ACKNOWLEDGEMENTS

This research was funded by the DIPA Limnologi LIPI 2016, Teknopark Samosir 2019 LIPI, Program Pendanaan Riset dan Inovasi untuk Indonesia Maju (RIIM) Lembaga Pengelola Dana Pendidikan (LPDP), Kementrian Keuangan Republik Indonesia number B-3838/II.7.5/FR.06.00/11/2023 and Rumah Program Organisasi Riset Hayati dan Lingkungan 2024 BRIN. The authors would like to thank those who have helped in the field. Special thanks to Daniel Lumbantobing for the discussion. Additionaly, thanks to Kunto Wibowo and Yudhi E. Sahari for their assistance in photographing the fish samples. Lastly, we appreciate the constructive comments from the anonymous reviewer (s) and editor (s).

REFERENCES

Ali, A., Dahanukar, A.N., Philip, S., Krishnakumar, K. & Raghavan, R. 2014. Distribution, threats and conservation status of the Wayanad Mahseer, *Neolissochilus wynaadensis* (Day, 1873) (Teleostei: Cyprinidae): An endemic large barb of the Western Ghats, India. *Journal of Threatened Taxa* 6(5): 5686-5699.

Almeida, R.B., Azambuja, M., Nogaroto, V., Oliveira, C., Roxo, F.F., Zawadzki, C.H. & Vicari, M.R. 2024. DNA barcode shows discordant cases among morphological and molecular species identification in *Isbrueckerichthys* (Siluriformes: Loricariidae). *Neotrop Ichthyol* 22(3): e240040. doi: 10. 1590/1982-0224-2024-0040

Barus, T.A., Wahyuningsih, H., Simanjuntak, B.N., Ginting, R.H., Batubara, A.S. & Hartanto, A. 2023. Genetic differentiation among Batak fish populations *(Neolissochilus sumatranus*, *Tor douronensis*, and *Tor soro*) in North Sumatra, Indonesia revealed by RAPD markers. *Biodiversitas* 24(3): 1327-1332. doi: 10.13057/biodiv/d240301

Cermakova, E., Lencova, S., Mukherjee, S., Horka, P., Vobruba, S., Demnerova, K. & Zdenkova, K. 2023. Identification of fish species and targeted genetic modifications based on DNA analysis: State of the art. *Foods* 12(1): 1-45. doi: 10.3390/foods12010228

Chasanah, E., Fithriani, D., Poernomo, A., Jeinie, M.H. & Huda, N. 2021. The nutritional profile of Indonesian salmon van Java mahseer *T. soro* species. *Potravinarstvo Slovak Journal of Food Sciences* 15: 566-574. doi:10.5219/1552

Chen, Y., Guo, Y., Wei, Z & Zhao, X. 2023. The complete mitochondrial genome of *Neolissochilus soroides* (Duncker, 1904) (Cypriniformes: Cyprinidae). *Mitochondrial DNA Part B* 8 (11): 1258-1262. doi: 10.1080/23802359.2023.2281032

Esa, Y.B., Siraj, S.S., Daud, S.K., Ryan, J.J.R., Rahim, K.A.A. & Tan, S.G. 2008. Molecular systematics of Mahseers (Cyprinidae) in Malaysia inferred from sequencing of a mitochondrial cytochrome C oxidase I (COI) gene. *Pertanika J. Trop. Agric. Sci.* 31(2): 263-269.

Everard, M., Pinder, A.C., Claussen, J.E. & Orr, S. 2021. Assessing the societal benefits of Mahseer (*Tor* spp.) fishes and habitats to strengthen the basis for their conservation. *Aquatic Conservation* 31(10): 2979-2986. doi:10.1002/aqc.3683

Froese, R. & Pauly, D. 2024. *FishBase.* World Wide Web electronic publication. www.fishbase.org (accessed on 1 June 2024).

Haryono & Tjakrawidjaja, A.H. 2006. Morphological study for identification improvement of Tambra fish (*Tor* spp.: Cyprinidae) from Indonesia. *Biodiversitas* 7(1): 59-62.

Hoàng, H.Đ., Phạm, H.M., Durand, J.D., Trần, N.T. & Phan, P.Đ. 2015. Mahseers genera *Tor* and *Neolissochilus* (Teleostei: Cyprinidae) from southern Vietnam. *Zootaxa* 3: 551-568. doi:10.11646/ zootaxa.4006.3.8

Hubert, N., Hanner, R., Holm, E., Mandrak, N.E., Taylor, E., Burridge, M., Watkinson, D., Dumont, P., Curry, A., Bentzen, P., Zhang, J., April, J. & Bernatchez, L. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS ONE* 3(6): e2490. doi: 10.1371/ journal.pone.0002490

Jaafar, F., Na-Nakorn, U., Srisapoome, P., Amornsakun, T., Duong, T.Y., Gonzales-Plasus, M.M., Hoang, D-H. & Parhar, I.S. 2021. A current update on the distribution, morphological features, and genetic identity of the southeast Asian mahseers, *Tor* species. *Biology* 10(4): 286. doi:10.3390/ biology10040286

Khai, N.X., Kusairi, M.N., Ahmad, S., Syahaneem, M.Z. & Fatimah, M.A. 2015. Market potential analysis for tengas (*Neolissochilus* sp.) in the Malaysian market. *International Food Research Journal* 22(4): 1429-1432.

Khaironizam, M.Z., Zakaria-Ismail, M. & Armbruster, J.W. 2015. Cyprinid fishes of the genus *Neolissochilus* in Peninsular Malaysia. *Zootaxa* 3962: 139-157. doi:10.11646/zootaxa.3962.1.7

Khudamrongsawat, J., Kettratad, J., Intasorn, P., Pinyo, N., Tapcheewin, S. & Wanusrut, P. 2021. Pattern of genetic structure of the common stream fish, *Neolissochilus soroides* (Pisces: Cyprinidae), addresses the importance of protected areas in eastern Thailand. *J. Fish. Biol*. 99(1): 175-185. doi: 10.1111/jfb.14709

Kottelat, M. 2013. 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. *The Raffles Bulletin of Zoology* Supplement 27: 1-663. http://zoobank.org/urn:lsid:zoobank.org:pub:0B66AE04-C644-43CD-9B76-043848FAA9FE

Kottelat, M., Whitten, A.J. & Kartikasari, S.N. 1993. *Freshwater Fishes of Western Indonesia and Sulawesi*. Singapore: Periplus Editions.

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura K. 2018. Mega X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35: 1547-1549. doi:10.1093/molbev/msy096

Lalramliana, Lalronunga, S., Kumar, S. & Singh, M. 2018. DNA barcoding reveal a new species of *Neolissochilus* Rainboth, 1985 from the Kaladan River of Mizoram, North East India. *Mitochondrial DNA Part A*. doi: 10.1080/24701394.2018.1450398

Larashati, S., Mayasari, N., Fitriana, Y.S., Wahyudewantoro, G., Widoretno, M.R. & Widiyanto, T. 2022. DNA barcoding of *Neolissochilus sumatranus* and *Tor douronensis* to support *in situ* conservation of Indonesian Mahseers. *Hayati Journal of Biosciences* 29(3): 370-377. doi:10.4308/hjb.29.3.370-377

Larashati, S., Sulastri, Ridwansyah, I., Afandi, A.Y. & Novianti, R. 2020. Conservation efforts of Ikan batak (*Tor* spp. and *Neolissochilus* spp.) and its prospects to support ecotourism in Samosir Regency, North Sumatra Indonesia. *IOP Conf. Series: Earth and Environmental Science* 535: 012041. doi:10.1088/1755- 1315/535/1/012041

Mohanty, M., Jayasankar, P., Sahoo, L. & Das, P. 2015. A comparative study of COI and 16S rRNA genes for DNA barcoding of cultivable carps in India. *Mitochondrial DNA* 26(1): 79-87.

Nasir, M., Munira, M. & Muchlisin, Z.A. 2018. Fish fauna in the Krueng Geumpang River, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* 139: 012023. doi:10.1088/1755-1315/139/1/012023

Nguyen, T.T.T., Na-Nakorn, U., Sukmanomon, S. & Chen, Z. 2008. A study on phylogeny and biogeography of mahseer species (Pisces: Cyprinidae) using sequences of three mitochondrial DNA gene regions. *Molecular Phylogenetics and Evolution* 48(3): 1223-1231. doi: 10.1016/j.ympev.2008.01.006

Palumbi, S.R., Martin, A.P., Romano, S., McMillan, W.O., Stice, L. & Grabowski, G. 1991. *The Simple Fool's Guide to PCR*, 2nd ed. Department of Zoology, University of Hawaii, Honolulu. pp. 1-45.

Pinder, A.C., Britton, J.R., Harrison, A.J., Nautiyal, P., Bower, S.D., Cooke, S.J., Lockett, S., Everard, M., Katwate, U., Ranjeet, K., Walton, S., Danylchuk, A.J., Dhanukar, N. & Raghavan, R. 2019. Mahseer (*Tor* spp.) fishes of the world: Status, challenges and opportunities for conservation. *Fish Biology and Fisheries* 29: 417-452. doi:10.1007/s11160-019-09566-y

Qi, P., Guo, B., Xie, C., Wu, C., Lu, S., Duan, Y. & Zhou, X. 2013. Assessing the genetic diversity and population structure of *Culter alburnus* in China based on mitochondrial 16S rRNA and COI gene sequences. *Biochemical Systematics and Ecology* 50: 390-396. doi: 10.1016/j.bse.2013.04.010

Rachmad, B., Sihombing, E. & Sabariyah, N. 2019. Pengelolaan ikan batak (*Neolissochilus thienemanni*, ahl 1933) di perairan umum daratan, Danau Toba, provinsi Sumatera Utara. *Jurnal Kelautan dan Perikanan Terapan* 2(2): 73-83.

Roberts, T.R. & Khaironizam, M.Z. 2008. Trophic polymorphism in the Malaysian fish *Neolissochilus soroides* and other old world barbs (Teleostei, Cyprinidae). *Fish. Nat. Hist. Bull. Siam Soc.* 56: 25-53.

Roesma, D.I., Chornelia, A. & Mursyid, A. 2019. Phenotype analysis of endemic mahseer fish (*Neolissochilus sumatranus*) from Batang Toru tributaries, North Sumatra, Indonesia. *J. Phys.: Conf. Ser.* 1317: 012099. doi:10.1088/1742-6596/1317/1/012099

Rumondang, A., Fuah, R.W. & Aidil Huda, M. 2023. Pemberdayaan masyarakat dalam pembuatan alat tangkap bubu yang ramah lingkungan untuk mendukung pemijahan ikan batak (*Neolissochilus thienemanni*) secara artificial spawning di Tapanuli Tengah. *Mattawang: Jurnal Penga,bdian Masyarakat* 4(3): 319-322. doi:10.35877/454RI.mattawang2096

Safitri, F.R., Sulistiono & Hariyadi, S. 2021. Aquatic physical and chemical characteristics of reservation and prohibited areas of mahseer (*Tor douronensis* Valenciennes, 1842) in Muara Bungo and Kerinci Regencies, Jambi Province of Indonesia. *E3S Web of Conferences* 322: 01010.

Sati, J., Sah, S., Pandey, H., Ali, S., Sahoo, P.K., Pande, V. & Barat, A. 2013. Phylogenetic relationship and molecular identification of five Indian Mahseer species using COI sequence. *Journal of Environmental Biology* 34: 933-939.

Scharpf, C. 2015. The authorship of *Neolissochilus soro* (Cypriniformes: Cyprinidae): A correction to Khaironizam et al. (2015). *Zootaxa* 3986(4): 499-500.

Steinke, D. & Hanner, R. 2011. The FISH-BOL collaborators' protocol. *Mitochondrial* 22: 10-14. doi: 10.3109/19401736.2010.536538

Walton, S.E., Gan, H.M., Raghavan, R., Pinder, A.C. & Ahmad, A. 2017. Disentangling the taxonomy of the mahseers (*Tor* spp.) of Malaysia: An integrated approach using morphology, genetics and historical records. *Reviews in Fisheries Science and Aquaculture* 25: 1-13. doi:10.1080/23308249.2016. 1251391

Ward, R.D., Zemlak, T., Innes, B., Last, P. & Hebert, P. 2005. DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. B. Biol. Sci.* 360: 1847-1857. doi: 10.1098/ rstb.2005.1716

Weber, M. & de Beaufort, L.F. 1916. *The Fishes of the Indo- Australian Archipelago III Ostariophysi: II Cyprinoidea (Synbranchi Leiden: Apodes)*. Leiden: E.J. Brill Ltd. pp. xv+455.

Wilson, J.M., Bunte, R.M. & Carty, A.J. 2009. Evaluation of rapid cooling and tricaine methanesulfonate (MS222) as methods of euthanasia in zebrafish (*Danio rerio*). *Journal of the American Association for Laboratory Animal Science* 48(6): 785-789.

World Conservation Monitoring Center. 1996. *Neolissochilus theinemanni*. *The IUCN Red List of Threatened Species 1996*: e.T14530A4442867

Wu, L., Qin, T., Hoang, H.D., Oo, T.N., Wang, X. & Chen, X. 2024. Molecular systematics and divergence of *Tor* and *Neolissochilus* fishes (Cypriniforms: Cyprinidae) from Southeast Asia and South China. *Journal of Zoological Systematics and Evolutionary Research* 2024: 5662867. doi: 10.1155/2024/5662867

Xia, Y., Gu, H.F., Peng, R., Chen, Q., Zheng, Y.C., Murphy, R.W. & Zeng, X.M. 2012. COI is better than 16S rRNA for DNA barcoding Asiatic salamanders (Amphibia: Caudata: Hynobiidae). *Mol. Ecol. Resour.* 12(1): 48-56. doi: 10.1111/j.1755-0998.2011.03055.x

Yang, L., Mayden, R.L., Sado, T., He, S., Saitoh, K. & Miya, M. 2010. Molecular phylogeny of the fishes traditionally referred to *Cyprinini sensu* stricto (Teleostei: Cypriniformes). *Zool. Scr.* 39(6): 527-550. doi:10.1111/j.1463-6409.2010. 00443.x

Zhang, J. & Hanner, R. 2012. Molecular approach to the identification of fish in the South China Sea. *PLoS ONE* 7(2): e30621. doi: 10.1371/journal. pone.0030621. 467

Zhou, C., He, J., Huang, H., Wang, H., Chu, Z., Zhao, B. & Guo, S. 2024. Phylogeny of *Neolissochilus* and studies on intergeneric kinship geography of Cyprinidae. *Hydrobiologia* 851: 4739-4759. doi:10.1007/s10750-024-05628-w

*Corresponding author; email: seka001@brin.go.id

SUPPLEMENTARY TABLE 1. Detailed sampling information where *Neolissochilus* specimens were obtained

Sampling data	Boho	Bonan Dolok	Tulas
Coordinate	98° 39' 31" E; 2° 35' 33" N	98° 40' 31" E; 2° 39' 37" N	98° 38' 58" E; 2° 37' 57" N
Sampling time	May, 2016	April, 2019	May, 2016
Sample size	8	9	11
Habitat characteristic:			
Riverbed	Boulder, gravel	boulder, gravel	sand, gravel
Land use at the upstream	Agriculture	forest	forest
Riparian zone	Settlement and road	forest	forest
Water quality parameters:			
Temperature (°C)	23.51	18.9-29.0	24.14
pH	7.68	6-8	7.18
Dissolved oxygen (mg/L)	8.8	7.20-8.61	8.9
Conductivity (mS/cm)	not measured	0.023-0.512	0.065
Specimen code	BH2-10	BD1, BD3-BD9	I6, I8, I9, I10, I11-13, I15, I18, I19, I22

The information on the coordinate, habitat characteristic, and water quality parameters were from Larashati and Ridwansyah (2017) and Larashati et al. (2020)

SUPPLEMENTARY TABLE 2. Source of published sequences of *Neolissochilus* species used for phylogenetic analysis and genetic distance for this study

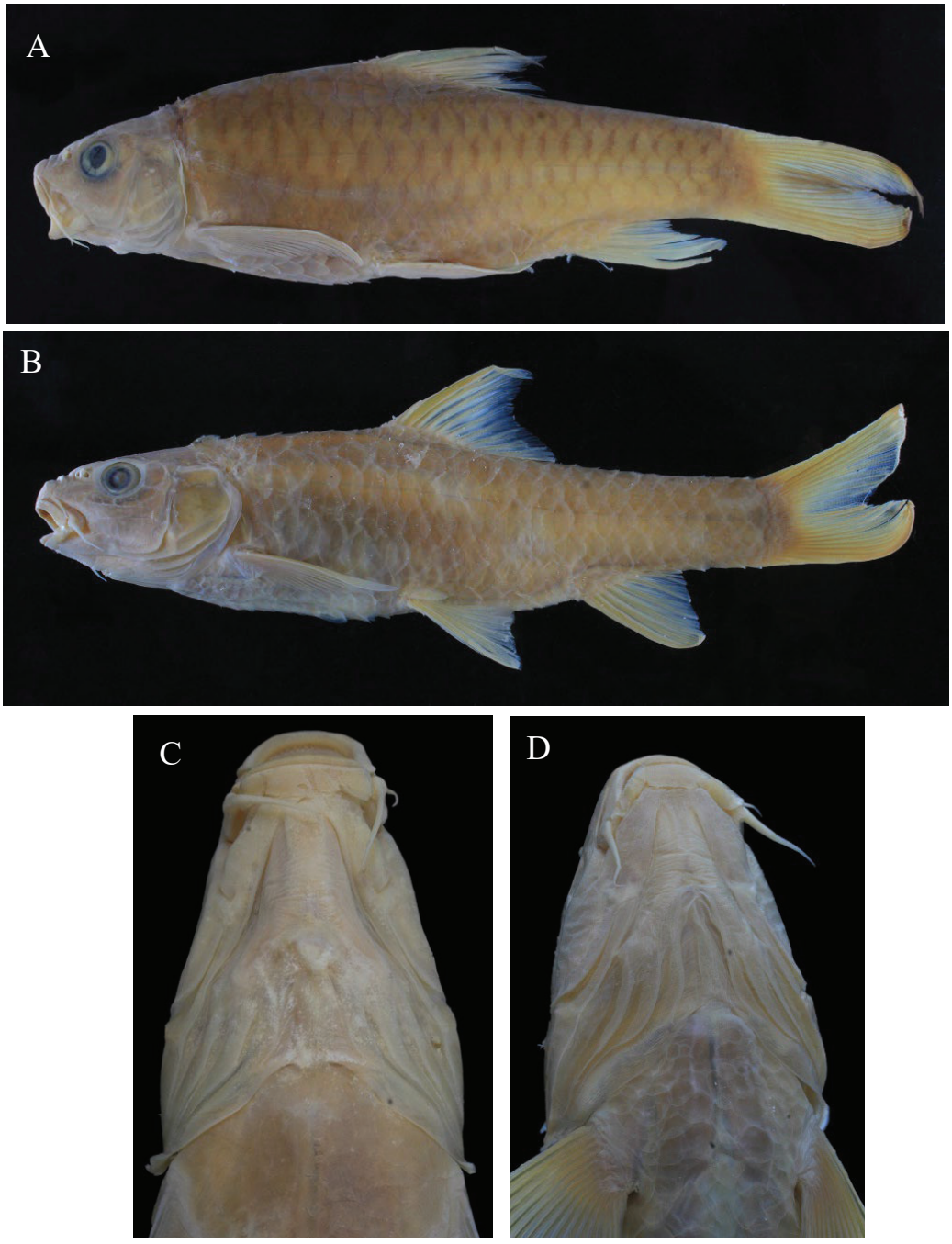
No.	Species	NCBI Acc No.	Sequence type and base position	Sample source
1	*N. sumatranus*	PQ237057.1	Partial/COI (bases 1 to 681)	Bonan Dolok River, Indonesia
2	*N. cf. soroides*	KT368153.1	Partial/COI (bases 1 to 654)	Terengganu, Malaysia
3	*N. hendersoni*	MW591110.1	Partial/COI (bases 1 to 655)	Pinang River, Malaysia
4	*N. stracheyi*	ON606046.1	Partial/COI (bases 1 to 701)	Dutah River, India
5	*N. hexastichus*	MZ617270.1	Partial/COI (bases 1 to 705)	Nagaland, India
6	*N. baoshanensis*	KJ994640.1	Partial/COI (bases 1 to 856)	China
7	*N. hexagonolepis*	OM646558.1	Partial/COI (bases 1 to 657)	India
8	*N. soroides*	KP659424.1	Partial/*Cyt b* (bases 1 to 1117)	unknown
9	*N. soroides*	MN612780.1–MN612782.1	Partial/*Cyt b* (bases 1 to 1117)	Thailand
10	*N. stracheyi*	JQ346146.1	Partial/*Cyt b* (bases 1 to 1118)	Laos
11	*N. heterostomus*	KJ994713.1	Partial/*Cyt b* (bases 1 to 1131)	China
12	*N. hexagonolepis*	MZ443538.1	Partial/*Cyt b* (bases 1 to 1140)	India
13	*N. baoshanensis*	KJ994692.1	Partial/*Cyt b* (based 1 to 1131)	China
14	*N. hexastichus*	KP712236.1	Partial/*Cyt b* (based 1 to 1141)	India
15	*N. pnar*	OQ357607.1	Partial/large subunit rRNA (bases 1 to 596)	India
16	*N. soroides*	AP011314.1	Complete mitochondrial genome/COI (bases 5479 to 7029), *Cyt b* (bases 14372 to 15512), 16s rRNA (bases 1098 to 2776)	unknown

continue to next page

continue from previous page

17	*N. soroides*	OM203154.1	Complete mitochondrial genome/COI (bases 5479 to 7029), *Cyt b* (bases 14372 to 15512), 16s rRNA (bases 1117 to 2752)	Daying River, China
18	*N. hendersoni*	OM202514.1	Complete mitochondrial genome/COI (bases 5479 to 7029), *Cyt b* (bases 14372 to 15512), 16s rRNA (bases 1117 to 2752)	Daying River, China
19	*N. stracheyi*	OM203155.1	Complete mitochondrial genome/COI (bases 5480 to 7030), *Cyt b* (bases 14373 to 15513), 16s rRNA (bases 1117 to 2753)	unknown
20	*N. stracheyi*	NC031555.1	Complete mitochondrial genome/COI (bases 5481 to 7031), *Cyt b* (bases 14374 to 15514), 16s rRNA (bases 1117 to 2753)	unknown
21	*N. heterostomus*	MW762597.1	Complete mitochondrial genome/COI (5480 to 7030), *Cyt b* (14373 to 15513), 16s rRNA (1118 to 2754)	unknown
22	*N. hexastichus*	NC056297.1/MN378521.1	Complete mitochondrial genome/COI (bases 5479 to 7029), *Cyt b* (bases 14373 to 15513), 16s rRNA (bases 1098 to 2776)	unknown
23	*N. hexagonolepis*	KM668070.1	Complete mitochondrial genome/COI (bases 5477 to 7027), *Cyt b* (bases 14370 to 15510), 16s rRNA (bases 1097 to 2773)	unknown
24	*N. hexagonolepis*	KU380329.1	Complete mitochondrial genome/COI (bases 5479 to 7029), *Cyt b* (bases 14372 to 15512), 16s rRNA (bases 1098 to 2776)	unknown

SUPPLEMENTARY FIGURE 1. Key characteristics in *N. sumatranus* distinguishing from *N. soro*. A: eye diameter, B: Length of pectoral fin, C: length of dorsal fin, D: Length of anal fin to the caudal fin. The specimen is *N. sumatranus* (BD5) collected from Bonan Dolok River



SUPPLEMENTARY FIGURE 2. Specimens of *N. soro* from Museum Zoologicum Bogoriense (MZB) collected from North Sumatra (A and C, MZB4055) and West Java (B and D, MZB1581) displaying truncated (C) and Tor-like (D) mouth type. Photo credit by Kunto Wibowo and Yudhi E. Sahari