



OPEN

Molecular diversity of Uzbekistan's fishes assessed with DNA barcoding

Bakhtiyor Sheraliev^{1,2} & Zuogang Peng¹✉

Uzbekistan is one of two doubly landlocked countries in the world, where all rivers are endorheic basins. Although fish diversity is relatively poor in Uzbekistan, the fish fauna of the region has not yet been fully studied. The aim of this study was to establish a reliable barcoding reference database for fish in Uzbekistan. A total of 666 specimens, belonging to 59 species within 39 genera, 17 families, and 9 orders, were subjected to polymerase chain reaction amplification in the barcode region and sequenced. The length of the 666 barcodes was 682 bp. The average K2P distances within species, genera, and families were 0.22%, 6.33%, and 16.46%, respectively. The average interspecific distance was approximately 28.8 times higher than the mean intraspecific distance. The Barcode Index Number (BIN) discordance report showed that 666 specimens represented 55 BINs, of which five were singletons, 45 were taxonomically concordant, and five were taxonomically discordant. The barcode gap analysis demonstrated that 89.3% of the fish species examined could be discriminated by DNA barcoding. These results provide new insights into fish diversity in the inland waters of Uzbekistan and can provide a basis for the development of further studies on fish fauna.

Spanning more than 35,900 species¹, fish account for half of all extant vertebrate species and are well known for their uneven distribution of species diversity². Consequently, fish constitute a significant component of biodiversity in the composition of animal taxa^{3,4}. Additionally, they have direct economic value and are important sources of animal protein for humans^{5,6}. However, the richness and abundance of fish biodiversity in aquatic ecosystems become more vulnerable, owing to human disturbances^{7,8}. Although approximately 400 new fish species have been described annually over the past 20 years¹, anthropogenic impacts, such as water pollution from plastic and other household waste, river dams, water withdrawal, overfishing, poaching, and habitat degradation have resulted in a catastrophic loss of fish diversity^{9–11}. In-depth taxonomic studies of species are key to conserving biodiversity.

Generally, fish species identification and taxonomy rely on morphometric and meristic characteristics, such as body shape, the number of fin rays or lateral line scales, allometric features, and colour patterns. However, morphological characters are not always stable during various developmental stages and often cannot be assessed in incomplete samples or rare and cryptic species. Moreover, fish identification can be challenging, owing to the similar morphology of congeners during their early life histories as well as due to contradictions in the existing literature and taxonomic history; this is true even if experienced taxonomists work with whole intact adults. In addition, different taxonomists may have different identification abilities and skills, thus even the same specimen may be identified inconsistently, thereby resulting in confusion when summarising and comparing data^{12–14}. However, environmental and conservation studies call for a high level of accuracy, requiring specimens to be identified entirely at the species level¹⁵. The inherent limitations of morphology-based taxonomy and the decreased number of taxonomists require molecular approaches for fish species identification¹⁶.

Molecular identification, which identifies species using molecular markers, is widely used today. Among the various molecular approaches used for species molecular identification, DNA barcoding based on mitochondrial DNA (mtDNA) is one of the most suitable tools for species-level identification^{17,18}. In addition, mtDNA-based molecular identification has several advantages over morphological approaches. First, species identification does not require complete specimens; however, a tiny piece of tissue such as muscle, skin, fin, or teeth is acceptable for DNA extraction^{18–20}. Second, DNA is more stable than morphological characters and is more resistant to degradation. For example, DNA can be extracted from water and soil previously occupied by an organism, or from samples that have been processed or digested^{21–24}. Third, it is difficult to distinguish some species with

¹Key Laboratory of Freshwater Fish Reproduction and Development (Ministry of Education), School of Life Sciences, Southwest University, Chongqing 400715, China. ²Faculty of Life Sciences, Fergana State University, Fergana, Uzbekistan. ✉email: pzg@swu.edu.cn

similar morphological characteristics, such as cryptic or sibling species^{25–27}. Molecular identification can help accurately distinguish among such species^{28,29}. Fourth, DNA is invariable throughout the developmental stages of an organism. In contrast, morphological characters can change during a life cycle, thereby leading to species misidentification¹². Therefore, molecular approaches can be applied in the identification of fish eggs, larvae, juveniles, and adults^{13,30}. Fifth, becoming a professional traditional taxonomist requires a lot of time, work, and resources^{31,32}. Advances in technology make it fairly easy to replicate and read DNA sequences, while bioinformatic software can automatically compare the resulting sequences; therefore, the training required to approach molecular identification is much less than that required for morphological identification. Molecular identification is widely used in a number of other fields besides species identification, including illegal species trade, food fraud, biological invasions, and biodiversity monitoring^{33–36}.

Hebert et al.¹⁷ pioneered the use of cytochrome *c* oxidase subunit I (COI) for molecular species identification, showing that this genetic marker can serve as a DNA barcode for biological identification in both invertebrates and vertebrates^{18,28,37–39}. The Fish Barcode of Life Initiative (FISH-BOL) is an international research collaboration aimed at creating a standardised reference library of DNA barcodes for all fish species^{40,41}. The main goal of this project is to enable the identification of fish species by comparing the sequence of queries against the database of reference sequences in the Barcode of Life Data Systems (BOLD)⁴². To date, many studies have been carried out worldwide on fish DNA barcoding dedicated to FISH-BOL^{3,4,18,43,44}. Compared to other regions of the world, studies devoted to fish barcoding are almost absent in Central Asia.

Uzbekistan is one of two doubly landlocked countries in the world, where all rivers are endorheic basins; therefore, fish biodiversity is poor. According to Mirabdullaev and Mullabaev⁴⁵, the total number of fish species in Uzbekistan exceeds 71, of which 26 species were introduced from other water bodies into the inland waters of the country. At the same time, the drying up of the Aral Sea, which is the largest water basin in the region, global climate change, population growth, river damming, water pollution, water withdrawals for agriculture, poaching, overfishing, and habitat destruction, all affect the fish species in the region^{46,47}. To date, studies on piscifauna have been based mainly on traditional morphological criteria and have not been comprehensively barcoded, except in our recent studies^{48–50}. Recently, molecular identification has been applied to identify mainly nematodes among animal species⁵¹.

Consequently, the main aim of the present study was to provide the first inventory of freshwater fish species in Uzbekistan based on DNA barcoding. This inventory could serve as a reference for screening DNA sequences in future studies. Additionally, we assessed the genetic diversity of freshwater fish species. The DNA barcode records generated in this study will be available to researchers for the monitoring and conservation of fish diversity in Uzbekistan.

Results

Morphology-based species identification. First, all collected specimens were identified using morphological approaches. Morphological identification classified all samples into 59 species belonging to 39 genera and 17 families that represented nine orders (Table 1). The identified specimens included 50 (84.75%) species identified to the species level and nine (15.25%) species that could not be identified to the species level (Tables 1 and S2). Approximately three-quarters of the species (44 species, 74.58%) belonged to the order Cypriniformes. The remaining eight orders included one or two species.

Of the 59 fish species collected from the inland waters of Uzbekistan, *Pseudoscaphirhynchus hermanni* and *P. kaufmanni* were classified as critically endangered (CR), *Acipenser baerii* and *Capoetobrama kuschakewitschi* were classified as endangered (EN), and *Cyprinus carpio* and *Luciobarbus brachycephalus* were classified as vulnerable (VU) according to International Union for Conservation of Nature's (IUCN) Red List of Threatened Species. The remaining species were grouped into the least concern (LC) and data deficient (DD) categories (Table 1).

Identification of fish species using DNA barcodes. A total of 666 fish samples were successfully amplified using three primers and PCR. After editing, all COI barcode sequences were 682 for each sample and the mean nucleotide frequencies of the entire dataset were A (24.49%), T (29.01%), G (18.50%), and C (28.00%). The genetic distance within species ranged from 0.000 to 0.0149.

For species identification at the species level, a total of 666 COI barcode sequences representing 59 different species were employed (mean of 11.3 samples per species). The GenBank and BOLD databases were used for species identification (Table S2). The GenBank-based identification of all species ranged from 98.58 to 100.00%. The COI sequences of 22 fish species had not been reported in the GenBank database. Among them, *P. hermanni* was identified as *P. kaufmanni*, *Cottus spinulosus* as *C. ricei*, *L. conocephalus* as *L. capito*, *Alburnus oblongus* and *A. taeniatus* as *A. escherichii*, *Leuciscus lehmanni* and *Petroleuciscus squaliusculus* as *L. baicalensis*, and *Triplophysa* sp. 1 as *T. aliensis* with 99.71%, 98.47%, 98.83–100%, 98.39–98.82%, 99.71–99.85%, and 98.37% similarity, respectively.

The BOLD-based identification of 46 fish species ranged from 98.36 to 100%. No matches were found for 13 species. *Pseudoscaphirhynchus hermanni* was identified as *P. kaufmanni*, *Cottus spinulosus* as *C. ricei*, *A. oblongus* and *A. taeniatus* as *A. escherichii*, *L. lehmanni* and *P. squaliusculus* as *L. baicalensis*, and *Triplophysa* sp. 1 as *T. aliensis* with 99.85–100%, 98.48%, 98.62–98.92%, 99.8%–100%, and 98.36% similarity, respectively. Despite the GenBank databases, *L. conocephalus*, *Neogobius pallasii*, and *Rhinogobius* sp. were identified with high similarities (>99.4%).

The Taxon ID tree shows that the specimens formed phylogenetic clusters that reflected previous taxonomic results based on morphology (Fig. S1). In turn, the barcode gap analysis revealed that five species lacked a barcode gap (intraspecific K2P distance \geq interspecific one), and four species had a low K2P distance to another species ($\leq 2\%$), which indicates that the majority of the investigated species could be identified by the DNA

Species no.	Order	Family	Genus	Species	Sample size	IUCN status
1	Acipenseriformes	Acipenseridae	<i>Acipenser</i>	<i>A. baerii</i>	2	EN
2			<i>Pseudoscaphirhynchus</i>	<i>P. hermanni</i>	3	CR
3				<i>P. kaufmanni</i>	2	CR
4	Anabantiformes	Channidae	<i>Channa</i>	<i>Ch. argus</i>	6	-
5	Cypriniformes	Acheilognathidae	<i>Rhodeus</i>	<i>R. ocellatus</i>	21	DD
6				<i>Rhodeus</i> sp.	1	-
7		Cobitidae	<i>Sabanejewia</i>	<i>S. aurata</i>	27	LC
8		Cyprinidae	<i>Capoeta</i>	<i>C. heratensis</i>	25	-
9			<i>Carassius</i>	<i>C. auratus</i>	7	LC
10				<i>C. gibelio</i>	28	-
11			<i>Cyprinus</i>	<i>C. carpio</i>	9	VU
12			<i>Luciobarbus</i>	<i>L. brachycephalus</i>	8	VU
13				<i>L. conocephalus</i>	15	-
14			<i>Schizothorax</i>	<i>S. eurystomus</i>	38	LC
15				<i>S. fedtschenkoi</i>	5	LC
16				<i>Schizothorax</i> sp.	5	-
17		Gobionidae	<i>Abbottina</i>	<i>A. rivularis</i>	11	-
18			<i>Gobio</i>	<i>G. lepidolaemus</i>	18	LC
19				<i>G. nigrescens</i>	8	-
20				<i>G. sibiricus</i>	2	LC
21			<i>Pseudorasbora</i>	<i>P. parva</i>	11	LC
22		Leuciscidae	<i>Abramis</i>	<i>A. brama</i>	2	LC
23			<i>Alburnoides</i>	<i>A. holciki</i>	41	-
24			<i>Alburnus</i>	<i>A. chalcoides</i>	12	-
25				<i>A. oblongus</i>	8	-
26				<i>A. taeniatus</i>	3	DD
27			<i>Capoetobrama</i>	<i>C. kuschakewitschi</i>	8	EN
28			<i>Leuciscus</i>	<i>L. aspius</i>	1	LC
29				<i>L. lehmanni</i>	10	LC
30			<i>Pelecus</i>	<i>P. cultratus</i>	4	LC
31			<i>Petroleuciscus</i>	<i>P. squaliusculus</i>	10	LC
32			<i>Rutilus</i>	<i>R. lacustris</i>	11	-
33		Nemacheilidae	<i>Dzihunia</i>	<i>D. amudarjensis</i>	11	LC
34				<i>Dzihunia</i> sp. 1	11	-
35				<i>Dzihunia</i> sp. 2	3	-
36				<i>Dzihunia</i> sp. 3	8	-
37			<i>Paracobitis</i>	<i>P. longicauda</i>	25	-
38			<i>Triplophysa</i>	<i>T. ferganaensis</i>	20	-
39				<i>T. strauchii</i>	29	LC
40				<i>Triplophysa</i> sp. 1	4	-
41				<i>Triplophysa</i> sp. 2	4	-
42		Xenocyprididae	<i>Ctenopharyngodon</i>	<i>C. idella</i>	8	-
43			<i>Hemiculter</i>	<i>H. leucisculus</i>	25	LC
44			<i>Hypophthalmichthys</i>	<i>H. molitrix</i>	8	NT
45				<i>H. nobilis</i>	5	DD
46			<i>Mylopharyngodon</i>	<i>M. piceus</i>	2	DD
47			<i>Opsariichthys</i>	<i>O. bidens</i>	7	LC
48			<i>Parabramis</i>	<i>P. pekinensis</i>	13	-
49	Cyprinodontiformes	Poeciliidae	<i>Gambusia</i>	<i>G. holbrooki</i>	44	LC
50	Esociformes	Esocidae	<i>Esox</i>	<i>E. lucius</i>	3	LC
51	Gobiiformes	Gobiidae	<i>Neogobius</i>	<i>N. melanostomus</i>	1	LC
52				<i>N. pallasii</i>	1	LC
53			<i>Rhinogobius</i>	<i>Rhinogobius</i> sp.	37	-
54	Perciformes	Cottidae	<i>Cottus</i>	<i>C. spinulosus</i>	6	LC
55		Percidae	<i>Sander</i>	<i>S. lucioperca</i>	10	LC
56	Salmoniformes	Salmonidae	<i>Oncorhynchus</i>	<i>O. mykiss</i>	2	-

Continued

Species no.	Order	Family	Genus	Species	Sample size	IUCN status
57	Siluriformes	Siluridae	<i>Silurus</i>	<i>S. glanis</i>	8	LC
58		Sisoridae	<i>Glyptosternon</i>	<i>G. oschanini</i>	1	LC
59				<i>Glyptosternon</i> sp.	8	–

Table 1. Classification, sample size and INCN status of the fish species of Uzbekistan using in this study. IUCN International Union for Conservation of Nature, CR critically endangered, EN endangered, VU vulnerable, NT near threatened, LC least concern, DD data deficient, – no assessment.

	N	Taxa	Comparisons	K2P genetic distance (%)		
				Minimum	Maximum	Mean and SD
Within species	661	54	6608	0.00	1.49	0.22 ± 0.00
Within genus	309	13	2534	0.00	11.78	6.33 ± 0.00
Within family	512	3	75,196	0.00	22.19	16.46 ± 0.00

Table 2. Summary of K2P genetic distances (%) calculated for different taxonomic levels.

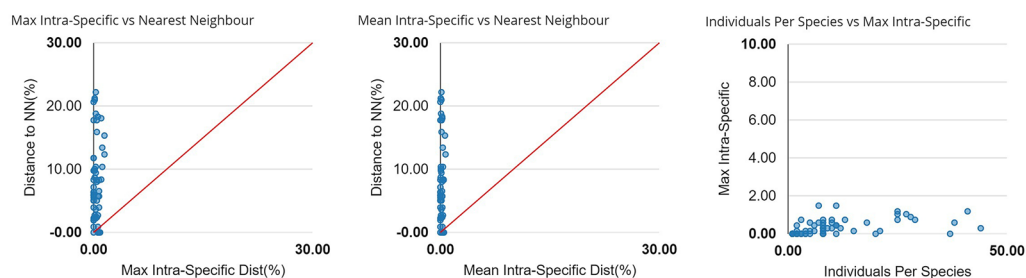


Figure 1. Barcoding gap: Maximum intraspecific Kimura 2-parameter (K2P) distances compared with the minimum interspecific K2P distances recorded in fish from Uzbekistan. The graphs show the overlap of the maximum and mean intra-specific distances with the inter-specific (NN = nearest neighbor) distances.

barcode approach (Table S3). Generally, the mean K2P distance of a species to its nearest neighbour (NN) was 8.04% (SD: 0.11%).

The mean K2P distances within species, within genera, and within families were 0.22%, 6.33%, and 16.46%, respectively (Table 2; Fig. 1). The largest intraspecific K2P distance was observed in *Opsariichthys bidens* (five specimens; Fig. 2; Table S3). The specimens obtained from several species, such as *Abramis brama* (two specimens), *Capoetobrama kuschakewitschi* (eight specimens), *Gobio nigrescens* (eight specimens), and *Rhinogobius* sp. (37 specimens), carried the same haplotype (Table S3). The average congeneric distance was approximately 28.8 times higher than the mean conspecific distance, but approximately 2.6 times less than the average genetic distance between families, thus the average genetic distance grew based on the taxonomic level.

The Barcode Index Number (BIN) discordance report showed that 666 specimens represented 55 BINs; among them, 45 BINs were taxonomically concordant, five BINs were taxonomically discordant, and five BINs were singletons. For the best match (BM), best close match (BCM), and all species barcodes (ASB) analyses of the 666 sequence data set with singletons, the percentages of correct identification were 94.74%, 94.74%, and 89.03%, respectively; those of ambiguous identification were 4.05%, 4.05%, and 10.51%, respectively; those of incorrect identification were 1.2%, 1.2%, and 0.44%, respectively. Moreover, for the same three analyses of the dataset without singletons (661 sequences), the percentages of correct identification were 95.46%, 95.46%, and 89.71%, respectively; those of ambiguous identification were 3.93%, 3.93%, and 10.13%, respectively; those of incorrect identification were 0.6%, 0.6%, and 0.15%, respectively (Table 3).

Automated barcode gap discovery (ABGD) analyses of species delimitation. The ABGD tool was used for species delimitation. A partition with prior maximal distance $P = 0.0359$ and 0.0046 delimited the entire dataset into 55 putative species (Table 4). Of the 59 morphological-based identified species, 55 (93.22%) were delimited clearly through the ABGD at a prior maximal distance of 0.0359 , which was consistent with the observations of genetic distance and neighbour-joining (NJ) and Bayesian inference (BI) analyses (Figs. S1 and 3). Furthermore, at a prior maximal distance of 0.0359 , few species, such as *Carassius auratus*, *C. gibelio*, *Gobio lepidolaemus*, *G. sibiricus*, *L. lehmanni*, *P. squaliusculus*, *P. hermanni*, and *P. kaufmanni* could not be delimited

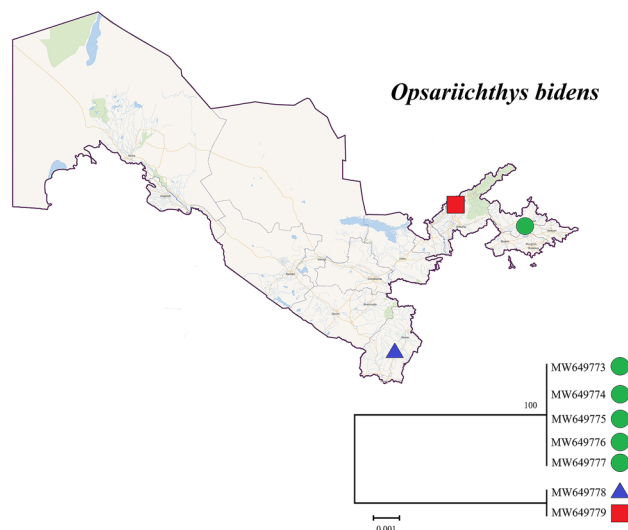


Figure 2. Neighbour-joining tree of *Opsariichthys bidens* from DNA barcode sequences with 100 000 bootstrapping replicates. Sampling localities: Syr Darya (green circle), Chirchik River (red square), and Surkhan Darya (blue triangle).

	Best match (%)	Best close match (%)	All species barcodes (%)
With singletons			
Correct identifications	631 (94.74%)	631 (94.74%)	593 (89.03%)
Ambiguous identifications	27 (4.05%)	27 (4.05%)	70 (10.51%)
Incorrect identifications	8 (1.2%)	8 (1.2%)	3 (0.44%)
Sequences without any match closer than threshold	NA	NA	NA
Without singletons			
Correct identifications	631 (95.46%)	631 (95.46%)	593 (89.71%)
Ambiguous identifications	26 (3.93%)	26 (3.93%)	67 (10.13%)
Incorrect identifications	4 (0.6%)	4 (0.6%)	1 (0.15%)
Sequences without any match closer than threshold	NA	NA	NA

Table 3. Results of identification success analysis for the criteria: best match, best close match and all species barcodes.

Partition	Groups number	Prior maximal distance (P)
1	165	0.001000
2	75	0.001668
3	75	0.002783
4	56	0.004642
5	56	0.007743
6	56	0.012915
7	55	0.021544
8	55	0.035938

Table 4. The number of groups inferred from ABGD analysis according to prior intraspecific divergence (P).

into different putative species. No clear divergence between these morphologically distinct species was observed in the NJ and BI analyses, with the exception of *Gobio* species.

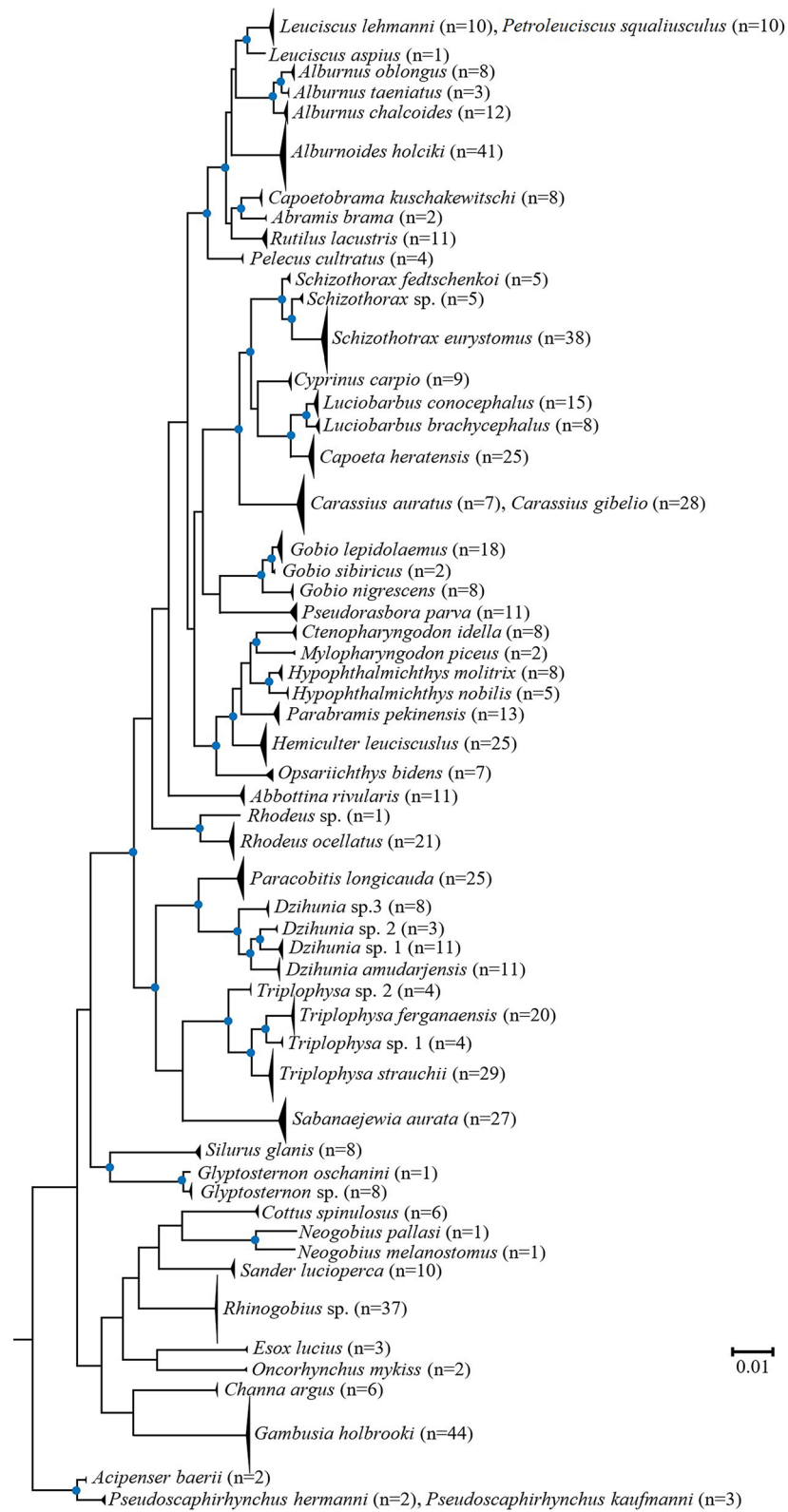


Figure 3. Bayesian inference (BI) consensus tree based on the COI partial gene sequences. The blue circle at nodes represents BI posterior probabilities values > 50%. Posterior probability values for all species are > 95%.

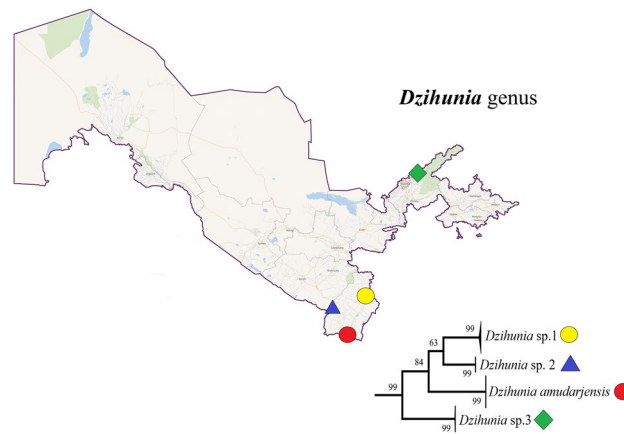


Figure 4. Neighbour-joining tree of *Dzhunzia* spp. from DNA barcode sequences with 100,000 bootstrapping replicates. Sampling localities: lower Surkhan Darya (red circle) upper Surkhan Darya (yellow circle), Sherabad River (blue triangle) and Chirchik River (green diamond).

Discussion

This study of the fish fauna of the inland waters of Uzbekistan is the first to compile the data in a sequence library, which contributes to the FISH-BOL in the BOLD system. This study included the molecular identification of 59 species. These 59 species included 83.1% of the reported fish fauna of the region⁴⁵. Relationships among species are shown in the topology of the BI tree (Fig. 3).

The gap between COI intraspecific and interspecific diversity is called the ‘barcode gap’, which is decisive for the discriminatory ability of DNA barcoding⁵². The barcode gap can be seen in our study (Table 2), as well as in many other previous studies^{3,44,53}, thereby further confirming that this approach is an effective way to distinguish between fish species.

This study clarified the taxonomic status of a number of taxa, such as *Alburnoides oblongus* and *A. taeniatus*, which belong to *Alburnus*, which is consistent with the results of Matveyev et al.⁵⁴ and Jouladeh-Roudbar et al.⁵⁵; *Schizothorax fedtschenkoi* is a valid species; another *Schizothorax* sp. from the southern part of the country is an undescribed species; the *Alburnoides* population (previously considered as *A. eichwaldii*) from the inland waters of Uzbekistan, is de facto *A. holciki*⁴⁹; three *Gobio* species occur in the inland waters of the country⁵⁰; *Glyptosternon* and *Rhodeus* each consist of two species and not just one, as previously believed; thus, additional taxonomic research is required; two species of the genus *Neogobius* (*N. melanostomus* and *N. pallasii*) (previously believed to belong to *N. melanostomus* and *N. fluviatilis*⁵⁶) occurred in the lower reaches of the Amu Darya; the population of *Opsariichthys* in Uzbekistan belongs to the same species, and *O. bidens* is not *O. uncistrostris* as previously believed⁵⁶; the entire *Rhinogobius* population in Uzbekistan belongs to the same species (*Rhinogobius* sp.), which is neither *R. brunneus* nor *R. similis* as previously thought^{56,57}; thus, taxonomic clarification is required (Figs. S1, 3; Table S2). Moreover, local researchers initially believed that *Gambusia affinis holbrooki* was introduced into the inland waters of the country last century. Later, the taxonomic status of this subspecies was raised to the valid species. Nevertheless, both *G. affinis* and *G. holbrooki* were considered to be found in the waters of the country^{56,58}. For the first time, our study proved that only one (*G. holbrooki*) of these species is found in Uzbekistan (Figs. S1, 3).

Only a single species of *Petroleuciscus* in Central Asia from the upper reaches of the Syr Darya, joined with *Leuciscus lehmanni* from the Zeravshan River in our phylogenetic analysis based on the COI barcode marker. *Petroleuciscus squaliusculus* was originally described in the genus *Squalius*. Previously, it was repeatedly assigned to the genus *Leuciscus*^{59,60}. Although these two species showed a very low genetic distance in our phylogenetic analysis, *P. squaliusculus* can be easily distinguished from *L. lehmanni* by processing convex posterior dorsal and anal-fin margins (vs. concave). However, our unpublished work (nuclear molecular and morphology) showed that they are two separate valid species, and *Petroleuciscus squaliusculus* belongs to *Leuciscus*.

Currently, three *Dzhunzia* Prokofiev, 2001 species are found in the Amu Darya (*D. amudarjensis*), Zeravshan (*D. ilan*), and Talas (*D. turdakovi*, outside Uzbekistan) rivers^{61,62}. Apparently, the species diversity of *Dzhunzia* seems to be much higher than previously thought (Fig. 3). In addition to *D. amudarjensis*, two more undescribed species were found in the upper reaches of Amu Darya. Another undescribed species was found in the Chirchik River; however, members of *Dzhunzia* had not previously been found in this river (Fig. 4). On the other hand, *D. ilan* was not found in two of our expeditions to the Zeravshan River; moreover, it is believed that this species may have become extinct⁶¹.

The inability of DNA barcodes to identify species may be due to incomplete sorting by lineage associated with recent speciation^{63,64} and haplotype sharing as a result of hybridisation⁶⁵. In our study, DNA barcodes of two *Leuciscus* and *Petroleuciscus* (*L. lehmanni* and *P. squaliusculus*), two *Carassius* (*C. auratus* and *C. gibelio*), and two *Pseudoscaphirhynchus* (*P. hermanni* and *P. kaufmanni*) species were sequenced, and the BIN discordance report illustrated that these six species could not be distinguished by the COI barcode gene (Figs. S1 and 3). In this case, a more rapidly evolving DNA fragment, such as the mitochondrial control region or the first internal transcribed ribosomal DNA spacer, may be better for identification³. A similar situation occurred with *Carassius*

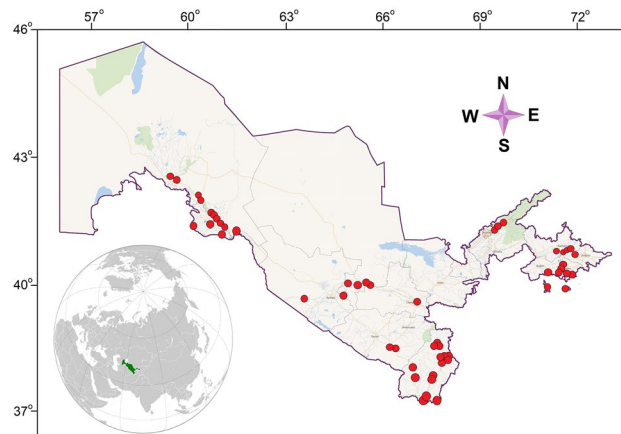


Figure 5. Map of sampling sites in Uzbekistan, 2016–2020. A symbol may cover more than one sampling site. The map was constructed in QGIS 2.18.27 (QGIS Development Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation. <http://qgis.org>), using free vector and raster map data made available by Wikimedia maps beta (<https://maps.wikimedia.org/>). All maps are in the public domain (<https://maps.wikimedia.org/>).

species collected in the Mediterranean basin⁶⁶. In addition, among the three *Leuciscus* (*L. baicalensis*, *L. cf. latus*, and *L. schmidti*) species from China, Kazakhstan, and Russia, very low interspecific differences were found based on the COI gene⁶⁷. However, in *Pseudoscaphirhynchus* species, no interspecies differences were found either when using other rapidly evolving mtDNA markers⁶⁸, the entire mtDNA genome⁶⁹, or nDNA markers (our unpublished data). In fact, these two sturgeon species are morphologically easy to distinguish from each other⁷⁰. Thus, the complete genome sequencing of *Pseudoscaphirhynchus* may be important for their molecular authentication.

Unexpectedly, *Abbottina rivularis* from Gobionidae is nested with members of the genus *Rhodeus* from Acheilognathidae in our NJ phylogenetic tree (Fig. S1). A similar result was achieved when we excluded morphological error or DNA contamination. Despite the sharp differences in morphology, the fact that these two genera are sister taxa has also been observed in previous studies^{71,72}.

The global fish diversity is currently a serious threat. Along with natural limiting factors to native species, the negative impact of introduced species is also increasing^{73–76}. At the same time, the negative impact of anthropogenic factors on the biodiversity of freshwater basins is also growing⁷⁷. The number of biological species is declining annually; therefore, DNA barcoding is becoming a versatile approach that can be used to assess fish biodiversity, monitor fish conservation, and manage fishery resources^{78–81}. While our DNA barcoding study is beneficial for the taxonomy of fishes in the Amu Darya and Syr Darya basins, it is also important to clarify the taxonomy of misidentified invasive species acclimatised to Central Asian watersheds⁵⁸.

Unfortunately, fish diversity in Uzbekistan has decreased in recent years. A rare sturgeon fish, *Acipenser nudiventris*, is completely extinct in the Aral Sea basin⁸². Another sturgeon species endemic to the Syr Darya, *Pseudoscaphirhynchus fedtschenkoi*, has been possibly extinct since the 1990s⁶⁹. The Syr Darya population of *Capoetobrama kuschakewitschi* has not been recorded in recent decades, and so far, this species has survived only in the lower reaches of the Amu Darya⁸³. *Gymnocephalus cernuus* and *Perca fluviatilis* have not been recorded in water bodies in the country since the late 1990s⁴⁵. Monitoring the existing populations of other rare native fish species and studying the negative impact of invasive species on them is advisable. The traditional monitoring of fish diversity is usually time-consuming, expensive, and labour intensive. However, with an ever-expanding barcode database and advances in biotechnology (such as environmental DNA analysis), the assessment of fish diversity is becoming more efficient^{84–86}. As our molecular study of fishes develops in Uzbekistan, data on fish species in this region will become more readily available than ever.

Methods

Ethical statement. Fieldwork in the inland waters of Uzbekistan was carried out in accordance with the requirements of the current Uzbek environmental legislation, with the Law of the Republic of Uzbekistan ‘On the protection and use of wildlife’ (No. 545-I 26.12.1997; <https://lex.uz/docs/-31719>) and the Guidelines for the Use of Fishes in Research (2013) (section 5) of the American Society of Ichthyologists and Herpetologists (ASIH) (<https://asih.org/animal-care-guidelines/>). Also, the competent authority (the Ethics Committee of the Southwest University (IACUC No. Approved: IACUC-20160201-01), China) approved the fish sampling and protocols of the present study and reporting herein follows the recommendations in the ARRIVE guidelines.

Sample collection and morphological identification. A total of 666 fish samples were collected from February 2016 to August 2020 using gill nets or cast nets from 53 distant locations in different rivers, tributaries, canals, springs, and lakes (Fig. 5). Information about the sampling stations, along with geographical coordinates and sampling dates, is given in Table S1.

Initially, all specimens were identified to the species level based on morphological characteristics following the identification keys of Berg^{59,70} and Mirabdullaev et al.⁸⁷. If identification was not correctly assigned to a specific species, the 'sp.' and 'cf.' abbreviations were applied⁸⁸. Two pieces of right pectoral fin tissue and muscle tissue were dissected from each fish specimen and stored in 99% ethanol at -20°C . Fin-clipped whole specimens and excess specimens for further morphological analyses were fixed in 10% formalin. After 5–7 days they were transferred to 70% ethanol for long-term storage and deposited in the Key Laboratory of Freshwater Fish Reproduction and Development at the Southwest University, School of Life Sciences (China), respectively, with the exception of sturgeon species, which were deposited in the Department of Biology at the Fergana State University, Faculty of Life Sciences (Uzbekistan).

DNA extraction, COI amplification, and DNA sequencing. Genomic DNA was extracted from muscle or fin tissues by proteinase K digestion followed by a standard phenol–chloroform method. The DNA concentration was estimated using a nano-volume spectrophotometer (NanoDrop 2000; Thermo Fisher Scientific Inc., Waltham, MA, USA) and stored at -20°C for further use. Approximately 680 bp were amplified from the 5' region of the COI gene using the fish-specific primers described by Ivanova et al.⁸⁹: FishF2_t1 TGT AAA ACG ACG GCC AGT CGA CTA ATC ATA AAG ATA TCG GCA C and FishR2_t1 CAG GAA ACA GCT ATG ACA CTT CAG GGT GAC CGA AGA ATC AGA A, respectively. The following primers¹⁸ were used for *Gambusia holbrooki*: FishF2-TCG ACT AAT CAT AAA GAT ATC GGC AC and FishR2-ACT TCA GGG TGA CCG AAG AAT CAG AA. The following primers⁹⁰ were used for sorigid catfishes: catF-TCT CAA CCA ACC ATA AAG ACA TTG G and catR-TAT ACT TCT GGG TGC CCA AAG AAT CA.

The PCR reactions were performed in a final volume of 25 μL , containing 10–100 ng template DNA, five μmol of each forward and reverse primer, while 12.5 μL of $2 \times \text{Taq}$ Master Mix (Novoprotein, Guangdong, China) and double-distilled water were also used. The thermal conditions consisted of an initial step of 3 min at 94°C followed by 35 cycles of 0.5 min at 94°C , 45 s at 54°C , and 1 min 10 s at 72°C , followed by a final extension of 7 min at 72°C . The reactions were performed in an Applied Biosystems thermocycler (Veriti™ 96-Well Thermal Cycler, Singapore), and the PCR products were evaluated by electrophoresis using 1% agarose gel stained with BioRAD (Universal Hood II; Des Plaines, IL, USA). The PCR products were sent to TsingKe Biological Technology Co., Ltd. (Chongqing) for sequencing.

Molecular data analysis. All sequences were manually edited using the SeqMan program (DNASTar software) combined with manual proofreading; all contig sequences started at the first codon position and ended at the third position; no stop codons were also detected. All obtained barcodes were uploaded to the BOLD and GenBank databases, and the details are given in Table S1.

The COI barcode sequence of each sample was identified by the scientific name or species using the BLAST and BOLD databases. Specimens were classified by family, genus, and species according to the fish taxonomic systems of Fricke et al.⁶², and their status was checked in the IUCN Red List of Threatened Species v. 2020-3. The results of species identification based on the BLAST and BOLD databases are presented in Table S2.

We uploaded the entire data set to BOLD under project title 'Freshwater fishes of Uzbekistan'. BOLD version 4 analytical tools were used for the following analyses. The distance summary with the parameter setting the Kalign alignment option⁹¹ and pairwise deletion (ambiguous base/gap handling) was employed to estimate the Kimura 2-parameter (K2P)⁹² distances for taxonomic ranks at the species, genus, and family levels. Barcode gap analysis was carried out with the setting of the parameter 'K2P; kalign alignment option; pairwise deletion (ambiguous base/gap handling)' to construct the distribution of intraspecific and interspecific genetic distances [nearest neighbour (NN) analysis]. The BIN discordance report was employed to confirm the exactness of species identification, as well as to check for cases of low levels of genetic differentiation between different species. The Taxon ID tree was used to construct an NJ tree of the entire 666 sequences with the parameter-setting K2P distance model, the Kalign alignment algorithm⁹¹, and pairwise deletion (ambiguous base/gap handling).

To verify intraspecific and interspecific genetic distances, we also used barcode gap analyses in ABGD (<https://bioinfo.mnhn.fr/abi/public/abgd/>)⁹³. ABGD was used with K2P with the transition/transversion ratio (TS/TV) set to 2.0, 10 recursive steps, X (relative gap width) = 1.0; the remaining parameters were set to default values (Pmin = 0.001, Pmax = 0.1, Nb bins = 20).

We also used SPECIESIDENTIFIER v1.7.8 (<http://taxondna.sourceforge.net/>)⁹⁴ to verify species identification success by applying three criteria (BM, BCM, and ASB) to the entire barcode dataset, following Meier et al.⁹⁴. Fish species that had only one sequence (singletons) were automatically assigned as 'incorrectly identified' under the BM and BCM criteria, as there were no conspecific barcoding sequences to match.

For phylogenetic reconstructions, the datasets were analysed based on the BI methodology using MrBayes 3.2⁹⁵. MrBayes was run with six substitution types (nst = 6), and we considered the gamma-distributed rate variation and the proportion of invariable positions (GTR + G + I) for the COI datasets. For BI, we ran four simultaneous Monte Carlo Markov chains for 25,000,000 generations, with sampling every 1000 generations. The chain temperature was set at 0.2. Log-likelihood stability was determined after 10,000 generations, and we excluded the first 1000 trees as burn-in. The remaining trees were used to compute a 50% majority-rule consensus tree. Moreover, to reveal the phylogenetic relationship of some fish species, the NJ tree of the K2P distance was constructed using MEGA7⁹⁶. Phylogenetic trees were visualised and edited using FigTree 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>)⁹⁷.

Data availability

All sequences and associated voucher data are available from BOLD (process ID from FFU001-20 to FFU666-21) and GenBank (accession numbers MN872388–MN872408, MW649153–MW649792). All other data are available

in Supplementary Information: Fig S1. Neighbor-joining tree based on the COI partial gene sequences; Table S1. Voucher metadata; Table S2. Fish species identification from GenBank and BOLD databases; Table S3. Barcode Index Number (BIN), average and maximum intraspecific distance and distance to nearest neighbor (NN).

Received: 8 May 2021; Accepted: 31 July 2021

Published online: 19 August 2021

References

1. Fricke, R., Eschmeyer, W. N. & Fong, J. D. *Eschmeyer's Catalog of Fishes: Genera/Species by Family/Subfamily*. <http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp> (2021).
2. Nelson, J. S., Grande, T. C. & Wilson, M. V. H. *Fishes of the World* 5th edn. (Wiley, Hoboken, 2016).
3. Chang, C. H. *et al.* DNA barcodes of the native ray-finned fishes in Taiwan. *Mol. Ecol. Resour.* **17**, 796–805 (2017).
4. Kundu, S., Chandra, K., Tyagi, K., Pakrashi, A. & Kumar, V. DNA barcoding of freshwater fishes from Brahmaputra River in Eastern Himalaya biodiversity hotspot. *Mitochondrial DNA Part B* **4**, 2411–2419 (2019).
5. Sargent, J. R. & Tacon, A. G. J. Development of farmed fish: a nutritionally necessary alternative to meat. *Proc. Nutr. Soc.* **58**, 377–383 (1999).
6. Tacon, A. G. J. & Metian, M. Food matters: Fish, income, and food supply: A comparative analysis. *Rev. Fish. Sci. Aquac.* **26**, 15–28 (2018).
7. Jackson, J. B. C. *et al.* Historical overfishing and the recent collapse of coastal ecosystems. *Science* **293**, 629–637 (2001).
8. Clausen, R. & York, R. Global biodiversity decline of marine and freshwater fish: a cross-national analysis of economic, demographic, and ecological influences. *Soc. Sci. Res.* **37**, 1310–1320 (2008).
9. Bryant, M. D. Global climate change and potential effects on Pacific salmonids in freshwater ecosystems of southeast Alaska. *Clim. Change* **95**, 169–193 (2009).
10. Dallas, H. F. & Rivers-Moore, N. A. Ecological consequences of global climate change for freshwater ecosystems in South Africa. *S. Afr. J. Sci.* **110**, 48–58 (2014).
11. Shelton, J. M. *et al.* Vulnerability of cape fold ecoregion freshwater fishes to climate change and other human impacts. *Aquat. Conserv.* **28**, 68–77 (2018).
12. Tillett, B. J. *et al.* Accuracy of species identification by fisheries observers in a north Australian shark fishery. *Fish. Res.* **127–128**, 109–115 (2012).
13. Ko, H. L. *et al.* Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS ONE* **8**, e53451. <https://doi.org/10.1371/journal.pone.0053451> (2013).
14. Hulley, E. N., Tharmalingam, S., Zarnke, A. & Boreham, D. R. Development and validation of probe-based multiplex real-time PCR assays for the rapid and accurate detection of freshwater fish species. *PLoS ONE* **14**, e0210165. <https://doi.org/10.1371/journal.pone.0210165> (2019).
15. Krishna Krishnamurthy, P. & Francis, R. A. A critical review on the utility of DNA barcoding in biodiversity conservation. *Biodivers. Conserv.* **21**, 1901–1919 (2012).
16. Hopkins, G. W. & Freckleton, R. P. Declines in the number of amateur and professional taxonomists: implications for conservation. *Anim. Conserv.* **5**, 245–249 (2002).
17. Hebert, P. D. N., Cywinska, A., Ball, S. L. & DeWaard, J. R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B. Biol. Sci.* **270**, 313–321 (2003).
18. Ward, R. D., Zemlak, T. S., Innes, H. B., Last, R. P. & Hebert, P. D. N. DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. B.* **360**, 1847–1857 (2005).
19. Lee, E. *et al.* The identification of elephant ivory evidences of illegal trade with mitochondrial cytochrome *b* gene and hypervariable D-loop region. *J. Forensic Leg. Med.* **20**, 174–178 (2013).
20. Khedkar, T. *et al.* DNA barcoding using skin exuviate can improve identification and biodiversity studies of snakes. *Mitochondrial DNA A* **27**, 499–507 (2016).
21. Chang, C. H., Jang-Liaw, N. H., Lin, Y. S., Fang, Y. C. & Shao, K. T. Authentication of the dried seahorses in traditional Chinese medicine market in Taiwan by molecular forensics. *J. Food Drug Anal.* **21**, 310–316 (2013).
22. Bohmann, K. *et al.* Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol. Evol.* **29**, 358–367 (2014).
23. Ford, M. J. *et al.* Estimation of a killer whale (*Orcinus orca*) population's diet using sequencing analysis of DNA from feces. *PLoS ONE* **11**, e0144956. <https://doi.org/10.1371/journal.pone.0144956> (2016).
24. Moran, Z., Orth, D. J., Schmitt, J. D., Hallerman, E. M. & Aguilar, R. Effectiveness of DNA barcoding for identifying piscine prey items in stomach contents of piscivorous catfishes. *Environ. Biol. Fishes* **99**, 161–167 (2016).
25. Mayr, E. *Systematics and the Origin of Species, from the Viewpoint of a Zoologist* (Harvard University Press, 1942).
26. Krück, N. C. *et al.* Multi-gene barcoding to discriminate sibling species within a morphologically difficult fish genus (*Sillago*). *Fish. Res.* **143**, 39–46 (2013).
27. Melo, B. F., Ochoa, L. E., Vari, R. P. & Oliveira, C. Cryptic species in the Neotropical fish genus *Curimatopsis* (Teleostei, Characiformes). *Zool. Scr.* **45**, 650–658 (2016).
28. Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. Identification of Birds through DNA Barcodes. *PLoS Biol.* **2**, e312. <https://doi.org/10.1371/journal.pbio.0020312> (2004).
29. Feng, C. *et al.* Multiple convergent events created a nominal widespread species: *Triplophysa stoliczkae* (Steindachner, 1866) (Cobitoidea: Nemacheilidae). *BMC Evol. Biol.* **19**, 177. <https://doi.org/10.1186/s12862-019-1503-3> (2019).
30. Hofmann, T., Kneibelsberger, T., Kloppmann, M., Ulleweit, J. & Raupach, M. J. Egg identification of three economical important fish species using DNA barcoding in comparison to a morphological determination. *J. Appl. Ichthyol.* **33**, 925–932 (2017).
31. Gaston, K. J. Taxonomy of taxonomists. *Nature* **356**, 281–282 (1992).
32. Bacher, S. Still not enough taxonomists: Reply to Joppa *et al.* *Trends Ecol. Evol.* **27**, 65–66 (2011).
33. Bariche, M. *et al.* Red Sea fishes in the Mediterranean Sea: A preliminary investigation of a biological invasion using DNA barcoding. *J. Biogeogr.* **42**, 2363–2373 (2015).
34. Gonçalves, P. F. M., Oliveira-Marques, A. R., Matsumoto, T. E. & Miyaki, C. Y. DNA barcoding identifies illegal parrot trade. *J. Hered.* **106**, 560–564 (2015).
35. Ruppert, K. M., Kline, R. J. & Rahman, M. S. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* **17**, e00547. <https://doi.org/10.1016/j.gecco.2019.e00547> (2019).
36. Van der Heyde, M. *et al.* Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA (eDNA) metabarcoding. *Mol. Ecol. Resour.* **20**, 732–745 (2020).
37. Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. T. species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *PNAS* **101**, 14812–14817 (2004).
38. Kochzius, M. *et al.* Identifying fishes through DNA barcodes and microarrays. *PLoS ONE* **5**, e12620. <https://doi.org/10.1371/journal.pone.0012620> (2010).

39. Brandon-Mong, G. J. *et al.* DNA metabarcoding of insects and allies: An evaluation of primers and pipelines. *Bull. Entomol. Res.* **105**, 717–727 (2015).
40. Ward, R. D., Hanner, R. & Hebert, P. D. N. The campaign to DNA barcode all fishes, FISH-BOL. *J. Fish Biol.* **74**, 329–356 (2009).
41. Becker, S., Hanner, R. & Steinke, D. Five years of FISH-BOL: Brief status report. *Mitochondrial DNA* **22**, 3–9 (2011).
42. Ratnasingham, S. & Hebert, P. D. N. BOLD: The barcode of life data system. *Mol. Ecol. Notes* **7**, 355–364 (2007).
43. Keskin, E. & Atar, H. H. DNA barcoding commercially important fish species of Turkey. *Mol. Ecol. Resour.* **13**, 788–797 (2013).
44. Ali, F. S., Ismail, M. & Aly, W. DNA barcoding to characterize biodiversity of freshwater fishes of Egypt. *Mol. Biol. Rep.* **47**, 5865–5877 (2020).
45. Mirabdullaev, I. M. & Mullabaev, N. R. Ichthyofauna of Uzbekistan: Modern state and taxonomy. *J. Biol. Uzbekistan* **5**, 43–49 (2020) **(In Russian with English abstract)**.
46. Aladin, N. V. & Potts, W. T. W. Changes in the Aral Sea ecosystems during the period 1960–1990. *Hydrobiologia* **237**, 67–79 (1992).
47. Karthe, D., Chalov, S. & Borchardt, D. Water resources and their management in central Asia in the early twenty first century: status, challenges and future prospects. *Environ. Earth Sci.* **73**, 487–499 (2014).
48. Sheraliev, B. & Peng, Z. *Triplophysa ferganaensis*, a new loach species from Fergana Valley in Central Asia (Teleostei: Nemacheilidae). *J. Fish Biol.* <https://doi.org/10.1111/jfb.14764> (2021).
49. Sheraliev, B., Allayarov, S. & Peng, Z. DNA barcoding revealed a wider distribution of *Alburnoides holciki* (Teleostei: Leuciscidae) in the inland waters of Uzbekistan. *J. Appl. Ichthyol.* <https://doi.org/10.1111/jai.14206> (2021).
50. Sheraliev, B., Allayarov, S. & Peng, Z. First records of *Gobio nigrescens* and *Gobio sibiricus* (Cypriniformes: Gobionidae) from the Amu Darya River basin, Uzbekistan. *J. Appl. Ichthyol.* **36**, 235–239 (2020).
51. Kuchboev, A. *et al.* Molecular analysis of polymorphic species of the genus *Marshallagia* (Nematoda: Ostertagiinae). *Parasites Vectors* **13**, 411. <https://doi.org/10.1186/s13071-020-04265-1> (2020).
52. Čandek, K. & Kuntner, M. DNA barcoding gap: Reliable species identification over morphological and geographical scales. *Mol. Ecol. Resour.* **15**, 268–277 (2014).
53. Guimaraes-Costa, A. J. *et al.* Fish diversity of the largest deltaic formation in the Americas: A description of the fish fauna of the Parnaíba Delta using DNA Barcoding. *Sci. Rep.* **9**, 7530. <https://doi.org/10.1038/s41598-019-43930-z> (2019).
54. Matveyev, M. P., Levina, M. A., Mamilov, N. S. & Levin, B. A. Tashkent riffle minnow *Alburnoides oblongus* belongs to the genus *Alburnus* (Osteichthyes: Cyprinidae) as inferred from mitochondrial and nuclear DNA markers. *Inland Water Biol.* **10**, 360–367 (2017).
55. Jouladeh-Roudbar, A., Ghanavi, H. R. & Doadrio, I. Ichthyofauna from Iranian freshwater: Annotated checklist, diagnosis, taxonomy, distribution and conservation assessment. *Zool. Stud.* **59**, 21. <https://doi.org/10.6620/ZS.2020.59-21> (2020).
56. Froese, R. & Pauly, D. FishBase. *World Wide Web electronic publication*. www.fishbase.org, version (02/2021) (2021).
57. Atamuratova, M. S. *Ichthyofauna of reservoirs in the Akhangaran river basin*. (Ph.D. dissertation, Institute of Zoology, Tashkent, Uzbekistan, 2021) **(In Uzbek)**.
58. Mauroner, A. Central Asia. *IUCN Freshwater Fish Specialist Group*. <http://www.iucnffsg.org/about-ffsg/ffsg-regions/centralasia/> (2014).
59. Berg, L. S. *Fishes of Fresh Waters of the USSR and Adjacent Countries*. Vol. 2. (Academy of Sciences of the USSR Press, 1949) **(In Russian)**.
60. Turdakov, F. A. *Fishes of Kirghizia* (Academy of Sciences of the Kyrgyz SSR Press, 1963) **(In Russian)**.
61. Prokofiev, A. M. *Loaches of the subfamily Nemacheilinae of the World* (Filigran, 2017) **(In Russian)**.
62. Fricke, R., Eschmeyer, W. N. & van der Laan, R. *Eschmeyer's catalog of fishes: Genera, species, references*. <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp> (2021).
63. Ramon, M. L., Lobel, P. S. & Sorenson, M. D. Lack of mitochondrial genetic structure in hamlets (*Hypoplectrus* spp.): Recent speciation or ongoing hybridization?. *Mol. Ecol.* **12**, 2975–2980 (2003).
64. Kashiwagi, T., Marshall, A. D., Bennett, M. B. & Ovenden, J. R. The genetic signature of recent speciation in manta rays (*Manta alfredi* and *M. birostris*). *Mol. Phylogenet. Evol.* **64**, 212–218 (2012).
65. Decru, E. *et al.* Taxonomic challenges in freshwater fishes: A mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo basin. *Mol. Ecol. Resour.* **16**, 342–352 (2015).
66. Geiger, M. F. *et al.* Spatial heterogeneity in the Mediterranean Biodiversity Hotspot affects barcoding accuracy of its freshwater fishes. *Mol. Ecol. Resour.* **14**, 1210–1221 (2014).
67. Pourshababan, A. *et al.* Morpho-species of the genus *Leuciscus* Cuvier, 1816 (Teleostei: Leuciscinae) from Iran revisited using molecular approaches. *Zool. Middle East* **67**, 126–132 (2021).
68. Dillman, C. B. *et al.* Molecular systematics of Scaphirhynchinae: An assessment of North American and Central Asian Freshwater Sturgeon Species. *J. Appl. Ichthyol.* **23**, 290–296 (2007).
69. Nedoluzhko, A. V. *et al.* Molecular phylogeny of one extinct and two critically endangered Central Asian sturgeon species (genus *Pseudoscaphirhynchus*) based on their mitochondrial genomes. *Sci. Rep.* **10**, 722. <https://doi.org/10.1038/s41598-020-57581-y> (2020).
70. Berg, L. S. *The Freshwater Fishes of the USSR and Adjacent Countries* (Academy of Sciences of the USSR Press, 1948) **(In Russian)**.
71. Chen, W., Ma, X., Shen, Y., Mao, Y. & He, S. The fish diversity in the upper reaches of the Salween River, Nujiang River, revealed by DNA barcoding. *Sci. Rep.* **5**, 17437. <https://doi.org/10.1038/srep17437> (2015).
72. Shen, Y., Guan, L., Wang, D. & Gan, X. DNA barcoding and evaluation of genetic diversity in Cyprinidae fish in the midstream of the Yangtze River. *Ecol. Evol.* **6**, 2702–2713 (2016).
73. Crivelli, A. J. Are fish introductions a threat to endemic freshwater fishes in the northern Mediterranean region?. *Biol. Conserv.* **72**, 311–319 (1995).
74. Rainbow, P. Impacts of invasions by alien species. *J. Zool.* **246**, 247–248 (1998).
75. Lusk, S., Lusková, V. & Hanel, L. Alien fish species in the Czech Republic and their impact on the native fish fauna. *Folia Zool.* **59**, 57–72 (2010).
76. Balestrieri, A., Remonti, L., Vezza, P., Prigioni, C. & Copp, G. H. Do non-native fish as prey favour the conservation of the threatened indigenous Eurasian otter?. *Freshw. Biol.* **58**, 995–1007 (2013).
77. Tickner, D. *et al.* Bending the curve of global freshwater biodiversity loss: An emergency recovery plan. *Bioscience* **70**, 330–342 (2020).
78. Takahara, T., Minamoto, T. & Doi, H. Using environmental DNA to estimate the distribution of an Invasive fish species in ponds. *PLoS ONE* **8**, e56584. <https://doi.org/10.1371/journal.pone.0056584> (2013).
79. Muchlisin, Z. A. *et al.* Assessing the species composition of tropical eels (Anguillidae) in Aceh Waters, Indonesia, with DNA barcoding gene *cox1*. *F1000 Res.* **6**, 258. <https://doi.org/10.12688/f1000research.10715.1> (2017).
80. Ran, K., Li, Q., Qi, L., Li, W. & Kong, L. DNA barcoding for identification of marine gastropod species from Hainan Island, China. *Fish. Res.* **225**, 105504. <https://doi.org/10.1016/j.fishres.2020.105504> (2020).
81. Gilbey, J. *et al.* Life in a drop: Sampling environmental DNA for marine fishery management and ecosystem monitoring. *Mar. Policy* **124**, 104331. <https://doi.org/10.1016/j.marpol.2020.104331> (2021).
82. Sturgeon Specialist Group. *Acipenser nudiiventris* Aral Sea stock. *The IUCN Red List of Threatened Species 1996*, e.T251A13048150. <https://doi.org/10.2305/IUCN.UK.1996.RLTS.T251A13048150.en> (1996).

83. Mamilov, N. & Karimov, B. *Capoetobrama kuschakewitschi*. *The IUCN Red List of Threatened Species*, 2020, e.T169620A156732450. <https://doi.org/10.2305/IUCN.UK.2020-3.RLTS.T169620A156732450.en> (2020).
84. Harper, L. R. *et al.* Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. *Hydrobiologia* **826**, 25–41 (2019).
85. Sales, N. G. *et al.* Fishing for mammals: Landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from riverine systems. *J. Appl. Ecol.* **57**, 707–716 (2020).
86. Zou, K. *et al.* eDNA metabarcoding as a promising conservation tool for monitoring fish diversity in a coastal wetland of the Pearl River Estuary compared to bottom trawling. *Sci. Total Environ.* **702**, 134704. <https://doi.org/10.1016/j.scitotenv.2019.134704> (2020).
87. Mirabdullaev, I. M., Mirzaev, U. T., Kuzmetov, A. R. & Kimsanov, Z. O. *Keys to Fishes of Uzbekistan and Adjacent Regions* (Sano-Standard Press, 2011) (In Uzbek).
88. Kottelat, M. & Freyhof, J. *Handbook of European freshwater fishes*. (Kottelat, Cornol & Freyhof, 2007).
89. Ivanova, N. V., Zemlak, T. S., Hanner, R. H. & Hebert, P. D. N. Universal primer cocktails for fish DNA barcoding. *Mol. Ecol. Notes* **7**, 544–548 (2007).
90. Sullivan, J. P., Peng, Z., Lundberg, J. G., Peng, J. & He, S. Molecular evidence for diphyly of the Asian catfish family Amblycipitidae (Teleostei: Siluriformes) and exclusion of the South American Aspredinidae from Sisoroidea. *Proc. Acad. Nat. Sci. Philadelphia* **157**, 51–65 (2008).
91. Lassmann, T. & Sonnhammer, E. L. Kalign: An accurate and fast multiple sequence alignment algorithm. *BMC Bioinform.* **6**, 298. <https://doi.org/10.1186/1471-2105-6-298> (2005).
92. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120 (1980).
93. Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* **21**, 1864–1877 (2011).
94. Meier, R., Shiyang, K., Vaidya, G. & Ng, P. K. L. DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. *Syst. Biol.* **55**, 715–728 (2006).
95. Ronquist, F. *et al.* MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
96. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
97. Rambaut, A. FigTree v1.4.2, a graphical viewer of phylogenetic trees. <http://tree.bio.ed.ac.uk/software/figtree/> (2014).

Acknowledgements

We are thankful to Elbek Jalolov, Zuhridin Obloqulov, Akbarjon Rozimov, Sirojiddin Allayarov, and Bahridin Karimov for assistance in the field work. This work was funded by the grant from the National Natural Science Foundation of China (No. 31872204).

Author contributions

Z.P. initiated the project, acquired funding, and managed project administration. B.S. collected specimens, performed DNA extraction, analyzed the results and wrote the manuscript. Z.P. reviewed the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-96487-1>.

Correspondence and requests for materials should be addressed to Z.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021