

Article

Not peer-reviewed version

DNA Barcoding of Catfishes (Order: Siluriformes) with New Records of Two Species from Eastern Nepal

[Jash Hang Limbu](#) *

Posted Date: 20 August 2024

doi: 10.20944/preprints202408.1398.v1

Keywords: Fish diversity, taxonomy, new records, stream



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

DNA Barcoding of Catfishes (Order: Siluriformes) with New Records of Two Species from Eastern Nepal

Jash Hang Limbu ^{1,*}, Dipak Rajbanshi ¹, Laxman Khanal ^{2,*}, Wang Hui ¹,
Ram Chandra Adhikari ³, Sunita Thapa ⁴, Jin Quan Yang ^{1,5} and Chenhong Li ^{1,5}

¹ College of Fisheries and Life Science, Shanghai Ocean University, Shanghai, China,

² Central Department of Zoology, Institute of Science and Technology, Tribhuvan University, Kathmandu 44618, Nepal,

³ Degree Campus, Department of Zoology, Tribhuvan University, Biratnagar, Nepal,

⁴ Manimukanda College, Department of Biology, Tribhuvan University, Butwal, Nepal

⁵ Shanghai Universities Key Laboratory of Marine Animal Taxonomy and Evolution, Shanghai Ocean University, Shanghai, China

* Correspondence: limbujash@gmail.com

Abstract: The freshwater bodies of Nepal host a great deal of fish diversity; however, the molecular study of Nepalese fish is still in its early stages. Conservation management tactics require knowledge of the real species composition, often impeded by descriptions of intermittent species based on traditional taxonomy. Catfish, one of the most overfished taxa on the planet are abundant and heavily exploited in the eastern region of Nepal. Hence, actual catfish diversity in this region needs to be examined using a combined morphological and molecular DNA barcoding approach. In this study, we sequenced the cytochrome oxidase subunit I (COI) gene in Nepalese fish species for the first time and used it for molecular identification. Altogether, 99 specimens representing 29 species belonging to 21 genera and 8 families were successfully barcoded. They were found to be 98.5–100% identical to the known sequences deposited in the GenBank and BOLD databases. Of these, two species, *Amblyceps arunachalense* and *Erethistoides sicula* were recorded for the first time in Nepal. Out of 29 species, one species was found to be Endangered, one species as Vulnerable, two species were classified as Near Threatened, one species as Data Deficient, and 24 species were categorized as Least Concern. The intraspecific genetic distances within species ranged from 0 to 5%. The highest intraspecific genetic distances (Kimura 2 parameter) were found among the individuals of two species, namely *E. sicula* (4%) and *Glyptothorax telchitta* (5%), indicating the presence of putative (sub) species. The interspecific genetic distances among the catfishes in eastern Nepal ranged between 8.47% (between *Glyptothorax trilineatus* and *G. cavia*) and 26.38% (between *Amblyceps arunachalense* and *Chaca chaca*). The GC content at the 1st, 2nd, and 3rd positions of the codons was found to be 42.4%, 49.3% and 43.8%, respectively. These findings will support future initiatives for fish resource inventory, monitoring and management. It is also an important input to the DNA barcode library of catfish in the eastern region of Nepal.

Keywords: fish diversity; taxonomy; new records; stream

1. Introduction

Catfish (Order: Siluriformes) is one of the diverse groups of ray-finned fishes [1]. It constitutes forty families, 500 genera and about 4300 valid species [2] and the catfishes are dispersed in all the mainland [3,4] except Australia [5]. Of these, some are poisonous or venomous (e.g., *Heteropneustes fossilis* and *Plotosus lineatus*) which can cause severe wounds by injecting poison with their spines. However, fish of the order Siluriformes are popular game fishes, food items, and tropical aquarium fishes in many areas [1]. Most of the catfishes are freshwater fish except for the family Ariidae and Plotosidae which are also considered representatives in freshwater and brackish water [5]. Catfishes are identified by the absence of scales, the presence of spines on the dorsal and pectoral fins, the

fusion of the third and fourth neural crests to produce the complicated centrum, and the fusion of the parietal and supraoccipital gyri to create a hard structure for the neurocranium [6]. They generally carry tiny eyes and barbels reaching out on each side of the upper jaw [7]. Furthermore, the absence of symplectic, sub-opercular, basihyal, and intermuscular bones are other important characteristic features of catfish.

Fishes are aquatic animals of great diversity in morphological appearances, and there are around 36,640 nominal species throughout the world, including 18,614 valid freshwater species [2], which contribute significantly to the existing vertebrates [8]. Fish species delimitation is most prime for fisheries direction, however, sporadically morphology-based species identification is laborious for a non-professional to execute [9]. Authentic identification of the fish is one of the utmost important missions for ichthyologists to execute proper research and conservation management [9,10]. In addition, the number of experts is not sufficient for accurate and multiplex morphological delimitation [11]. Therefore, swift, authentic and duplicatable molecular tests are needed to identify species in many areas [12]. From the several hundred years, taxonomic descriptions were widely promised through morphometric and meristic characterizations [13], where species misidentification occurred due to their phenotypic and genotypic variations, cryptic or different ecological life stages [14].

DNA barcoding with the mitochondrial cytochrome c oxidase I gene (COI), has become a predominant approach for rapid and accurate identification for most of the fishes [14–16]. This technology has substantiated a high success rate of fish taxonomic delimitations with mean levels of COI diversity within fish species of approximately 0.3–0.4% [17]. Moreover, DNA barcoding not only does accurate species identification but also helps to identify the probable existence of morphologically cryptic species [14].

The eastern region of Nepal (ERN) has three ecological zones: the Terai, mid-hills and the mountains that extend between an altitude of 60 m to 8,848.86 m asl. The Koshi (Tamor, Arun, Tamakoshi, Dudkoshi, Bhotekoshi, Sunkoshi and Indrawati), Kankai (Deumai, Puwamai, Jogmai and Mai), and Mechi River are important water resources for ERN, which serve as the habitat for small-sized fishes (e.g. *Puntius phutunio*) to large-sized fishes (e.g. *Bagarius bagarius*). Studies on fish taxonomy and systematics of Nepal are in the preliminary stage. Several morphology-based ichthyofaunal surveys were done in ERN, which may have been prone to errors and might not have identified the cryptic species. For instance, Limbu et al. [18] misidentified several fish species collected from Phew Khola stream (*Schistura rupecula* as *S. sovana*; *Glyptothorax* unnamed as *G. trilineatus*; *Garra lissorhynchus* as *G. annandalei*, etc.). Similarly, until recently, Khatri et al. [19] misidentified *Tariqilabeo latius* as *T. macmahoni*, which is an endemic fish species of Dasht River, Pakistan. To date, studies have not focused on the identification of Nepalese fish using DNA barcoding. Therefore, there may be some cryptic fish species as well as some undiscovered species. This study aimed to identify catfish diversity in eastern Nepal by conducting extensive sampling and complementing morphological diagnosis with DNA barcoding technology. Here, for the first time, we used mitochondrial COI for fish species delimitation for Nepal, covering whole Siluriformes fishes collected from ERN.

2. Materials and Methods

2.1. Study Area and Sample Collection

This study collected samples in the Eastern Region of Nepal (ERN) from 45 sampling stations in the Koshi, Kankai, Lohandra, Mechi, Ratuwa, Singhiya and Tamor rivers (Figure 1). A total of around 700 specimens were collected from July 2019 to August 2023. Fishes were captured by cast net, gill net, and local fishing tools (Dhadiya, Ghorlang, and Mosquito net). The fish specimens were photographed in the field for the best living colour representation. For the molecular study, the muscle tissue or caudal fin of 2–6 individuals of each species (except for *Ailia coila* and *Mystus bleekeri*) were preserved in 95% ethanol in the field and then transferred to 75% ethanol. All specimens were assigned a collection number to facilitate sample tracking. Voucher specimens were deposited at the Museum of the Central Department of Zoology, Tribhuvan University, Kirtipur, Kathmandu,

Nepal. Voucher specimens were identified with the help of morphological measurements including total length (TL), standard length (SL), and meristic observations using Talwar and Jhingran [20] and Jayram [21]. We followed FishBase (www.fishbase.org) and Fricke et al. [2], for the recent valid name of species, genus and family. Subsequently, we checked the International Union for the Conservation of Nature (IUCN) (<https://www.iucnredlist.org/>) status of all the studied species.

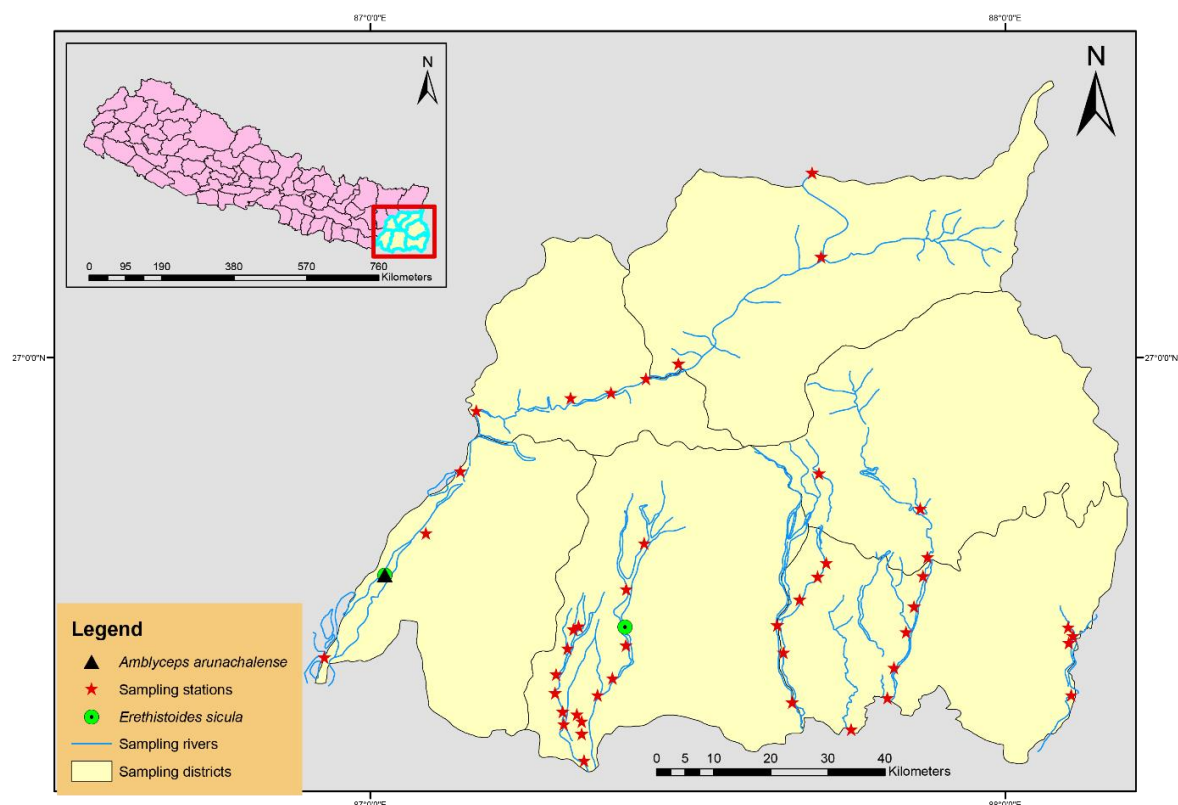


Figure 1. Map of the study area showing catfish sampling stations in water bodies of eastern Nepal.

2.2. DNA Extraction and PCR

Total genomic DNA was extracted from the muscle tissue or caudal fin using Tiangen Genomic DNA Purification Kit (Tiangen Biotech, Beijing, China) following the manufacturer's protocol. Partial sequences (~661bp) of cytochrome c oxidase subunit I (COI) were amplified with a forward primer, CGTGATTCTTCTCTACCAAYCAYAAAGA, and a reverse primer, ACTTCTGGGTGGCCGAAGAAYCARAA, from 2–6 individuals of each species using a volume of 20 µl, containing 10 µl of 2× TaqMaster Mix (Vazyme), 6 µl deionized distilled water, 2 µl DNA and 1 µl of each forward (F) and reverse (R) primers. The Polymerase Chain Reactions (PCR) were conducted for 35 cycles with the following parameters: initial denaturation step at 95 °C for 2 minutes, denaturation at 95 °C for 30 s, annealing at 55 °C for 45 s, elongation at 72 °C for 45 s and final elongation at 72 °C for 5 minutes. The amplified products were checked on 1% agarose gel and the successful amplicons were sequenced in both directions with the same pair of primers as for PCR amplification using a BigDye Terminator Cycle Kit v.3.1 (Invitrogen) on an ABI 3730XL sequencer (Applied Biosystems).

2.3. Data Analysis

All the resulting sequences were assembled using 'Geneious Prime 9.0.2' software and aligned and analyzed using MEGA 11 [22]. The noisy sequences of both ends were trimmed before subsequent sequence analysis. The 99 sequences obtained and analyzed in the present study were submitted to the NCBI GenBank database and the assigned accession numbers are given in Table 1. The maximum likelihood phylogenetic trees of the newly recorded species, namely *Amblyceps*

arunachalense and *Erethistoides sicula* were constructed using homologous sequences from their respective congeners through IQ-Tree v1.6.12 [23], with HKY+G evolutionary model and 1000 bootstrap replication. For the entire data set of 99 sequences, phylogenetic analysis was performed using maximum likelihood (ML) methods through IQ-TREE v1.6.12 [23] using the 1000 bootstrap replications implemented in Phylosuite 1.2.3 [24]. We used the evolutionary model TIM2+F+I+G4 as the best-fit model selected by Model Finder [25], applying the Bayesian information criterion. The resulting trees were visualized in FigTree v. 1.4.4. Through MEGA 11 software [22], we calculated sequence nucleotide composition, and the mean genetic distances of various taxonomic levels, within and between species, genera and families, using K2P model with 1000 bootstrap replicates. To detect whether the DNA barcode gaps existed, we calculated intra-specific and inter-specific K2P and p-distances including minimum, maximum and mean values. The pairwise genetic distances among the 29 species were calculated using K2P parameter and 500 bootstrap replicates in MEGA 11. The matrix of pairwise genetic distances was used to visualize in a heatmap using ‘heatmaply’ package [26] in R statistical software [27].

Table 1. Fish species collected from ERN with their IUCN status and GenBank accession numbers.

S.N.	Family	Species	IUCN, 2022	GenBank Accession No.
1	Amblycipitidae	<i>Amblyceps mangois</i> (Hamilton 1822)	LC	OR793874–OR793877
2		* <i>Amblyceps arunachalense</i> Nath & Dey 1989	EN	OR801071–OR801073
3	Ailiidae	<i>Ailia coila</i> (Hamilton 1822)	NT	OR793880
4		<i>Clupisoma garua</i> (Hamilton 1822)	LC	OR726087, OR799676
5		<i>Eutropiichthys vacha</i> (Hamilton 1822)	LC	OR733549, OR799677
6		<i>Eutropiichthys murius</i> (Hamilton 1822)	LC	OR794001–OR794003,
7	Bagridae	<i>Batasio batasio</i> (Hamilton 1822)	LC	OR793868–OR793871
8		<i>Mystus bleekeri</i> (Day 1877)	LC	OR726099
9		<i>Mystus vittatus</i> (Bloch 1794)	LC	OR801519–OR801522
10		<i>Mystus cavasius</i> (Hamilton 1822)	LC	OR793864–OR793867
11		<i>Rita rita</i> (Hamilton 1822)	LC	OR801529–OR801531
12		<i>Sperata aor</i> (Hamilton 1822)	LC	OR801372–OR801374
13		<i>Sperata seenghala</i> (Sykes 1839)	LC	OR801376–OR801381
14		<i>Olyra longicaudata</i> McClelland 1842	LC	OR801375, OR801515
15	Siluridae	<i>Ompok bimaculatus</i> (Bloch 1794)	NT	OR801463–OR801468
16		<i>Bagarius bagarius</i> (Hamilton 1822)	VU	OR801523–OR801526
17		<i>Erethistes pusillus</i> Müller & Troschel 1849	LC	OR801436–OR801437
18		* <i>Erethistoides sicula</i> Ng 2005	DD	OR801443–OR801454
19		<i>Gagata cenia</i> (Hamilton 1822)	LC	OR801457–OR801459
20		<i>Gogangra viridescens</i> (Hamilton 1822)	LC	OR801461–OR801462
21		<i>Nangra assamensis</i> Sen & Biswas 1994	LC	OR801527, OR801553
22		<i>Glyptothorax telchitta</i> (Hamilton 1822)	LC	OR801537–OR801541
23		<i>Glyptothorax cavia</i> (Hamilton 1822)	LC	OR801516–OR801517
24		<i>Glyptothorax trilineatus</i> Blyth 1860	LC	OR801440–OR801441
25		<i>Glyptothorax indicus</i> Talwar 1991	LC	OR801455–OR801456
26		<i>Pseudecheneis sulcata</i> (McClelland 1842)	LC	OR801199–OR801203
27	Heteropneustidae	<i>Heteropneustes fossilis</i> (Bloch 1794)	LC	OR801151–OR801155
28		<i>Pachypterus atherinoides</i> (Bloch 1794)	LC	OR801194–OR801198
29	Chacidae	<i>Chaca chaca</i> (Hamilton 1822)	LC	OR801438–OR801439

*= New records for Nepal.

3. Results

3.1. Identification of Catfishes in ERN

The taxonomic delimitation and nomenclature of the collected fish species by morphological and molecular categorization affirmed that all 99 sequences belonged to 29 species, 21 genera and 8 families. Of these, two species viz: *Amblyceps arunachalense*, and *Erethistoides sicula* were recorded for the first time from Nepal (Figures 3 and 5). Based on the International Union for the Conservation of Nature (IUCN 2022), 1 species was found to be Endangered (EN), 2 species under Near Threatened (NT), 1 species under Vulnerable (VU), 1 species under Data Deficient (DD) and 24 species under Least Concern (LC) (Table 1).

3.2. Additional Ichthyofauna for Nepal

Amblycipitidae

Amblyceps arunachalense Nath & Dey 1989

Material examined

KR130, 130b, 130c, 130d, 130e; 60.81–95.5mm SL; Koshi River; Nepal 50 km from Biratnagar Sub-metropolitan City, 26.65861°N and 87.02277°E, 96 masl; Limbu and Rajbanshi, 12 January 2021.

Diagnosis. *Amblyceps arunachalense* has a slim and posteriorly compressed body (Figure 3A). The head is short and nearly rounded. The mouth is terminal having double folded lips with lower jaw longer than the upper jaw. Teeth in the lower jaw have a bowed band which is divided in the central, whereas teeth in the upper jaw have a thin band with hind extensions on both edges isolated by an extremely narrow aperture. Eyes are small and subcutaneous. It has a somewhat convex dorsal body contour, while the profile from the pectoral fin to the base of the pelvic fin is convex, the ventral body profile from the pelvic fin to the base of the anal fin is slightly concave. There are four pairs of barbels which are compressed throughout their length. Nasal and inner mandibular barbels are nearly equal to the head length. The maxillary barbel extends just beyond the central part of the pectoral fin base whereas the outer mandibular barbel extends just beyond the posterior end of the pectoral fin base. The dorsal fin is vertically located almost mid-part of the pectoral and pelvic fin, containing I, 7 rays. The spine is weak and concealed in the small bulging skin. The adipose fin is located at the vertical level of the anal fin origin, with its posterior part rounded and far from the caudal fin. Pectoral fin has I, 7 rays, a comparatively smooth and stronger spine than dorsal spine. It extends at the back of dorsal fin base. Pelvic fin contains I, 5 rays, while the anal fin with V, 8 rays. The caudal fin is deeply forked, having the upper lobe longer than the lower lobe bearing 6 + 9 rays. Vertebrae are 40. The lateral line is absent.

Coloration. Its live coloration is darkish brown. After preservation in 10% formaldehyde solution, its color changes into grey on lateral side; white creamy on the ventral side.

Distribution. Sunsari and Saptari districts (Commonly available in Koshi River).

Remarks. Collected from slowly moving water near Koshi Barrage.

Economic importance. This fish is a locally famous food fish.

Phylogenetic relationship. For the phylogenetic analysis, fifteen *Amblyceps* cytochrome c oxidase subunit I (COI) sequences were employed (Figure 2). For the phylogenetic tree reconstruction, homologous sequence of *Danio rerio* (Hamilton, 1822) was used as the out-group. With a strong bootstrap value (>95%), the phylogenetic tree revealed that the recently discovered Nepalese *Amblyceps* species is congruent with an Indian *Amblyceps* species known as *A. arunachalense* (Figure 3). Nepalese *A. arunachalense* had a genetic distance of 8.8% with *A. mangois*, 1.7% with Indian *A. arunachalense* (retrieved from NCBI), 13.1% with *A. apangi*, 10.5% with *A. laticeps*, 10.1% with *A. hmolaii*, 11.8% with *A. mucronatum*, and 11.1% with *A. serratum*. The overall mean genetic distance of tested *Amblyceps* species was found to be 0.11±0.01.

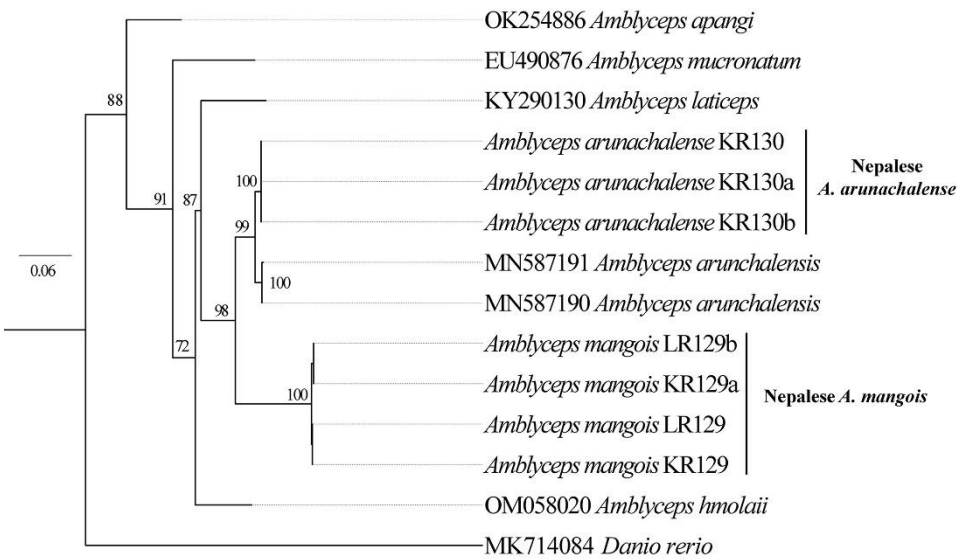


Figure 2. Maximum-likelihood (ML) tree of newly recorded *A. arunachalense* and other species of the Amblycipitidae, based on COI gene sequences covering the partial barcoding region (661 bp).



Figure 3. (A) *Amblyceps arunachalense*, KR-130 captured from the Koshi River in eastern Nepal (TL = 73.6 mm and SL= 66.8 mm); (B) *A. mangois* LR-129 captured from the Lohandra River in eastern Nepal (TL = 98.5 mm and SL=88.7 mm).

Sisoridae

Erethistoides sicula Ng 2005

Material examined

KR120, 120a, 120b; 27–40.5mm SL; Koshi River; Nepal, 50 km from Biratnagar Sub-metropolitan City, 26.65861°N and 87.02277°E, 96 masl; Limbu and Rajbanshi, 12 January 2021. LR120, 120b; 32–47 mm SL; Lohandra River, Nepal, 6 km from Biratnagar Sub-metropolitan City, 26.54611°N and 86.9383°E, 70 masl; Limbu and Rajbanshi, 12 January 2021.

Diagnosis. *Erethistoides sicula* has a broad and strongly depressed head (Figure 5). The body is broad and compressed with more compression in the caudal peduncle. Its dorsal profile is slightly

convex while the ventral profile is nearly straight. Eyes are ovoid having the longest axis on the horizontal side. Gill openings are narrow. The mouth is small with the upper jaw projected on the far side of the lower jaw with papillated lips. Four pairs of barbels are present with short and slim nasal barbels which extend to the posterior margin of posterior nares. The Maxillary barbel is slim, extending to the base of the pectoral fin base, whereas the outer mandibular barbel extends on the far side of the base of the posterior-most pectoral fin ray. Similarly, the inner mandibular barbell is short and can nearly reach the base of the pectoral spine. On the flanks, along with skin, numerous elongated and flattened plaque-like tubercles are present in the longitudinal rows. The lateral line is complete and located in the mid-lateral position. Vertebrae 29–31. All the fin numbers were found to be the same as described by the original author. Dorsal fin with 5(12) rays and its margin is straight. Furthermore, it has robust, straight and compressed spine, having smooth spine on its anterior region and posterior region with 24 small serrations. Pectoral fin with 5(5) rays, anal fin with 10–11 rays, and caudal fin forked having both lobes’ tips pointed, with the lower lobe slightly longer than the upper lobe.

Coloration. The dorsolateral side of the body and head is light chocolate brown as described by [35]. Both paired and unpaired fins are hyaline. The dorsal fin has transverse faint brown bands in the direction of the base and the center portion of the dorsal fin. The dark brown color is often present in the pelvic, pectoral and anal fins. All barbels are creamy white but sometimes maxillary barbel is composed of brown rings.

Distribution. Known from the Koshi River which is the largest river of Nepal, and also known from the Lohandra River of Morang district, six km away from Biratnagar Sub-metropolitan City.

Remarks. *E. sicula* was collected from Koshi and Lohandra rivers. This can be found in fast-flowing water with gravelly bottom (Figure 1).

Economic importance. This fish is locally famous for food.

Phylogenetic analysis. For the phylogenetic analysis of *E. sicula*, we employed five sequences of newly generated and five from NCBI public database (Figure 4). There was not any mitochondrial cytochrome c oxidase subunit I (COI) sequence for other *Erethistoides* spp. in the GenBank. So, we used other closely related species for phylogenetic tree construction. Newly generated sequences were tested using the BLAST tool in both NCBI GenBank and BOLD databases, which showed 100% match with *E. sicula* from other countries. Further, the K2P genetic distance of Nepalese *E. sicula* and retrieved data from GenBank was found to be 0%, indicating the same species.

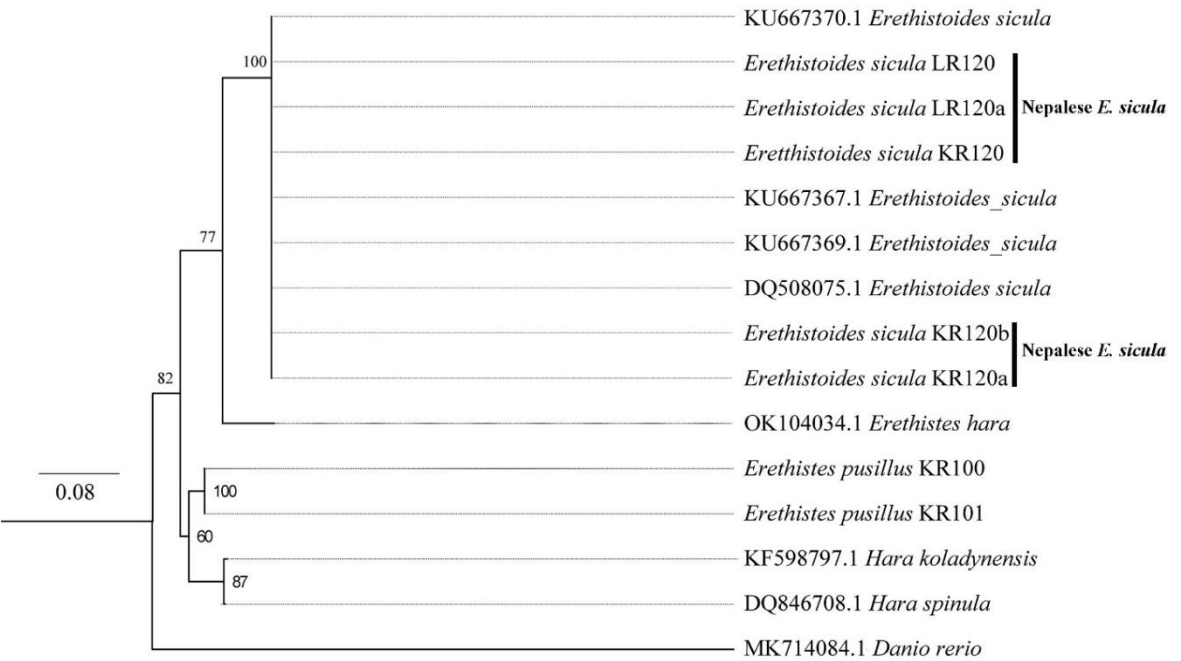


Figure 4. Maximum-likelihood (ML) tree of newly recorded *E. sicula* and other species of the Sisoridae, based on COI gene sequences covering the partial barcoding region (~661 bp).



Figure 5. *Erethistoides sicula*, KR120 captured from the Koshi River in eastern Nepal. The TL and SL of the specimen were 42.3 mm and 35.8 mm, respectively.

3.3. DNA Barcoding of Catfishes from Eastern Nepal

A subsample of 2–6 specimens for each species representing all the sampling localities from where they were recorded was barcoded (except *A. coila* and *M. bleekeri*). Due to the limited availability of tissue samples, we used a single specimen from each species. We created barcodes for 99 specimens belonging to 29 species, representing 21 genera and eight families from 45 locations in the eastern region of Nepal (ERN) (Figure 1; Table 1). The sequences were aligned to NCBI and BOLD public databases using the BLAST tool, which displayed a similarity of 98.5–100% and matched with the obtained morphological data, except for one species, *Glyptothorax indicus*. The sequences generated in this study have been deposited in the GenBank database (Table 1).

In the analysis of nucleotide base composition for an alignment of 99 sequences containing 648 sites, 318 (49.07%) were invariable sites. Variable sites were 247 (38.11%) which included 15 (2.31%) singleton sites and 232 (35.8%) parsimony informative sites. The average nucleotide composition showed that the composition of T was the highest (29.0%), while that of G was the lowest (18.0%). The base composition of C and A were 27.2% and 25.8% respectively. It was also found that AT content (54.8%) was higher than that of GC content (45.2%). GC content at 1st, 2nd, and 3rd codon positions was 42.4%, 49.3% and 43.8%, respectively.

3.4. Phylogenetic Relationship among Catfishes in Eastern Nepal

The maximum likelihood (ML) (Figure 6) tree was constructed for the 99 sequences belonging to 29 species, 21 genera and 8 families. The family Sisoridae with 11 species, Bagridae with 9 species, Ailiidae with 5 species, Amblycipitidae with 2 species, and Siluridae, Horabagridae, Heteropneustidae and Chacidae with single species each. In the phylogenetic tree, COI barcoding region sequences discriminated all the species and clustered the similar species with significant bootstrap values under the same nodes (Figure 6). The Sisoridae family formed a paraphyletic clade in which *Pseudechenies sulcata* formed a separate group from the rest of the species in the family. The species under the family Bagridae also formed a paraphyletic clade, *Olyra longicaudata* and *Batasio batasio* remained distinct from the clade of the rest of the species *Batasio batasio* distantly separated from the other members of the Bagridae family. The members of the family Ailiidae also clustered into one clade with strong bootstrap support. Two species of the *Amblyceps* genus, namely *A. arunachalense* and *A. mangois* formed a distinct monophyletic clade. The species belonging to the genus *Glyptothorax*, namely *G. indicus*, *G. cavia*, *G. trilineatus*, and *G. telchitta* formed a strong

monophyletic clade. Similarly, the species belonging to the genus *Mystus* (*M. vittatus*, *M. bleekeri* and *M. cavasius*) formed a strong monophyletic clade. Similarly, sequences of the genus *Sperata* (*S. aor* and *S. seenghala*) also formed a strong monophyletic clade. The fish belonging to the genera *Ompok*, *Batasio*, *Rita* and *Chaca* formed a paraphyletic clade.

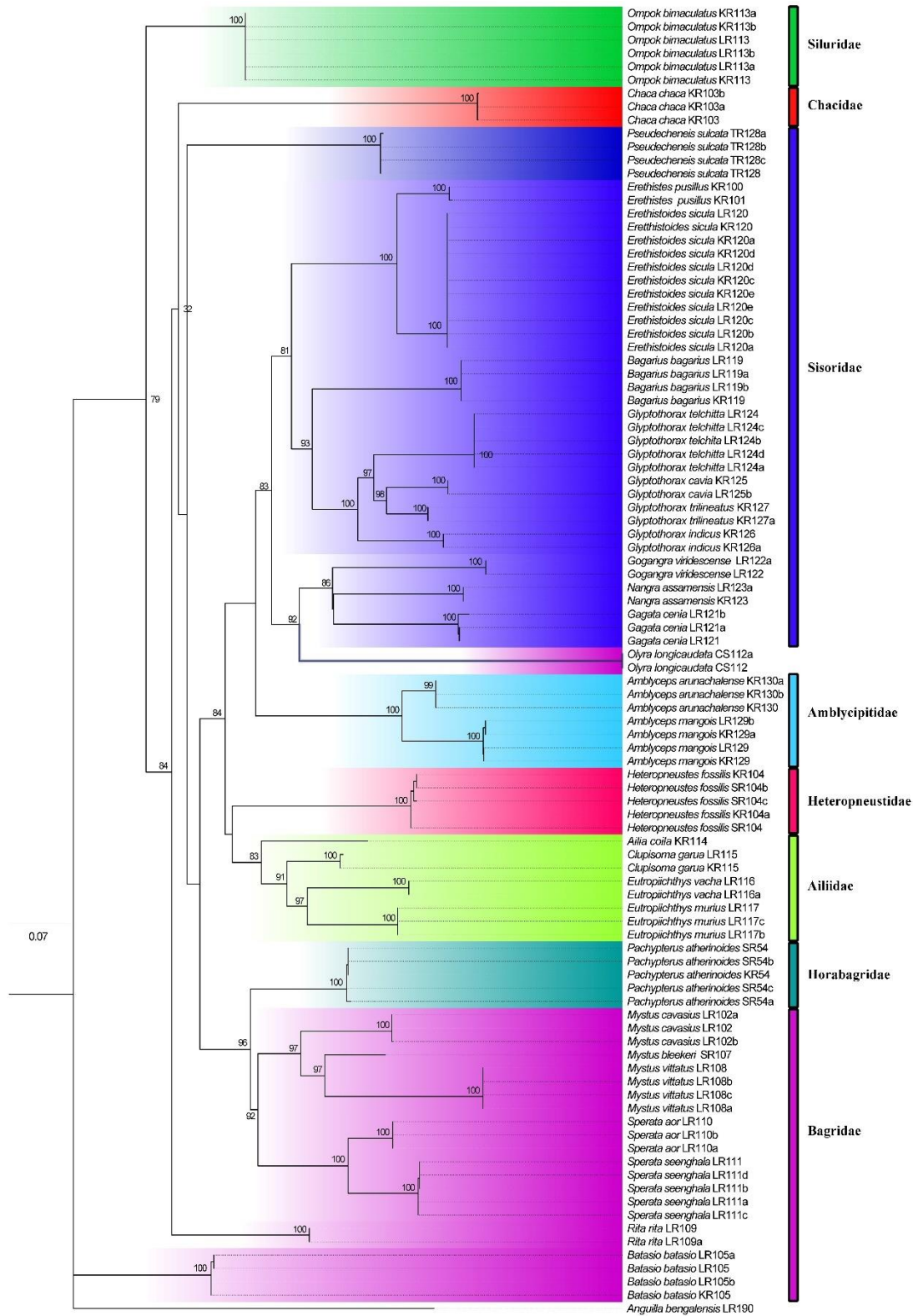


Figure 6. ML tree showing the phylogenetic relationship among the catfishes from eastern Nepal using the COI sequences.

3.5. Genetic Divergence among Catfishes of Eastern Nepal

The K2P and p-distances were computed at various taxonomic levels, including species, genus, and family (Table 2). The mean genetic distances increased hierarchically from within species to within the family. The overall mean distances of K2P and p-distance for the COI sequences were 22 ± 0.02 (mean $\pm$ SE) and 17 ± 0.01 (mean $\pm$ SE). The intraspecific K2P genetic distances ranged from 0 to 5% (mean = 0.52, standard error [SE] = 0.0023) at the species level, whereas the intraspecific p-distances ranged from 0 to 4% (mean = 0.45, standard error [SE] = 0.0011). The highest intraspecific K2P and p-distances were found among the two species, namely *E. sicula* (4%) and *G. telchita* (5%), indicating the presence of putative species/subspecies. The interspecific K2P distances ranged from 0 to 17% (mean = 2.6%, standard error [SE] = 0.0098) at the genus level, while the interspecific p-distances ranged from 0 to 15% (mean = 2.14%, standard error [SE] = 0.0073). The pairwise interspecific genetic distances among the studied 29 species of catfishes ranged between 8.47% (between *Glyptothorax trilineatus* and *G. cavia*) and 26.38% (between *Amblyceps arunachalense* and *Chaca chaca*) (Figure 7).

Table 2. Mean K2P and p-distances for within species, genus and family of fish collected from ERN.

	Taxonomic level	Min (%)	Mean (%)	Max (%)	SE
K2P distance					
	Within species	0.00	0.52	5.0	0.0023
	Within genus	0.00	2.6	17.0	0.0098
	Within family	0.00	8	21.0	0.0031
P-distance					
	Within species	0.00	0.45	4.0	0.0011
	Within genus	0.00	2.12	14.0	0.0073
	Within family	0.00	7.25	15.0	0.0014

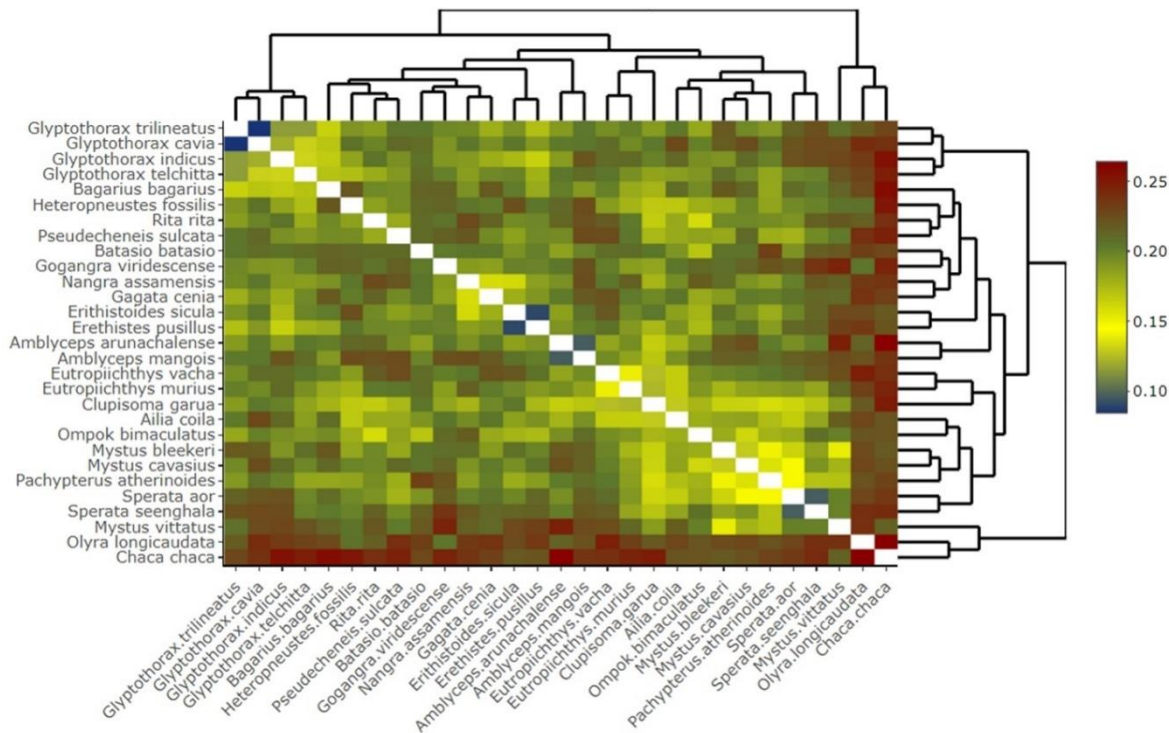


Figure 7. Heatmap showing the pairwise genetic distance among the catfish species in eastern Nepal based on barcoding region of the COI gene sequences.

4. Discussion

Nepal, uniquely placed in the Himalayan region is known for its rich freshwater resources that support diverse flora and fauna. Despite covering relatively a small surface area, Nepal hosts a great deal of freshwater fish diversity. The ichthyofaunal diversity of the country is poorly documented and lacks their comprehensive taxonomic identification. To date, Nepalese ichthyologists have been using morphology-based identification, which is not sufficient for accurate and multiplex morphological identification. Therefore, there might be a larger number of species that are yet to be discovered. For the first time, we have used the COI gene for species identification of catfishes from eastern Nepal. We also recorded the occurrence of two species of catfishes, namely *Amblyceps arunachalense* and *Erithistoides sicula* which were not previously reported from Nepal. Nearly for two decades, *A. arunachalense* was treated as a synonymy of a *A. mangois* [29]. However, Vishwanath and Linthoingambi [30] redescribed and resurrected it as a valid separate species. Ng [29] described a new species named *E. sicula* from West Bengal, India. Almost 18 years after its original description, we recorded this species from the southeastern region of Nepal. It's not questionable because the southeastern part of Nepal is proximal to the West Bengal of India and southward-flowing rivers from these regions join to form big rivers before pouring into the Bay of Bengal. There are numerous rivers including Nepal's biggest Koshi River, which flow downward through West Bengal. Such connectivity of water bodies might have facilitated the dispersal of *E. sicula* into the water bodies of Nepal. Both morphology and COI gene confirmed that these two newly recorded species exist in Nepal. Moreover, we have three *Amblyceps* and two *Erithistoides* undescribed putative species. Due to the lack of ethanol-preserved tissue, we could not include those species in this study.

The fish species collected from ERN were successfully identified with 98.5–100% identity both in NCBI GenBank and BOLD databases. However, one species of *Glyptothorax* was unable to be identified through both databases. Morphologically, we identified this species as a *G. indicus*, but the results from both databases suggested it matches with the genus *Glyptothorax* but below 92% at the species level. While retrieving the sequences from databases, we found one sequence of *G. indicus* but the constructed tree suggested that this earlier submitted sequence was found to be species *G. telchitta*. In the current DNA barcoding, and high throughput sequencing era, there is the possibility of wrong submissions to public databases [31]. With accurate taxonomic and molecular identification, our work was able to enhance the number of COI gene sequences of important catfish species in the ERN in both databases. Therefore, fish species identified in this work may solve the problem of misidentifications in Nepal. In this study, 99 sequences were generated with 2–6 specimens from each species. A total of 29 Siluriformes fish species were successfully identified representing 21 genera and eight families. DNA barcoding results are in support of morphology-based identification. The utility of DNA barcoding with the COI gene was displayed to be an efficient species identification, nearly 100% success rate recorded in the present study, and this corresponds with other reports [33,34]. Intraspecific genetic divergence in each of *E. sicula* and *G. telchitta* were found to be more than 2%, indicating that there might be putative species. Ward [17] suggested that if there is more than 2% of the genetic divergence with more than 95% probability, then these two specimens can be regarded as different species. Therefore, a detailed taxonomic study focusing on the *E. sicula* and *G. telchitta* is warranted.

The phylogenetic analysis of the fish of order Siluriformes in the ERN disclosed that the species belonging to genus *Eutropiichthys* are closely related to the members of the family, Ailiidae. However, the family, Ailiidae, [35], was set up with three genera (*Ailia*, *Laidess* and *Clupisoma*), but *Eutropiichthys*, *Proeutropiichthys* and *Silonia* are kept in Schilbeidae, where the taxonomic status of four genera were not discussed/unclear. Fricke et al. [2] placed *Ailiichthys*, *Proeutropiichthys*, *Silonia* and *Eutropiichthys* in Ailiidae, where they have included seven genera (*Ailia*, *Laidess*, *Clupisoma*, and *Ailiichthys*, *Proeutropiichthys*, *Silonia* and *Eutropiichthys*), with 25 legitimate nominal species. Our study supports the taxonomic placement of the genus *Eutropiichthys* in the family Ailiidae as suggested in the Eschmeyer fish catalogue. It is pellucid that this was based on circular deducing which also manifests that further research is obligatory to determine the taxonomic position of each genus. The genus *Eutropiichthys* has been described to be sister group of the genus *Clupisoma* and *Clupisoma sinense* is

more closely related to *Lalates longibarbis* with 85.63% similar sequence identity [36]. Mitochondrial gene, though, cannot alone be taken into consideration for phylogenetic studies. Therefore, it is suggested that more studies are needed for both families i.e. Schilbeidae and Ailiidae by using nuclear genes for taxonomic confirmation. The Bagridae and Sisoridae families formed distinct paraphyletic clades with interspecific separation. Several families are represented by the single species viz: Chacidae, Horabagridae, Heteropneutidae and Siluridae. It is recommended that further studies with extensive sampling is essential to understand the freshwater fish diversity in Nepal.

This study used the DNA barcoding technique to understand catfish diversity in eastern Nepal as it provides a reliable, efficient, and cost-effective species identification [38]. DNA barcoding offers a precise method for distinguishing between closely related species and ensures accurate biodiversity assessments [39]. It has helped us to identify two new records of catfishes in eastern Nepal which signifies the possibility of identifying undescribed fish species from the water bodies of Nepal. The freshwater bodies of Nepal are experiencing the problems of biological invasions [40] and the barcoding techniques could be helpful in the early identification of invasive alien species and assist in their timely control. The sequences generated during this study have been deposited on freely accessible databases like GenBank and BOLD which facilitate international collaboration by compiling DNA barcodes of species globally. This collective effort enhances ichthyofaunal diversity research and conservation strategies.

Author Contributions: Jash Hang Limbu performed fieldwork and lab work, data curation and analysis. Jash Hang Limbu and Laxman Khanal analyzed the data and prepared the manuscript. Jin Q. Yang and Chenhong Li supervised the study and improved the manuscript. Jash Hang Limbu, Dipak Rajbanshi, Laxman Khanal, Wang Hui, Sunita Thapa revised and finalized the manuscript. All authors approved the manuscript for submission.

Funding: This study has no any specific funding.

Data Availability: The information about the data used to support the findings of this study is available in the main text. The DNA sequences generated during this study have been submitted to GenBank and respective accession numbers have been provided in the paper.

Acknowledgments: We would like to thank local fishermen Phulchhan Tajpuriya, and Dipendra Tumbapo for support in specimen collection during the study.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Nelson JS, Grande TC, Wilson MVH (2016). *Fishes of the World*. 5th Edition, John Wiley and Sons, Hoboken. <https://doi.org/10.1002/9781119174844>
2. Fricke R, Eschmeyer WN, Laan RV (2024). Eschmeyer's catalog of fishes: Genera, Species. Electronic version accessed 9 July 2024.
3. Grande L, Eastman JT (1986). A review of Antarctic ichthyofaunas in the light of new fossil discoveries. *Palaeontology*, 29(1):113–137.
4. Diogo R (2004). Osteology and myology of the cephalic region and pectoral girdle of *Batrochoglanis raninus*, with a discussion on the synapomorphies and phylogenetic relationships of the Pseudopimelodinae and Pimelodidae (Teleostei: Siluriformes). *Animal Biology*, 54(3):261–280.
5. Malabarba LR, & Malabarba MC (2020). Phylogeny and classification of Neotropical fish. In B. Baldisserotto, E. C. Urbinati, & J. E. P. Cyrino (Eds.), *Biology and Physiology of Freshwater Neotropical Fish* (pp. 1-19). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-815872-2.00001-4>.
6. Ballen GA, De Pinna MCC (2022). A standardized terminology of spines in the order Siluriformes (Actinopterygii: Ostariophysi). *Zoological Journal of the Linnean Society*, 194(2): 601–625. <https://doi.org/10.1093/zoolinnean/zlab008>
7. Shrestha TK (2019). *Ichthyology of Nepal, a study of fishes of the Himalayan waters*. Prism Color Scanning and Press Supportive Pvt. Ltd, Kathmandu, Nepal, 400 pp.
8. Zhang J, Hanner R (2012). Molecular approach to the identification of fish in the South China Sea. *PLoS One*, 7(2): e30621. <https://doi.org/10.1371/journal.pone.0030621>

9. Pandey PK, Singh YS, Tripathy PS, Kumar R, Abujam SK, Parhi J (2020). DNA barcoding and phylogenetics of freshwater fish fauna of Ranganadi River, Arunachal Pradesh. *Gene*, 754:144860. <https://doi.org/10.1016/j.gene.2020.144860>
10. Shen Y, Guan L, Wang D, Gan X (2016). DNA barcoding and evaluation of genetic diversity in Cyprinidae fish in the midstream of the Yangtze River. *Ecology and Evolution*, 6(9):2702–2713. <https://doi.org/10.1002/ece3.2060>.
11. de Carvalho DC, Oliveira DA, Pompeu PS, Leal CG, Oliveira C, Hanner R (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. *Mitochondrial DNA*, 22:80–86. <https://doi.org/10.3109/19401736.2011.588214>
12. Steinke D, Zemlak TS, Hebert PDN (2009). Barcoding Nemo: DNA-based identifications for the ornamental fish trade. *PLoS One*, 4(7): e6300. <https://doi.org/10.1371/journal.pone.0006300>
13. Rosso JJ, Mabragaña E, Castro MG, de Astarloa JM (2012). DNA barcoding Neotropical fishes: recent advances from the Pampa Plain, Argentina. *Molecular Ecology Resources*, 12(6): 999–1011. <https://doi.org/10.1111/1755-0998.12010>
14. Chen W, Ma X, Shen Y, Mao Y, He, S (2015). The fish diversity in the upper reaches of the Salween River, Nujiang River, revealed by DNA barcoding. *Scientific Reports*, 5: 17437. <https://doi.org/10.1038/srep17437>
15. Hebert PD, Cywinska A, Ball SL, de Waard JR (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society: Biological Sciences*, 270(1512): 313–321. <https://doi.org/10.1098/rspb.2002.2218>
16. Kundu, S., Chandra, K., Tyagi, K., Pakrashi, A., and Kumar, V., 2019. DNA barcoding of freshwater fishes from Brahmaputra River in Eastern Himalaya biodiversity hotspot. *Mitochondrial DNA Part B*, 4(2), 2411–2419. <https://doi.org/10.1080/23802359.2019.1637290>
17. Ward RD (2009). DNA barcode divergence among species and genera of birds and fishes. *Molecular Ecology Resources*, 9(4): 1077–1085. <https://doi.org/10.1111/j.1755-0998.2009.02541.x>
18. Limbu JH, Subba S, Gurung JK, Tumbahangfe J, Subba BR (2021b). Correlation of fish assemblages with habitat and environmental variables in the Phewa Khola stream of Mangsebung rural municipality, Ilam, Nepal. *Journal of Animal Diversity*, 3(1): 27–36. <http://dx.doi.org/10.52547/JAD.2021.3.1.5>
19. Khatri, K., B.R. Jha, S. Gurung & U.R. Khadka (2024). Freshwater fish diversity and IUCN Red List status of glacial-fed (Bheri) and spring-fed (Babai) rivers in the wake of inter-basin water transfer. *Journal of Threatened Taxa*, 16(1): 24535–24549. <https://doi.org/10.11609/jott.8084.16.1.24535–24549>
20. Talwar PK, Jhingran AG (1991). Inland Fisheries of India and Adjacent Countries. Volume I and II. Oxford and IBH Publishing Corporation. India, 1158pp.
21. Jayaram KC (2010). The freshwater fishes of Indian region. Narendra Publishing House, Delhi, India, 614pp.
22. Tamura K, Stecher G, Kumar S (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38(7), 3022–3027. <https://doi.org/10.1093/molbev/msab120>.
23. Nguyen LT, Schmidt HA, von Haeseler A, & Minh BQ (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1): 268–274. <https://doi.org/10.1093/molbev/msu300>.
24. Zhang D, F. Gao, I. Jakovlić, H. Zou, J. Zhang, W.X. Li, and G.T. Wang (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources*, 20(1): 348–355. <https://doi.org/10.1111/1755-0998.13096>.
25. Kalyaanamoorthy S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14: 587–589.
26. Galili T, O'Callaghan A, Sidi J, & Sievert C (2018). heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics*, 34(9): 1600–1602. <https://doi.org/10.1093/bioinformatics/btx657>
27. R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
28. Puillandre N, Lambert A, Brouillet S, Achaz G (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, 21: 1864–1877.
29. Ng HH (2005). *Erethistoides sicala*, a new catfish (Teleostei: Erethistidae) from India. *Zootaxa*, 1021, 1–12. <https://doi.org/10.11646/zootaxa.1021.1.1>
30. Vishwanath W, Linthoingambi I (2007). Redescription of catfish *Amblyceps arunachalensis* nath & dey and *Amblyceps apangi* nath & dey (teleostei: amblycipitidae). *Zoos Print Journal* 22(4): 2662–2664.

31. Wong EHK, Hanner RH (2008). DNA barcoding detects market substitution in North American seafood. *Food Research International*, 41:828–837.
32. Lakra WS, Singh M, Goswami M, Gopalakrishnan A, Lal KK, Mohindra V, Ayyappan S (2015). DNA barcoding Indian freshwater fishes. *Mitochondrial DNA*, 27: 4510–4517.
33. Shen Y, Guan L, Wang D, Gan X (2016). DNA barcoding and evaluation of genetic diversity in Cyprinidae fish in the midstream of the Yangtze River. *Ecology and Evolution*, 6(9): 2702–2713.
34. Ude GN, Igwe DO, Brown C. *et al.* (2020). DNA barcoding for identification of fish species from freshwater in Enugu and Anambra States of Nigeria. *Conservation Genetic Resources*, 12: 643–658. <https://doi.org/10.1007/s12686-020-01155-7>
35. Wang J, Lu B, Zan R, Chai J, Ma W *et al.* (2016). Phylogenetic Relationships of Five Asian Schilbid Genera Including *Clupisoma* (Siluriformes: Schilbeidae). *PLoS One*, 11(1): e0145675. <https://doi.org/10.1371/journal.pone.0145675>
36. Karinthanyakit W, Jondeung A (2012). Molecular phylogenetic relationships of pangasiid and schilbid catfish in Thailand. *Journal of Fish Biology*, 80(7): 2549–2570. <https://doi.org/10.1111/j.1095-8649.2012.03303.x>
37. Song J, Hou F, Zhang X, Yue B, Song Z (2014). Mitochondrial genetic diversity and population structure of a vulnerable freshwater fish, rock carp (*Procypris rabaudi*) in upper Yangtze River drainage. *Biochemical Systematics and Ecology*, 55: 1–9. <https://doi.org/10.1016/j.bse.2014.02.008>
38. Araújo RG, Silva RDF, Sampaio I & Guimarães - Costa A (2019). Does DNA barcoding offer meaningful insights into the diversity of the parrotfish of the genus *Sparisoma* (Scaridae)? *Journal of Applied Ichthyology*, 35: 1029 – 1033. <https://doi.org/10.1111/jai.13932>
39. Kumar A, Adhavan DAV & Prakash S (2020). DNA barcoding revealed first record of the ‘fine spotted whiplay’ *Himantura tutul* (Myliobatoidei: Dasyatidae) in the Indian coastal waters. *Journal of Applied Ichthyology*, 36(4): 515-518. <https://doi.org/10.1111/jai.14060>
40. Limbu J, Rajbanshi D, Subba A, Khanal L, Yang JQ & Li C (2024). First record of the non-native vermiculated sailfin catfish *Pterygoplichthys disjunctivus* (Weber, 1991) from Lohandra River, Eastern Nepal. *BioInvasions Records*, 13(2): 557-564. <https://doi.org/10.3391/bir.2024.13.2.20>

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.