



## Full Length Article

Assessment of genetic diversity of freshwater genus *Labeo* through DNA barcoding and phylogenetic analysis inhabited in the River IndusSaba Zafar<sup>a</sup>, Farhat Jabeen<sup>a,\*</sup>, Muhammad Ali<sup>b</sup>, Aiman Sajjad<sup>a</sup><sup>a</sup> Department of Zoology, Government College University Faisalabad, Pakistan<sup>b</sup> Bahauddin Zakariya University, Multan, Pakistan

## ARTICLE INFO

## Keywords:

Morphology and DNA barcoding  
Freshwater fish diversity  
Phylogenetic analysis  
Fish conservation  
Genus *Labeo*

## ABSTRACT

**Objectives:** Genus *Labeo* of family Cyprinidae comprising of important group of fish species are known for ornamental purpose, food, extracting oil, medicinal value and for potential viable aquaculture. The present study was conducted to assess the morphometric and molecular identification of freshwater fishes of genus *Labeo* including *Catla catla*, *Labeo rohita*, *Labeo calbasu*, *Labeo gnius*, *Labeo bata* from the three fish resources of River Indus, Punjab, Pakistan.

**Methods:** Out of 75, thirty fish specimen were amplified and sequenced with 650 bp of mitochondrial *COI* gene for the evaluation of fish biodiversity, management and conservation by using different bioinformatics techniques.

**Results:** Maximum and minimum genetic divergence between fishes of *C. catla*, *L. calbasu*, *L. boga*, *L. gonius* and *L. rohita*, were 0.052 to 0.105, respectively. Highest divergence was found between *Labeo bata* and *Labeo rohita*. Interspecific mean distance was found higher than intraspecific distance. In nucleotide base composition, AT contents was higher than GC contents indicated the clear pattern of anti-G bias. Transition rate was  $k_1 = 8.84$  for purines and transversional rate was  $k_2 = 15.47$  for pyrimidines with transition/transversion bias  $R = 5.62$ . Tajima D test and nucleotide diversity was measured  $D = 1.382883$  and  $\pi = 0.049981$  respectively. Positive value of D indicated a strong selective pressure and lack of population expansion. Thirty species were identified with e-value 0.0 and similar identity 99.85%–100% with maximum 600 bp length. Phylogenetic tree revealed two major clades with 95% high bootstrap values identity showed closet relationship among these species.

**Conclusions:** Current study supported the efficacy of the *COI* gene for the species identification in River Indus and develop the concept for establishing Pakistani fish gene bank which is still needed.

## 1. Introduction

River Indus is one of longest river system in Pakistan. It is recognized for its freshwater biodiversity, including fishes (Abro et al., 2020). In Pakistan, total 531 species have been recognized, of which 298 are marine and 233 are freshwater fish species. Economically 78 out of 233 have been recognized as “species of special importance” based on endemism (Ghouri et al., 2020). The knowledge about freshwater fish diversity of River Indus mainly focus on the identification for fish conservation and management.

The Cyprinidae family of freshwater fishes is the most varied, with 377 genera and more than 3038 known species (Alam et al., 2021).

Maximum species of family cyprinidae can be found in North America, Asia, Europe and Africa (Kaur et al., 2021). In addition to their great value in the recreational fisheries, some fish species of family cyrinidae are economically significant in aquarium trade, aquaculture, angling and serve as a main source of proteins and food for Pakistan (Adeoba et al., 2018). Among species of family cyprinidae, *Labeo bata*, *Catla catla*, *Labeo gonius*, *Labeo rohita* and *Labeo calbasu* are appropriate for fish culture in Pakistan. These fishes are belongs to genus *labeo* and only occur in South East Asia and Africa.

After the Barbiinae, the second most vital group of fishes is *Labeo* lineage of family Cyprinidae. Cyprinid species represent 19.6% occurrence in Asia (Sarma et al., 2017). Genus *Labeo* is the most dominating

**Abbreviations:** COI, cytochrome oxidase I; PCR, polymerase chain reaction; ABGD, automated barcode gap discovery analysis; K2P, kimura 2 parameter; PrPecL, pre-pectoral length; PrPevL, pre-pelvic length.

\* Corresponding author.

E-mail address: [farhatjabeen@gcuf.edu.pk](mailto:farhatjabeen@gcuf.edu.pk) (F. Jabeen).

<https://doi.org/10.1016/j.jksus.2024.103123>

Received 6 September 2023; Received in revised form 24 January 2024; Accepted 3 February 2024

Available online 5 February 2024

1018-3647/© 2024 The Author(s). Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

group of family Cyprinidae. Out of 105 fish species of *Labeo* genus around the world, 36 fish species of *Labeo* genus are represented throughout the South and South East Asia and 69 occur in Africa (Sudha et al., 2015).

Fish is an essential aquatic organisms having high nutritional and protein value for human consumption (Selamoglu, 2018; Kakoolaki et al., 2013). *Labeo rohita* is favored for its public preference and taste. This is a candidate specie for aquaculture practices in Pakistan, Myanmar, Bangladesh, and India. Researchers have created genetically modified rohu for critical qualities including growth and disease resistance due to consumer interest in this species (Rasal and Sundaray, 2020). *Catla catla* is considered an ideal dietetic food and an important source of calcium, protein and low fatty acid. Therefore, in composition of animal taxa, fish constitute an essential component of biodiversity and have economical values and are significant source of protein for humans (Hossain et al., 2015).

However, the River Indus's diversity and abundance of these freshwater fishes become more susceptible due to human involvement. Annually almost 400 species have been defined over past 20 years (Fricke et al., 2021). Anthropogenic effects such as habitat degradation, overfishing, water pollution from plastics, poaching, ports, pollution of water, noise, household wastage and water withdrawal have resulting the loss of fish diversity. Comprehensively, taxonomic study of fish species are key to conserving and managing biodiversity (Sheraliev and Peng, 2021; Selamoglu, 2021a,b). Freshwater fishes of the genus *Labeo* are traditionally identified through their visual morphology characteristics such as, features, shape, size, number of fin rays and different body measurements. This system is quite valuable for gathering details about an organism's shape.

Classifications and identifications, moreover, can be time consuming and challenging based on the morphological features; although species documentations usually require significant taxonomic data (Bilgin et al., 2014). Thus, accurate specimen identification at the species level is necessary for environmental studies and conservation. The inherent boundaries of traditional taxonomy and minimum number of taxonomists need genetic techniques for fish species identifications (Sheraliev and Peng, 2021).

The most used molecular identification technique is DNA barcoding, which is based on the mitochondrial *COI* gene. This technique is effective for identification at the level of larvae that can't be identified through traditional morphological approaches. This technique has potential to reveal the misidentification of species that happened by means of traditional identification ways (Jaonaison et al., 2022).

This region received certification at the first international conference on DNA Barcode Life (Trivediet et al., 2016). Different studies revealed that the sequences diversity in the range of approximately 650 bp from the 5' area of the *COI* provides great resolution for various animal groups at species level (Cermakova et al., 2023). To identify the eggs, larvae and broken specimens by traditional method cannot be done properly but the DNA barcoding method makes it easier and authentic (Jaonaison et al., 2022). Fishes are the most diverse and largest vertebrate class as well as easiest group to create DNA barcode records. By the use of DNA barcode method, it is estimated that species number was exceeding at the rate of 300 new species per year (Weigt et al., 2012).

The efficacy of this method has been verified by various studies, it is possible to correctly identify more than 90 % species at the species. It can expose cryptic diversity and also identify fish specimens when morphometric characters cannot, like treated commercial products. Furthermore, DNA barcoding suggests the occurrence of species in any given area by evaluating ambient DNA. This method is known as DNA meta-barcoding (Tsoupas et al., 2022).

Environmental factors and human exploitation have created significant impacts on structure and biodiversity of fish community in River Indus. Decline in genetic variations in any population reduces fish capability to adapt the environmental variations and also decrease the species chances of long term survival (Tickner et al., 2020). However, to

support sustainability, improved control and management of fishes should be applied. Still, one of the most crucial and fundamental issues in fisheries management is the identification of species. The goal of current study was to check the biodiversity of commercially important fishes of *Labeo* genus for conservation from three different populations of River Indus Punjab, Pakistan and to provide new information on its genetic diversity by analyzing additional samples from previously studied areas as well as by studying new and unexplored population. In this study, DNA barcoding records will be available to investigators and researchers to conserve and monitor the biodiversity of fishes in these studied areas.

## 2. Materials and methods

### 2.1. Ethic statement

Fish specimens utilized in this study were treated and handled after the approval from Ethics Review Committee of the Government College University Faisalabad, Pakistan (Ref. No. GCUF/ERC/19/017 dated 04-2-2019).

### 2.2. Study area

The current research was carried out from the three different sites Mianwali (upstream), kallurKot(midstream) and Dera Ghazi Ghaat (downstream) of riverine system of River Indus, Punjab, Pakistan in October 2020 to February 2022. Basic data, such as the location's latitude and longitude, were noted at each site (Fig. 1).

### 2.3. Fish sampling

Total 75 specimens of fish species were collected from three different sampling areas that cover the main course of River Indus, Punjab, Pakistan. The fish samples were collected randomly with casting net, gill nets, and baskets with the help of fisherman. Samples were labelled and packed in polyethylene bags. Then fish samples were photographed and preserved as voucher specimen to fisheries laboratory, GC University Faisalabad. For DNA barcoding, at least two individual per common species according to available number of each species were preserved in alcohol.

### 2.4. Morphological analysis

Morphological identification of fish samples was done with help of morphological characters and standard taxonomic keys provided by Punjab Fisheries Department (Mirza and Sandhu, 2007). The morphometric measurements, standard length, total length, forked length, head length, body depth, snout length, pre-pectoral length, pre-pelvic length, pre-dorsal length and pre-anal length of *Labeo calbasu*, *Labeo gonius*, *Labeo rohita*, *Labeo bata* and *Catla catla* were measured in (mm) with help of measuring board, measuring tape, vernier caliper and divider (Table 1).

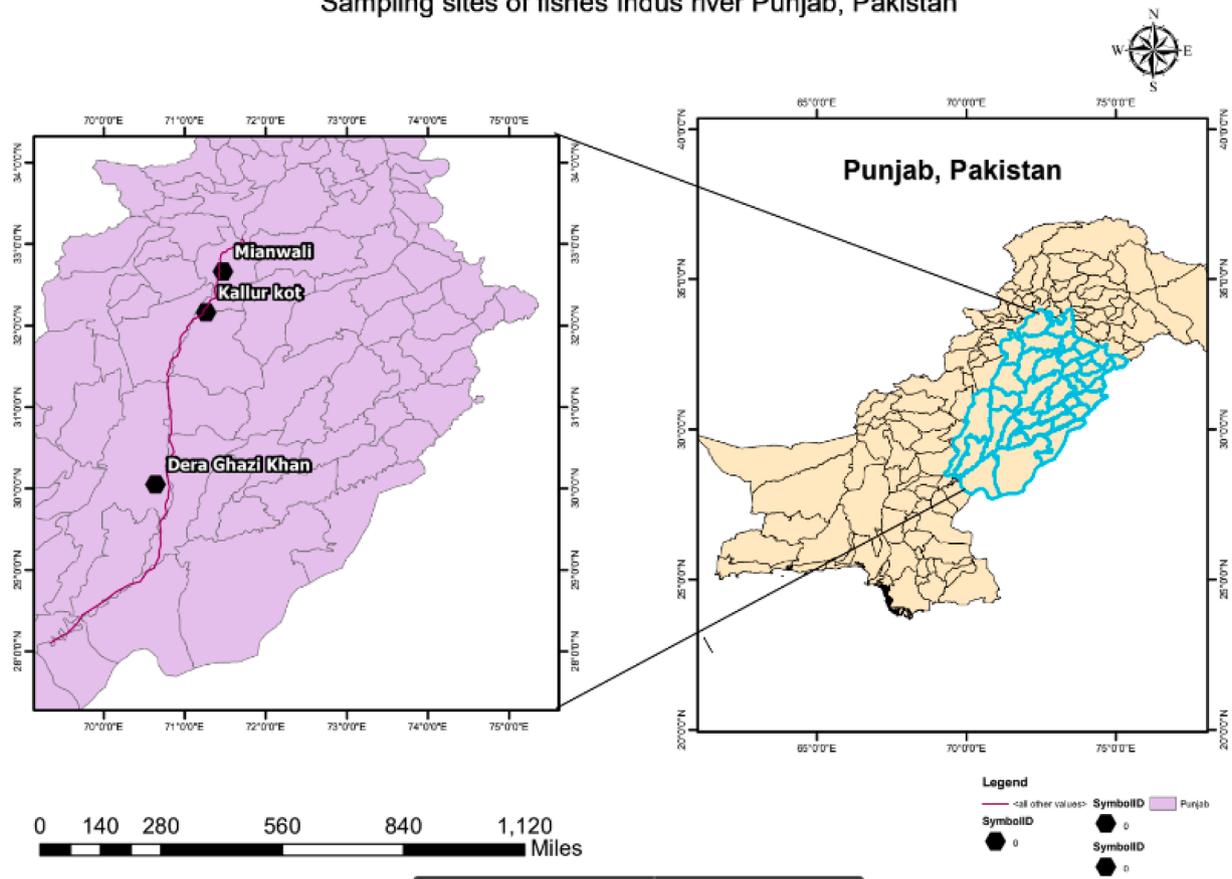
### 2.5. DNA extraction

After morphological measurements, fin tissue of each specimens were dissected and then preserved at  $-20^{\circ}\text{C}$  in 95 % ethanol. Fin tissues of voucher specimens were cut with the help of sterilized scissor and total genomic DNA was extracted according to manufacturer's protocol of Tissue Kit method (QIAamp™ DNA Extraction Mini Kit). DNA extraction consists of three steps: lyses, precipitation, and purification of DNA.

### 2.6. DNA visualization

After the DNA extraction, purity and quality of DNA specimens were

## Sampling sites of fishes Indus river Punjab, Pakistan



**Fig. 1.** Map of Sampling sites. Map showing three different locations (Mianwali, KallurKot, Dera Ghazi Khan) for collecting five species of labeo genus. River Indus are marked with purple and sites from which samples were collected are marked with black symbol. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**

Descriptive data of samples collected from three sampling sites of River Indus, Punjab.

Family	Genus	Species	Common name	Commercial value	Sample size	Distribution	IUCN status
Cyprinidae	Labeo	<i>Labeo rohita</i>	Rahu	Very high	15	AJK, Baluchistan, Punjab, Sindh	LC
		<i>Labeo calbasu</i>	Calbans, Dahi	High	15	KP, Punjab, Sindh	LC
		<i>Labeo gonius</i>	Sereha	High	15	Punjab, Sindh	LC
		<i>Labeo bata</i>	Bata	High	15	KP, Punjab, Sindh	LC
		<i>Catla catla</i>	Thaila	High	15	AJK, Baluchistan, Punjab, Sindh	LC

checked with NanoDrop values (ND- 1000 spectrophotometer). Samples were viewed on 0.8 % agarose gel to prevent degradation possibility and set final concentration to 50 ng/ml to prevent PCR inhibition because of excessive quantity of template DNA.

## 2.7. PCR amplification

The polymerase chain reaction amplification process was then utilized with DNA samples. The universal primers FishF1 and FishR1 were used to amplify a nearly 650 bp sequence of cytochrome c oxidase subunit I from the 5'.

FishF1: 5'/TCAACCAACCACAAAGACATTGGCAC3/

Fish R1: 5'/TAGACTTCTGGGTGGCCAAAGAATCA3/

Mitochondrial COI marker was selected because of its well-known sequence, quick primer construction and higher interspecies variability than intraspecies. The mitochondrial gene COI is one of the most useful markers for wide range of fish species identification. Its 650 bp segment serves as the foundation for taxonomic fish differentiation using DNA barcoding (Cermakova et al., 2023). The Bio-Rad I-Cycler™

Thermo cycler (Bio-Rad Inc., Waltham, MA) was used to carry out the PCR amplification reaction. The COI gene is amplified by initial denaturation at 95 °C for 3.00 min, 36 cycles at 95 °C for 0.30 s, annealing at 58 °C for 0.45 s, and extension at 72 °C for 0.45 s. A final expansion lasting 10 min at 72 °C was noticed before being recorded at 4 °C. Prior to sanger sequencing, amplified PCR products were sized and observed on a 0.8 % agarose gel, and the COI gene containing products were chosen.

## 2.8. Mitochondrial COI region sequencing

Polymerase chain reaction Using a genetic analyzer (ABI PRISM 3730XL Analyzer 96 capillary type), sequencing was carried out. The Big Dye (R) Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) was used for the cycle sequencing. Fish species from all habitats were used in the current investigation, and mitochondrial DNA was used to amplify and sequence the COI gene. These sequences were analyzed to determine genetic diversity and evolutionary history of the *L. rohita*, *L. calbasu*, *L. gonius*, *L. bata*, and *C. catla* species.

### 3. Statistical analysis

Morphological data was statistically analyzed measuring mean and standard deviation by plotting data on Excel sheet. Sequence with their accession number were submitted to GeneBank Barcode database. Consensus sequence of studied fishes were generated and submitted in NCBI than these sequence along with voucher sequence from GenBank and reference sequence from BOLD were aligned and compared by using ClusterW program of MEGA. 11. Accession number of studied fishes and reference sequence number were used to generate multiple sequence alignment for *Labeo rohita*, *Labeo bata*, *Labeo gonius*, *Labeo calbasu* and *Catla catla*.

Using MEGA11.0, pairwise genetic distances were calculated for sequence comparisons based on the Kimura 2-parameter (K2P) distance model. The species of *L. rohita*, *L. calbasu*, *L. gonius*, *L. bata*, and *C. catla* that were examined here are represented graphically by a UPGMA tree. For intraspecific and interspecific divergence, automated barcode gap discovery (ABGD) was used with K2P model with transition and transversion ratio (TS/TV) set to 1.5, relative gap width (x) = 1.0, recursive steps 10, and remaining parameters set to default value with pmax = 0.1, pmin = 0.001 and Nb bins = 20 ABGD (<https://bioinfo.mnhn.fr/ab/public/abgd/>). QR code generator (aspose.app/barcode/generate) was used to generate unique codes when species gets separated.

### 4. Results

#### 4.1. Morphometric identification

In *Labeo rohita*, back was black and sides became brownish when preserved in formalin. Fins were dark gray and iris rims was red. The edge of each scale has a dark brown tinge. In *Catla catla*, dorsal side (back) was dirty greenish and ventral side was silvery brown. Fins color was black with light base color. *Labeo bata* has a narrow mouth with no lateral lobe and has two minute maxillary barbels. Body color was dark gray to orange and fins with reddish tinge. The body color of *Labeo calbasu* was blackish with lighter on ventral side. It has four pairs of barbels with very minute gill rakers. The mouth was narrow with depressed snout and has thick lips. In *Labeo gonius*, body color was reddish golden to orange with convex dorsal profile than abdomen. The mouth is narrow with short barbels. Fins color was reddish brown and pectoral fin was as long as head.

Total 75 fishes of *Labeo rohita*, *Catla catla*, *Labeo bata*, *Labeo calbasu* and *Labeo gonius* were collected from three different sites Mianwali, KallurKot and Dera Ghazi Ghat of River Indus, Puanjab, Pakistan (Fig. 1). At taxonomic level, samples were morphological identified by measuring total length (TL), standard length (SL), fork length (FL), body depth (BD), head length (HL), snout length (SnL), pre-dorsal length (PrDL), pre-pectoral length (PrPecL), pre-pelvic length (PrPevL) and pre-anal length (PrAL). Within species maximum mean value of TL, SL, FL, BD, HL, SnL, PrDL, PrPecL, PrPevL and PrAL was observed in *Catla catla*, *Labeo calbasu*, *Labeo gonius* and *Labeo rohita* from Mianwali populations which shows statistically significant difference ( $p \leq 0.05$ ) when compared with the mean values of Kallur Kot and Dera Ghazi Ghat fish populations (Table 2). Maximum mean values of morphometric characters were observed in KallurKot populations of *Labeo bata* except head length which has maximum mean value among fishes of Dera Ghazi Ghat population. Among all fish populations maximum mean values were observed in *L. rohita* which showed statistically significance difference when compared with other species of genus labeo (Table 2).

#### 4.2. DNA barcoding for fish species identification

Total 30 mitochondrial *COI* barcodes sequence were successfully obtained and analyzed from fishes of family cyprinidae and amplified using two primers in PCR (Table 3). All sequences were 650 bp in length after aditing and trimming, and sequences had no stop codons,

**Table 2** Morphological traits (Mean  $\pm$  SD) of five species of labeo genus collected from three different sites of River Indus, Punjab, Pakistan.

Morphometric characters	<i>Catla catla</i>			<i>Labeo calbasu</i>			<i>Labeo gonius</i>			<i>Labeo bata</i>			<i>Labeo rohita</i>		
	MW	KK	DGG	MW	KK	DGG	MW	KK	DGG	MW	KK	DGG	MW	KK	DGG
TL	215.33 $\pm$ 6.51 <sup>a</sup>	105.02 $\pm$ 1.77	95.19 $\pm$ 0.78	223.33 $\pm$ 9.5 <sup>a</sup>	183.03 $\pm$ 6.38	160.79 $\pm$ 3.23	172.08 $\pm$ 2.12 <sup>a</sup>	162.57 $\pm$ 4.72	154.61 $\pm$ 1.71	224.77 $\pm$ 12.59	278.45 $\pm$ 2.0 <sup>a</sup>	241.1 $\pm$ 1.65	286.66 $\pm$ 10.97 <sup>a</sup>	255.1 $\pm$ 7.45	226.81 $\pm$ 6.01
SL	167.3 $\pm$ 6.35 <sup>a</sup>	85.00 $\pm$ 5.69	73.03 $\pm$ 0.53	173.7 $\pm$ 12.85 <sup>a</sup>	139.01 $\pm$ 3.77	132.28 $\pm$ 1.2	135.47 $\pm$ 2.53 <sup>a</sup>	128.05 $\pm$ 3.27	120 $\pm$ 2.44	170 $\pm$ 5	215.67 $\pm$ 5.86 <sup>a</sup>	189 $\pm$ 3	218.93 $\pm$ 12.28 <sup>a</sup>	191.72 $\pm$ 6.86	179.78 $\pm$ 3.33
FL	186.66 $\pm$ 6.0 <sup>a</sup>	91.83 $\pm$ 3.13	78.07 $\pm$ 0.72	195.45 $\pm$ 12.2 <sup>a</sup>	158.13 $\pm$ 5.68	145.58 $\pm$ 2.23	150.23 $\pm$ 3.42 <sup>a</sup>	139.33 $\pm$ 4.04	130.9 $\pm$ 1.83	187.74 $\pm$ 3.26	238.68 $\pm$ 12.72 <sup>a</sup>	207.92 $\pm$ 2.88	249.6 $\pm$ 15.75 <sup>a</sup>	225.5 $\pm$ 6.01	203.09 $\pm$ 6.97
BD	54.87 $\pm$ 4.22 <sup>a</sup>	20.47 $\pm$ 0.74	20.5 $\pm$ 0.67	58.89 $\pm$ 6.05 <sup>a</sup>	45.27 $\pm$ 1.43	46.74 $\pm$ 0.5	40.66 $\pm$ 2.71 <sup>a</sup>	35.9 $\pm$ 1.96	37.17 $\pm$ 0.3	48.07 $\pm$ 2.54	51.28 $\pm$ 1.36 <sup>a</sup>	50.73 $\pm$ 1.17	68.77 $\pm$ 4.06 <sup>a</sup>	61.47 $\pm$ 1.75	57.02 $\pm$ 2.42
HL	51.54 $\pm$ 2.99 <sup>a</sup>	28.70 $\pm$ 2.13	26.41 $\pm$ 0.39	41.24 $\pm$ 2.73 <sup>a</sup>	35.19 $\pm$ 1.36	29.87 $\pm$ 1.84	29.56 $\pm$ 1.68 <sup>a</sup>	27.42 $\pm$ 0.51	27.56 $\pm$ 0.07	38.05 $\pm$ 3.35	41.02 $\pm$ 11.48	42.77 $\pm$ 0.87 <sup>a</sup>	62.77 $\pm$ 5.22 <sup>a</sup>	56.88 $\pm$ 3.69	49.97 $\pm$ 1.18
SnL	23.11 $\pm$ 0.68 <sup>a</sup>	11.26 $\pm$ 0.86	10.06 $\pm$ 0.07	20.62 $\pm$ 0.87 <sup>a</sup>	19.7 $\pm$ 0.42	17.28 $\pm$ 0.78	13.15 $\pm$ 1.04 <sup>a</sup>	11.71 $\pm$ 1.35	13.19 $\pm$ 0.17 <sup>a</sup>	17.87 $\pm$ 2.96	20.97 $\pm$ 91.36 $\pm$ 0.65 <sup>a</sup>	20.31 $\pm$ 2.22 <sup>a</sup>	29.23 $\pm$ 115.24 $\pm$ 0.46	26.47 $\pm$ 102.23 $\pm$ 0.77	23.74 $\pm$ 91.55 $\pm$ 0.46
PrDL	80.3 $\pm$ 2.36 <sup>a</sup>	36.19 $\pm$ 0.99	35.46 $\pm$ 0.42	85.05 $\pm$ 1.71 <sup>a</sup>	74.18 $\pm$ 3.27	66.15 $\pm$ 2.69	61.7 $\pm$ 2.3 <sup>a</sup>	55.89 $\pm$ 2.72	51.76 $\pm$ 0.92	75.35 $\pm$ 3.82	91.36 $\pm$ 0.31 <sup>a</sup>	82.79 $\pm$ 3.46 <sup>a</sup>	115.24 $\pm$ 62.17 $\pm$ 2.22	102.23 $\pm$ 55 $\pm$ 1.57	91.55 $\pm$ 7.07
PrPecL	54.23 $\pm$ 1.22 <sup>a</sup>	26.23 $\pm$ 1.74	23.76 $\pm$ 0.27	43.08 $\pm$ 1.33 <sup>a</sup>	35.48 $\pm$ 2.3	28.79 $\pm$ 2.24	30.59 $\pm$ 2.04 <sup>a</sup>	28.24 $\pm$ 1.54	29.65 $\pm$ 0.4	40.05 $\pm$ 2.09	44.37 $\pm$ 2.46 <sup>a</sup>	42.55 $\pm$ 2.91 <sup>a</sup>	62.17 $\pm$ 118.29 $\pm$ 0.64	55 $\pm$ 103.5 $\pm$ 2.86	47.64 $\pm$ 2.86
PrPevL	91.67 $\pm$ 2.62 <sup>a</sup>	38.87 $\pm$ 2.45	35.61 $\pm$ 0.09	83.44 $\pm$ 3.13 <sup>a</sup>	76.71 $\pm$ 1.9	64.51 $\pm$ 3.47	68.47 $\pm$ 0.15 <sup>a</sup>	65.7 $\pm$ 0.25	65.1 $\pm$ 0.17	84.25 $\pm$ 7.64	97.87 $\pm$ 0.42 <sup>a</sup>	94.03 $\pm$ 0.9	118.29 $\pm$ 4.34 <sup>a</sup>	103.5 $\pm$ 1.99	91.68 $\pm$ 4.76
PrAL	136.82 $\pm$ 2.1 <sup>a</sup>	58.36 $\pm$ 1.51	55.87 $\pm$ 0.29	124.62 $\pm$ 3.66 <sup>a</sup>	111.23 $\pm$ 2.67	96.44 $\pm$ 2.42	105.69 $\pm$ 0.75 <sup>a</sup>	98 $\pm$ 4.07	94.8 $\pm$ 0.38	135.24 $\pm$ 7.14	158.35 $\pm$ 2.08 <sup>a</sup>	147.27 $\pm$ 0.72	184.69 $\pm$ 2.33 <sup>a</sup>	147.58 $\pm$ 1.23	133.23 $\pm$ 5.82

**Table 3**Accession No. and BLAST results of *C. catla*, *L. bata*, *L. calbasu*, *L. gonius*, *L. rohita*.

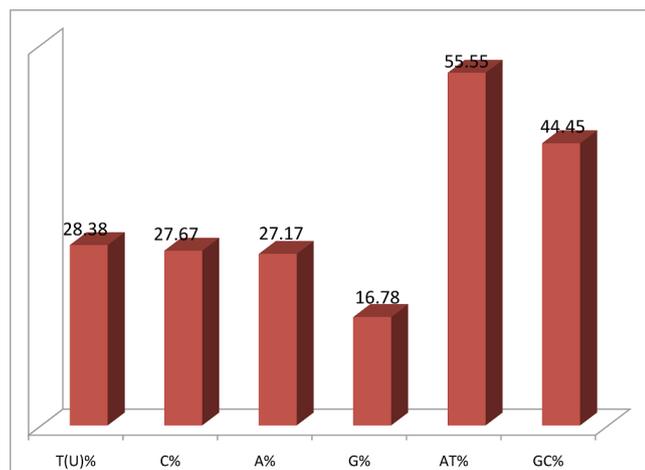
Query sequence ID	Accession No.	Query Length	Percentage similarity	E-value	Accession No. of best match	Query cover
<i>Catla catla</i>	OP536010	623	99.84 %	0.0	KC757310.1	100 %
<i>Catla catla</i>	OP536108	401	100 %	0.0	MK416162.1	100 %
<i>Catla catla</i>	OP536113	401	100 %	0.0	KX946601.1	100 %
<i>Catla catla</i>	OP536142	657	99.85 %	0.0	JX074196.1	100 %
<i>Catla catla</i>	OP536585	630	100 %	0.0	JX083155.1	98 %
<i>Catla catla</i>	OP536587	612	100 %	0.0	KJ406530.1	100 %
<i>Labeo bata</i>	OP737758	629	100 %	0.0	MT492421.1	100 %
<i>Labeo bata</i>	OP737886	642	100 %	0.0	OP737886.1	100 %
<i>Labeo bata</i>	OP737975	628	100 %	0.0	MN083094.1	100 %
<i>Labeo bata</i>	OP738086	627	100 %	0.0	OP738086.1	100 %
<i>Labeo bata</i>	OP738176	624	100 %	0.0	KT353105.1	100 %
<i>Labeo bata</i>	OP738298	614	99.84 %	0.0	MT492421.1	100 %
<i>Labeo calbasu</i>	OP709361	590	100 %	0.0	MN395836.1	100 %
<i>Labeo calbasu</i>	OP709370	615	100 %	0.0	MN178274.1	100 %
<i>Labeo calbasu</i>	OP710249	633	100 %	0.0	OP710249.1	100 %
<i>Labeo calbasu</i>	OP711819	628	100 %	0.0	KT364767.1	100 %
<i>Labeo calbasu</i>	OP711820	589	100 %	0.0	KX946689.1	100 %
<i>Labeo calbasu</i>	OP712184	600	100 %	0.0	MK238493.1	99 %
<i>Labeo gonius</i>	OP558489	626	100 %	0.0	OP558489.1	100 %
<i>Labeo gonius</i>	OP558529	610	100 %	0.0	MN200474.1	100 %
<i>Labeo gonius</i>	OP558530	628	99.84	0.0	MK572282.1	100 %
<i>Labeo gonius</i>	OP679801	576	100 %	0.0	OP679801.1	100 %
<i>Labeo gonius</i>	OP680014	605	100 %	0.0	HQ645092.2	100 %
<i>Labeo gonius</i>	OP680528	634	100 %	0.0	KC757287.1	99 %
<i>Labeo rohita</i>	OP550321	606	100 %	0.0	MN605503.2	100 %
<i>Labeo rohita</i>	OP550421	622	100 %	0.0	MK732922.1	100 %
<i>Labeo rohita</i>	OP550423	620	100 %	0.0	OP550423.1	100 %
<i>Labeo rohita</i>	OP550426	481	100 %	0.0	MN964133.1	100 %
<i>Labeo rohita</i>	OP550442	609	100 %	0.0	MN964135.1	100 %
<i>Labeo rohita</i>	OP550553	635	100 %	0.0	MK572283.1	100 %

insertions and deletions. The length of almost all analyzed sequences were greater than 600 bp.

#### 4.3. Nucleotide base composition analysis

The average nucleotides base frequencies of entire dataset observed were C (27.67 %), T (28.38 %), A (27.17 %) and G (16.78 %). The nucleotide base composition analysis of the *COI* gene sequence showed that the average G content was lowest and the highest T content, indicating that the AT content (55.55 %) was higher than the GC content (44.45 %) (Fig. 2) exhibiting a anti-G bias pattern. The rate of transition/transversion ratios were  $k_1 = 8.84$  and  $k_2 = 15.47$ .

The overall transition/transversion bias was  $R = 5.62$ , where  $R = [A^*G*k_1 + T^*C*k_2]/[(A + G)*(T + C)]$ . These biases of *COI* gene



**Fig. 2.** Average nucleotide bases, AT content and GC content percentage of cytochrome oxidase 1 (COX1) genes: T = thymine, A = adenine, C = cytosine, G = guanine.

substitution were measured by using a K2P model (Kimura 1980). Tajima neutrality test statistic and nucleotide diversity was also measured as Tajima  $D = 1.382883$  and  $\pi = 0.049981$  respectively. The transitional substitution ratio was shown in bold while transversional substitution ratio was shown in italics. The transitional substitution ratios from A to G, T to C, C to T and G to A were 16.49, 29.22, 28.30 and 10.43 respectively. While the transversional substitution ratios from A to T, A to C, T to A, T to G, C to A, C to G, G to T and G to C were 2.10, 2.10, 2.21, 2.21, 2.14, 2.14, 1.33 and 1.33, respectively (Table 4).

Nucleotide base pair frequency analysis of entire data set evaluated 571 out of 681 conserved sites, 99 out of 681 variable sites, 91 out of 681 parsimony sites and 7 of 681 singleton sites were observed. In this study genetic distance within and between fishes of *Catla catla*, *Labeo calbasu*, *Labeo bata*, *Labeo gonius* and *Labeo rohita* were measured using the Kimura 2-P model. The genetic distance within species were found to be 0 for all fishes. While the mean distance between species ranged from 0.052 to 0.105. The genetic variations between species is greater than mean distance within species which is also less than 1 %. The overall mean distance among all fishes of *L. rohita*, *C. catla*, *L. calbasu*, *L. bata*, and *L. gonius* was calculated as 0.07.

The data also shows that K2P distance increased with taxonomic level but above species level were smaller. In this study genetic distance were found to be less between *C. catla*, *L. calbasu*, *L. gonius*, *L. bata*, and *L. rohita*. For species level identification, a total of 30 *COI* barcode sequences from 5 distinct species were used. Identification of species were

**Table 4**

Maximum Likelihood Estimate of Substitution Matrix.

	A	T	C	G
A	–	<i>2.21</i>	<i>2.14</i>	<b>10.42</b>
T/U	<i>2.10</i>	–	<b>28.30</b>	<i>1.33</i>
C	<i>2.10</i>	<b>29.22</b>	–	<i>1.33</i>
G	<b>16.49</b>	<i>2.21</i>	<i>2.14</i>	–

Different transitional substitutions rates are shown in bold and transversional substitutions rates are shown in italics.

done by using GenBank and BOLD database. The GenBank and BOLD based identification of all specimens were ranged from 98.9 to 100 %. While the GenBank based identification of studied fishes of genus *Labeo* ranged from 99.85 to 100 %. The evolutionary divergence estimation between *COI* gene sequences at species level were determined by using K2P model. The lowest and highest values of evolutionary divergence were found between 0.04 and 0.11. The highest estimates was observed in *Labeo rohita* (0.110) (Table 5).

#### 4.4. Barcode gap analysis

Automated barcode gap discovery analysis was used for species identification and delimitation. Morphological identified fish species were accurately delimited between 7 and 9 putative species partition with prior maximal intraspecific divergence = 0.001668 and  $p = 0.001000$ . Each delimited species were classified into a distinct group (Table 6). Although four fish species with accession number (*Catla catla* OP536010, *Labeo bata* OP738298, *Labeo calbasu* OP712184, *Labeo rohita* OP550423) were assigned into separate groups with single sequence. Every other groups of each species contained sequences of the same species.

#### 4.5. Phylogenetic analysis

Phylogenetic tree was constructed to establish the relationships among species. UPGMA Tree was constructed among 30 species belongs to family cyprinidae with all haplotypes and 15 species from NCBI used as reference sequence which is provided in Fig. 3. The branch length of Phylogenetic reconstruction was 0.02 of with the bootstrap test 1000 replicates in which associated sequence clustered together. This phylogenetic analysis divided cyprinidae fishes into five major clades that shown a complex relationship among fishes.

Specimens of the same species clustered under the same node while the different species formed separate nodes. Species of *L. gonius* along with reference sequence form one cluster upside-down, *L. calbasu* form second cluster along with reference sequence, *L. rohita* form third cluster along with their reference sequence, *C. catla* form fourth cluster and above out group, *L. bata* form separate cluster along with their reference sequence. These nodes were supported by bootstrap values lies between 59 % and 98 %. *Salmostoma bacalia* belongs to a separate genus was used as outgroup to understand the evolutionary relationship among fishes of genus *Labeo*. The primary goal of the barcode analysis was to distinguish between the various species, even those that exhibit a high degree of taxonomic similarity. Barcode approach revealed the accurate and authentic identification of the examined species.

#### 4.6. DNA based QR code generation

In the present study, *COI*-based sequences of studied *Labeo* species were encoded in a QR code generator (aspose.app/barcode/generate) for the first time to generate QR codes for precise species identification. Each specimen has an own QR code that may be scanned using a mobile application (Fig. 4). For accurate and quick identification of fish in the *Labeo* genus, a DNA based QR code is produced. These barcode images make it possible for technical and scientific applications to use a QR code reader to retrieve information.

### 5. Discussion

Total 75 fish species were recorded from the three populations of the River Indus Punjab, Pakistan and these belongs to family Cyprinidae. Due to high protein contents, fish is an essential organisms of aquatic ecosystem (Gulhan and Selamoglu, 2016). Fishes shows high degree of variations within and among fish population than other vertebrate and are more liable to morphological variations induced by local habitats (Biswas et al., 2018). Such variation particularly occurs due to

separation of small portions of a population within habitat may cause notable genetic and phenotypic differentiation among populations within species (Turan et al., 2004). The abundance of fishes has been dropped due to alterations in the ecosystem of rivers and overharvesting which is caused by anthropogenic factors. Fish species that are specific and native to a particular region could improve both native species conservation and production (Sheraliev and Peng, 2021). The assessment of fish stock conformation is a useful tool for conserving and managing natural group's population.

Different morphometric characters were also measured for *Labeo* genus fishes. Morphometric differences are expected within and among species because fishes can adapt quickly themselves by changing required morphometric due to environmental changes. When the average measurements of *Labeo rohita* from the three population sites was compared with other four species (Table 2), it was found that *L. rohita* (MW) has highest TL, SL, FL, BD, HL, SnL, PrDL, PrPecL, PrPevL, PrAL than other four species. When the mean values within species was compared which found highest in Mianwali population except *Labeo bata* which was found higher in KallurKot population. It is difficult to determine and explain the reasons of morphological variations among different fish population. It has been recommended that morphology of fish species are determined by genetics, environment and interaction among them. Meanwhile, a fish shows higher plasticity in morphometric parameters to environmental fluctuations (Barman and Sharma, 2017).

Identification of animal species through DNA barcoding is a methodological technique based on a specific gene of mtDNA helps in the accurate identification and discrimination of unknown species (Elsaied et al., 2021). Molecular barcodes approach are useful for overlapped species discrimination, unidentified fish species identification and also for comparing with traditional taxonomy to define species limitations (Elsaied et al., 2021). For species identification, genetic barcodes method with cytochrome oxidase I gene has been well documented and acknowledged particularly in fishery. Approximately 650 bp of Cytochrome oxidase I gene (*COI*) is considered to be the most effective - barcoding gene. Because of various reasons such as hypermutation, absence of introns, maternal inheritance, lack of fast nucleotide substitution, slow rate of variations in amino acid sequence of *COI* gene as compare to other mtDNA genome (Cermakova et al., 2023). Through this approach, we can detect endangered and invasive species. Barcoding was helpful in phylogenetic analysis and identifying the diversity of the cryptic species (Ghouri et al., 2020). Almost 31,000 fish species are presents globally, out of which 25 % species have been administered experimentally with at least one fish species has been barcode from each of all 89 % fish families (Bhattacharya et al., 2016).

In the current study, five freshwater fish species of genus *Labeo*, belonging to family Cyprinidae were studied for the generation of DNA barcoding (Table 3). For all fish samples, two primers Fish F1 and Fish R1 were used to generate the sequence with the average length of 650 bp of *COI* gene. There were 571 conserved sites, 99 variable sites, 91 parsimony sites and 7 singleton sites in the barcode sequence generated. No deletion, pseudogenes, and stop codons in analyzed sequences were observed. All these views supports the amplified sequences constitute the functional *COI* gene sequences of mtDNA. This study is in agreement with the work of barcoding of freshwater fishes from Pakistan reported by Karim et al. (2016). Transitional substitute ( $si = 84.43$ ) were found to be high than transversional substitute ( $sv = 15.46$ ). Similar nucleotide pair frequency was reported in the studies of fishes in Turkey. The transitional and transversional bias in this study was  $R = 5.62$ .

The nucleotide base analysis of *COI* gene sequence investigated that AT contents (55.55 %) was higher than GC contents (44.45 %). Similar average nucleotide composition (AT = 55.5, GC = 44.5) was found in previous study of snow trout fishes from Kashmir valley, India (Bashir et al., 2016). These base composition indicated clear pattern of the anti-G bias and same pattern has also been reported in Cyprinidae (Alam et al., 2021). Changes in codon positions of mitochondrial gene are due to vary degree of selection pressure of base mutation and base usage bias in

**Table 5**  
Genetic distance using KP2 Model.

◆◆	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5
1																																
1	0.00																															
1	0.00	0.00																														
1	0.00	0.00	0.00																													
1	0.00	0.00	0.00	0.00																												
1	0.00	0.00	0.00	0.00	0.00																											
2	0.10	0.10	0.10	0.10	0.09	0.09																										
2	0.10	0.10	0.10	0.10	0.09	0.09	0.00																									
2	0.10	0.10	0.10	0.10	0.09	0.09	0.00	0.00																								
2	0.10	0.10	0.10	0.10	0.09	0.09	0.00	0.00	0.00																							
2	0.10	0.10	0.10	0.10	0.09	0.10	0.00	0.00	0.00	0.00																						
2	0.10	0.10	0.10	0.10	0.10	0.10	0.00	0.00	0.00	0.00	0.00																					
3	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.08	0.08	0.08																					
3	0.07	0.07	0.07	0.06	0.06	0.06	0.08	0.08	0.08	0.08	0.08	0.08																				
3	0.07	0.07	0.07	0.06	0.06	0.06	0.08	0.08	0.08	0.08	0.08	0.08	0.00																			
3	0.07	0.07	0.07	0.06	0.06	0.06	0.08	0.08	0.08	0.08	0.08	0.08	0.00	0.00																		
3	0.07	0.07	0.07	0.06	0.06	0.06	0.08	0.08	0.08	0.08	0.08	0.08	0.00	0.00	0.00																	
3	0.07	0.07	0.07	0.06	0.06	0.06	0.08	0.08	0.08	0.08	0.08	0.08	0.00	0.00	0.00	0.00																
4	0.07	0.07	0.07	0.07	0.07	0.07	0.09	0.09	0.09	0.09	0.09	0.09	0.05	0.05	0.05	0.05	0.05															
4	0.08	0.07	0.07	0.07	0.07	0.07	0.09	0.09	0.09	0.09	0.09	0.09	0.05	0.05	0.05	0.05	0.05	0.05														
4	0.07	0.07	0.07	0.07	0.07	0.07	0.09	0.09	0.09	0.09	0.09	0.09	0.05	0.05	0.05	0.05	0.05	0.05	0.05													
4	0.08	0.07	0.07	0.07	0.07	0.07	0.09	0.09	0.09	0.09	0.09	0.09	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.00												
4	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.09	0.08	0.09	0.08	0.08	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.00	0.00											
4	0.07	0.07	0.07	0.07	0.07	0.07	0.09	0.08	0.09	0.09	0.09	0.09	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.00	0.00	0.00										
5	0.07	0.07	0.07	0.06	0.06	0.06	0.10	0.10	0.10	0.10	0.10	0.10	0.07	0.07	0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	
5	0.07	0.07	0.07	0.06	0.06	0.06	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.07	0.07	0.07	0.07	0.07	0.07	0.05	0.06	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.00	
5	0.07	0.07	0.07	0.07	0.07	0.07	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.07	0.07	0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.00	0.00
5	0.07	0.07	0.07	0.07	0.07	0.07	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.07	0.076	0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.00	0.00	0.00
5	0.07	0.07	0.07	0.07	0.07	0.07	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.07	0.06	0.069	0.06	0.07	0.07	0.06	0.06	0.05	0.06	0.06	0.06	0.05	0.00	0.00	0.00	0.00	0.00	
5	0.07	0.07	0.07	0.06	0.07	0.07	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.07	0.07	0.070	0.06	0.07	0.07	0.05	0.06	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.00	

◆◆ Evaluates the highest and lowest genetic distance. ◆◆ Number of respective species: 1 = *Catla catla*, 2 = *Labeo bata*, 3 = *Labeo calbasu*, 4 = *Labeo gonius*, 5 = *Labeo rohita*.

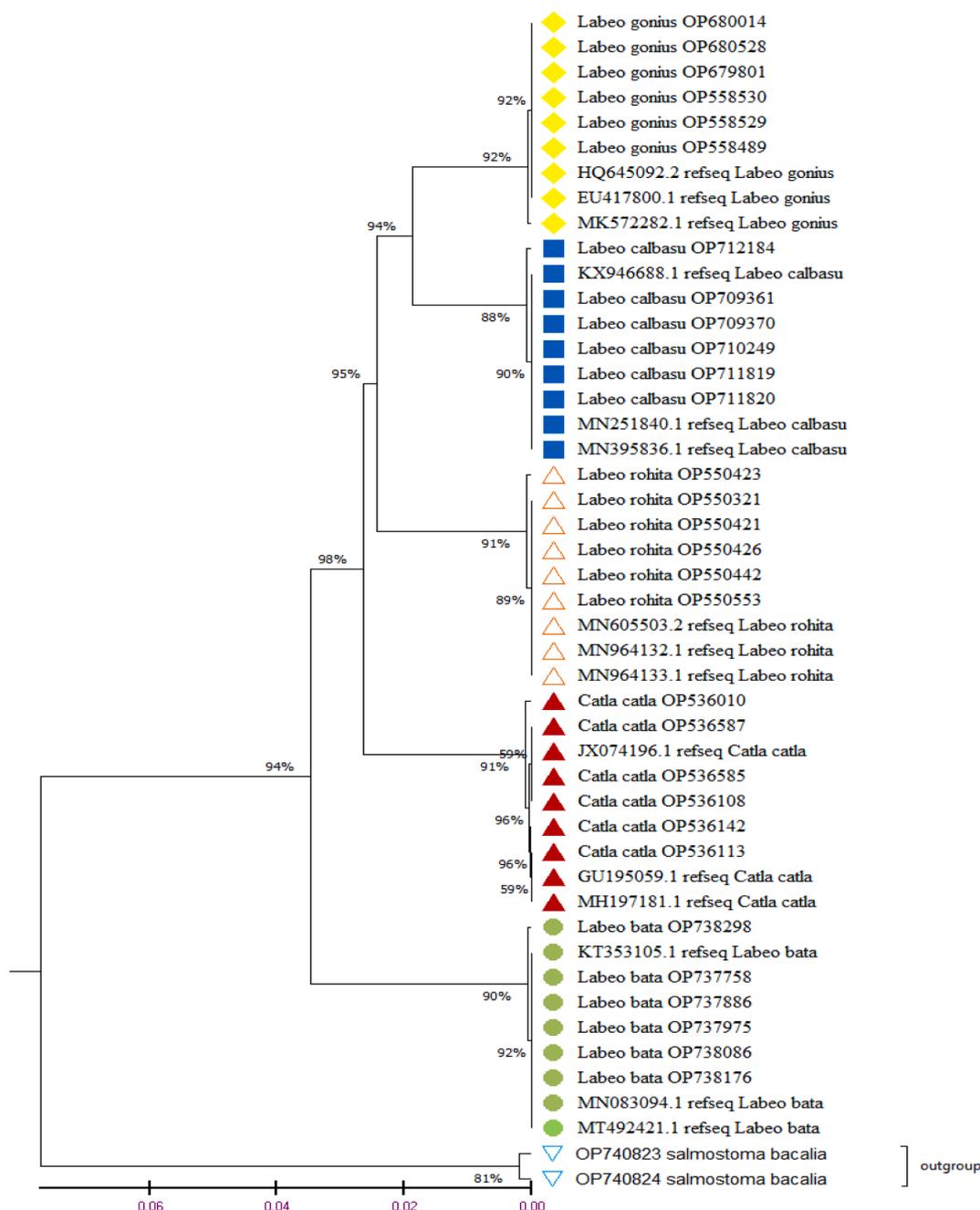
**Table 6**

The number of groups concreted from automated barcode discovery analysis according to prior intraspecific divergence.

Partition	Groups	Prior maximal distance (P)
Partition 1	9	P = 0.001000
Partition 2	7	P = 0.001668
Partition 3	5	P = 0.002783
Partition 4	5	P = 0.004642
Partition 5	5	P = 0.007743
Partition 6	5	P = 0.012915
Partition 7	5	P = 0.021544
Partition 8	5	P = 0.035938

codon position may be due to base mutation pressure. Tajima neutrality test was identified as  $D = 1.382883$  for all species. Positive value of  $D$  indicated a strong selective pressure and lack of population expansion within examined species. That's why we consider the hypothesis of selective neutrality. In this study nucleotide diversity among five species of family Cyprinidae was also found  $\pi = 0.049981$  which is higher than the nucleotide diversity was reported in 68 species of *Sicyopterus stenorhynchus*  $\pi = 0.0039$  by Taillebois et al. (2013).

The accuracy of DNA barcodes for the identification of fish depends on the and interspecific and intraspecific divergence. To delineate species, barcode analysis was used to identify fish boundaries which parallels to divergence between the closet species within groups. Still there is a no universal threshold defined for the interspecies delineation. To define barcoding gap, lowest congeneric and highest conspecific divergence were used. This genetic difference was the more accurate than



**Fig. 3.** Phylogenetic relationship of fishes of family Cyprinidae from three different locations of River Indus, Pakistan by using UPGMA tree method based on the COI partial gene sequences. Adjacent numbers of the branches of tree shows  $\geq 50\%$  bootstrap values and different colored symbols represents each identified fish species.

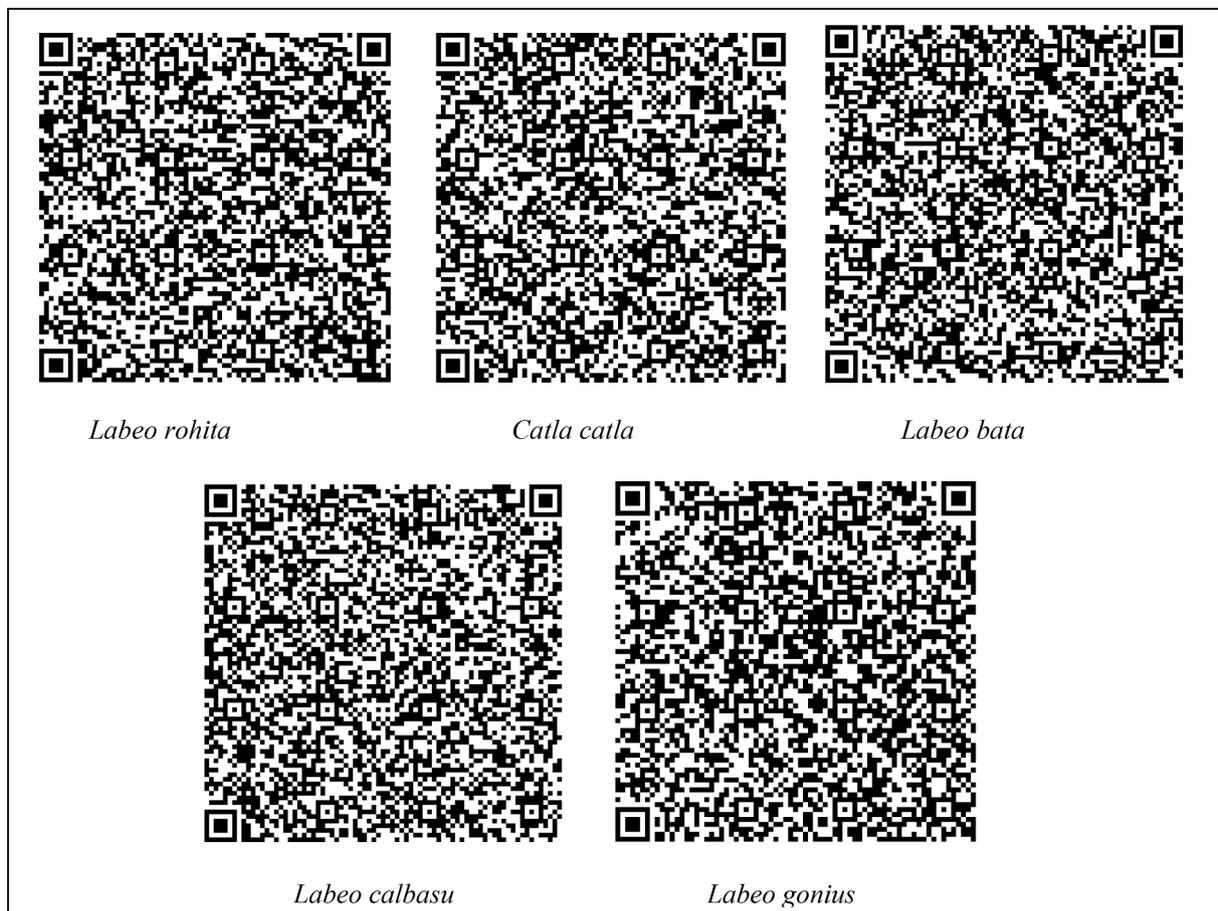


Fig. 4. DNA based QR code of *Labeo rohita*, *Catla catla*, *Labeo bata*, *Labeo calbasu*, *Labeo gonius*.

mean of intraspecific and the interspecific gene sequence variability (Bhattacharjee et al., 2012).

The mean distance within species was calculated as 0.00. The minimum genetic distance (0.04 %) was measured between two closely related taxa *Labeo gonius* and *Labeo calbasu* which was compatible with the study of Esa et al. (2012) where he reported the minimum genetic divergence (0.05 %) between two cyprinid fish species. Maximum evolutionary divergence (0.11 %) was found between *Labeo rohita* and *Labeo bata*. These values varies from species to species and low among closely related species but relatively higher among different ancestral species indicating the relationship between species. The overall mean distance among all fishes of *L. rohita*, *C. catla*, *L. bata*, *L. calbasu* and *L. gonius* was calculated as 0.07.

According to the Pereira et al. (2013), minimum values of congeneric divergence indicates minimum numbers of congeneric species of few genera with three or more species were analyzed. In present study, only one genus was studied, that's why low values of congeneric divergence was estimated. However, low values of congeneric divergence possibly due to recent radiation of congeneric species like in *Labeo* genus.

Automated barcode gap discovery investigated the number of species between 7 and 9 which is based on morphological and DNA barcode identification by constructing accession number. However, the placement of four species *Catla catla* OP536010, *Labeo bata* OP738298, *Labeo calbasu* OP712184, *Labeo rohita* OP550423) into 4 different ABGD groups revealed the presence of four different species. (Ali et al., 2020) also reported ABGD analysis to check the freshwater fish biodiversity of Egypt. Tree was constructed by using Kimura-2 Model to investigate the heirarchical relationships between species along with outgroup. Phylogenetic tree showed different lineage and each lineage further divided into different genus. These further divided into their corresponding

species which formed a monophyletic group. Phylogenetic results showed two major clades at  $\geq 90$  % bootstrap value in *Labeo* genus which describing the clear cut resolution and diagnostic use of cytochrome oxidase I gene as earlier reported (Ude et al., 2020).

One clade contained four species *C. catla*, *L. rohita*, *L. calbasu* and *L. gonius* and second clade contained *L. bata*. These results showing close relationship among species of *C. catla*, *L. rohita*, *L. calbasu* and *L. gonius* (Fig. 3). Thus, the use of DNA barcoding was considered the most efficient and accurate for identification of species due to 100 % success rate which is recored in various study (Ude et al., 2020). Some study showed the success rate from 90 % to 99.8 % (Iyiola et al., 2018). But our study revealed the success rate 99.85 % to 100 % among species. These five species was also identified by phylogenetic analysis. However, morphometric and DNA barcoding clearly support each other for species identification and phylogeny of *Labeo* genus species from River Indus, Punjab Pakistan. In the current study, first time QR code generator was used for generating DNA based QR codes barcodes of studied species for the fast and precise identification of labeo species (Fig. 4). Which help to increase the consumer awareness about food quality. Because fish and fishery products are among the most internationally traded products which are at forefront for food safety (Selamoglu, 2021a,b). This is the first study to create QR codes for the molecular approaches-based identification of the examined labeo species. A DNA barcode was previously created by Yang et al. (2019) as an illustration for the accurate identification of Teleost fish species. In contrast to Yang et al. (2019), our strategy generated the DNA sequence-based QR codes rather than employing a Bio-Rad DNA barcode generator to produce barcodes.

Globally, there is a major threat to fish variety. Along with the detrimental effects of studied species, limited natural factors to native species are also getting worse (Tickner et al., 2020). Many toxic

chemicals, pesticides and other inorganic compounds cause many genetic alteration in fishes species (Orun and Erdogan, 2014; Orun et al., 2005). As a result of organic waste and other human activities, the aquatic ecosystem may be greatly contaminated (Caglar et al., 2019). That's why the number of biological important species is decline annually (Tickner et al., 2020). Therefore DNA barcoding is a versatile and accurate approach to assess fish biodiversity, to manage fish resources and for fish conservation (Jaonalison et al., 2022). Unfortunately, fish diversity of Indus River in these areas has been decreased in recent years due to various purpose. In recent flood situation has also created imbalance in fish diversity. In Pakistan vast number of studied fishes has used for various purpose like eating, medicinal, decoration and for research laboratory purposes. Traditional monitoring for fish management is usually expensive, time consuming. However barcode database for the conservation and assessment of fish diversity becoming the more efficient and helpful for misidentification. Molecular study for conservation of studied fishes was developed in River Indus and these data in this region will become more readily available than ever.

## 6. Conclusion

DNA barcoding is a beneficial and accurate approach for the species identification as compared to traditional method. The current study was first designed approach for freshwater fishes conservation and management of River Indus, Pakistan through DNA barcoding. Phylogenetic tree analysis with other published cytochrome oxidase I sequence revealed clear identity of five studied species of *Labeo* genus. Lack of fish species coverage at BOLD and NCBI created essential difficulties in accurate identification if used solely. Therefore different identification methods are recommended for proper and clear species identification. Thus morphometric and DNA barcoding data support each other in species identification of these studied fishes from River Indus, Pakistan.

Future research should utilize a larger sample size of fish to evaluate genetic diversity more thoroughly for the use of nuclear genes and other molecular markers to confirm the boundaries of these species. In conclusion, the fish species in these area can be accurately identified by combining morphological methods with DNA delimitation methods.

## CRedit authorship contribution statement

**Saba Zafar:** Formal analysis, Investigation, Methodology, Validation, Writing – original draft. **Farhat Jabeen:** Conceptualization, Funding acquisition, Project administration, Supervision, Visualization, Writing – review & editing. **Muhammad Ali:** Conceptualization, Funding acquisition, Writing – review & editing. **Aiman sajjad:** Formal analysis, Investigation.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors are grateful to Higher Education Commission of Pakistan for funding this research project under grant number 5698/PUNJAB/NRPU/R&D/HEC/2016 and the Punjab Fisheries Department for helping sample collection.

## References

Abro, N.A., Waryani, B., Narejo, N.T., Ferrando, S., Abro, S.A., Abbasi, A.R., Ul-Hassan, H., 2020. Diversity of freshwater fish in the lower reach of Indus River, Sindh province section, Pakistan. *Egypt. J. Aquac. Biol. Fish.* 24 (6), 243–265. <https://doi.org/10.21608/EJABF.2020.111114>.

Adeoba, M.I., Kabongo, R., Van der Bank, H., Yessoufou, K., 2018. Re-evaluation of the discriminatory power of DNA barcoding on some specimens of African Cyprinidae (subfamilies Cyprininae and Danioninae). *ZooKeys* 746, 105. <https://doi.org/10.3897/zookeys.746.13502>.

Alam, M.A., Parvez, I., Ara, Y., Khan, M.N., Nehrin, S., Mahajebin, T., Hassan, M., 2021. Phylogenetic relations of the cyprinid fishes (Cyprinidae) in Bangladesh inferred from morphological traits and cytochrome b gene sequences. *AAFL Bioflux* 14, 1631–1644. <http://www.bioflux.com.ro/docs/2021.1631-1644.pdf>.

Ali, F.S., Ismail, M., Aly, W., 2020. DNA barcoding to characterize biodiversity of freshwater fishes of Egypt. *Mol. Bio. Rep.* 47, 5865–5877. <https://doi.org/10.1007/s11033-020-05657-3>.

Barman, A., Sharma, D.K., 2017. Morphometric study of *Labeonandina*, the fresh water endangered cyprinid from Brahmaputra River, Assam and a comparative analysis with four locally available *Labeo* species. *J. Biol. Sci.* 3 (5), 9. <https://doi.org/10.53555/bs.v3i5.1625>.

Bashir, A., Bisht, B.S., Mir, J.I., Patiyal, R.S., Kumar, R., 2016. Morphometric variation and molecular characterization of snow trout species from Kashmir valley. *India. Mit. DNA Part A* 27 (6), 4492–4497. <https://doi.org/10.3109/19401736.2015.1101537>.

Bhattacharjee, M.J., Laskar, B.A., Dhar, B., Ghosh, S.K., 2012. Identification and re-evaluation of freshwater catfishes through DNA barcoding. *PlosOne* 7 (11). <https://doi.org/10.1371/journal.pone.0049950>.

Bhattacharya, M., Sharma, A.R., Patra, B.C., Sharma, G., Seo, E.M., Nam, J.S., Lee, S.S., 2016. DNA barcoding to fishes: current status and future directions. *Mitochondrial DNA Part A* 27 (4), 2744–2752. <https://doi.org/10.3109/19401736.2015.1046175>.

Bilgin, R., Utkan, M.A., Kalkan, E., Karhan, S.U., Bekbölet, M., 2014. DNA barcoding of twelve shrimp species (Crustacea: Decapoda) from Turkish seas reveals cryptic diversity. *Medit. Mar. Sci.* 16 (1), 36–45. <https://doi.org/10.12681/mms.548>.

Biswas, I., Nagesh, T.S., Sajina, A.M., Radhakrishnan, K., 2018. Morphometric variations and meristic counts of *Clupisomagarua*: an approach for assessing stock structure in the River Ganga. *J. Exp. Zool. India* 21, 813–822. <http://www.connectjournals.com/jez>.

Caglar, M., Canpolat, O., Selamoglu, Z., 2019. Determination of some heavy metal levels in three freshwater fish in Keban Dam Lake (Turkey) for public consumption. *Iran. J. Fish. Sci.* 18 (1), 188–198. <https://doi.org/10.22092/IJFS.2018.117890>.

Cermakova, E., Lencova, S., Mukherjee, S., Horak, P., Vobruba, S., Demnerova, K., Zdenkova, K., 2023. Identification of fish species and targeted genetic modifications based on DNA analysis: state of the art. *Foods* 12 (1), 228. <https://doi.org/10.3390/foods12010228>.

Elsaided, H., Soliman, T., Abdelmageed, A.A., Abu-Taleb, H.T., 2021. Applications and challenges of DNA barcoding and metabarcoding in African fisheries. *Egypt. J. Aquac. Res.* 47 (1), 1–12. <https://doi.org/10.1016/j.ejar.2021.02.003>.

Esa, Y.B., Japning, J.R.R., Rahim, K.A.A., Siraj, S.S., Daud, S.K., Tan, S.G., Sungan, S., 2012. Phylogenetic relationships among several freshwater fishes (Family: Cyprinidae) in Malaysia inferred from partial sequencing of the cytochrome b mitochondrial DNA (mtDNA) gene. *Pertanika J. Trop. Agric. Sci.* 35 (2), 307–318.

Fricke, R., Eschmeyer, W.N., Fong, J.D., 2021. Eschmeyer's Catalog of Fishes: genera/Species by Family/Subfamily. <http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp>.

Ghouri, M.Z., Ismail, M., Javed, M.A., Khan, S.H., Munawar, N., Umar, A.B., Ahmad, A., 2020. Identification of edible fish species of Pakistan through DNA barcoding. *Front. Mar. Sci.* 7. <https://doi.org/10.3389/fmars.2020.5541>.

Gulhan, F., Selamoglu, Z., 2016. Comparison of the effects of propolis and pollen extracts in the same concentrations on some biochemical and hematological parameters in rainbow trout (*Oncorhynchus mykiss*). *J. Surv. Fish. Sci.* 3 (1), 1–8. <https://doi.org/10.17762/sfs.v3i1.101>.

Hossain, M., Islam, R., Hossen, M., Rahman, M., Ohtomi, J., 2015. Threatened fishes of the world: *Cirrhinusreba* (Cypriniformes: Cyprinidae). *Croatian J. Fish. Ribarstvo* 73 (1), 40–42. <https://doi.org/10.14798/73.1.786>.

Iyola, O.A., Nneji, L.M., Mustapha, M.K., Nzeh, C.G., Oladipo, S.O., Nneji, I.C., Adeola, A.C., 2018. DNA barcoding of economically important freshwater fish species from north-central Nigeria uncovers cryptic diversity. *Ecol. Evol.* 8 (14), 6932–6951. <https://doi.org/10.1002/ece3.4210>.

Jaonalison, H., Durand, J.D., Mahafina, J., Valade, P., Collet, A., Cerqueira, F., Ponton, D., 2022. Application of DNA barcoding for monitoring Madagascar fish biodiversity in coastal areas. *Diversity* 14 (5), 377. <https://doi.org/10.3390/d14050377>.

Kakoolaki, S., Talas, Z.S., Cakir, O., Ciftci, O., Ozdemir, I., 2013. Role of propolis on oxidative stress in fish brain. *Basic Clin. Neurosci.* 4 (2), 153.

Karim, A., Iqbal, A., Akhtar, R., Rizwan, M., Amar, A., Qamar, U., Jahan, S., 2016. Barcoding of fresh water fishes from Pakistan. *Mitochondrial DNA Part. A* 27 (4), 2685–2688. <https://doi.org/10.3109/19401736.2015.1043544>.

Kaur, A., Datta, S.N., Tyagi, A., 2021. Study on biometric parameters and genetic diversity of *Labeo rohita* from harike wetland-a Ramsar site. *Indian J. Animal Res.* 55, 617–623. <https://doi.org/10.18805/ijar.B-4002>.

Orun, I., Erdogan, K., 2014. Role of propolis on biochemical and hematological parameters of *Oncorhynchus mykiss* exposed to cypermethrin. *J. Surv. Fish. Sci.* 1 (1), 21–35. <https://doi.org/10.17762/sfs.v1i1.36>.

Orun, I., Ates, B., Selamoglu, Z., Yazlak, H., Ozturk, E., Yilmaz, İ., 2005. Effects of various sodium selenite concentrations on some biochemical and hematological parameters of rainbow trout (*Oncorhynchus mykiss*). *Fresenius Environ. Bull.* 14 (1), 18–22.

Pereira, L.H., Hanner, R., Foresti, F., Oliveira, C., 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Gen.* 14 (1), 1–14. <https://doi.org/10.1186/1471-2156-14-20>.

Rasal, K.D., Sundaray, J.K., 2020. Status of genetic and genomic approaches for delineating biological information and improving aquaculture production of farmed

- rohu, Labeorohita (Ham, 1822). *Rev. Aquacul.* 12 (4), 2466–2480. <https://doi.org/10.1111/raq.12444>.
- Sarma, K.J., Prajapati, M., Mankodi, P.C., 2017. Morphological description and taxonomic account of Labeo species (Cypriniformes, Family: Cyprinidae) from Gujarat, India. *J. Entomol. Zool. Stud.* 5, 1120–1125. Corpus ID: 217180949.
- Selamoglu, Z., 2018. Selenium compounds for fish health: an update. *J. Survey Fish. Sci.* 4 (2), 1–4. <https://doi.org/10.17762/sfs.v4i2.144>.
- Selamoglu, M., 2021a. Importance of the cold chain logistics in the marketing process of aquatic products: an update study. *J. Surv. Fish. Sci.* 8 (1), 25–29. <https://doi.org/10.17762/sfs.v8i1.64>.
- Selamoglu, M., 2021b. The effects of the Ports and water transportation on the Aquatic ecosystem. *Op Acc J Bio Sci & Res.* 10 (1) <https://doi.org/10.46718/JBGSR.2021.10.000239>.
- Sheraliev, B., Peng, Z., 2021. Molecular diversity of Uzbekistan's fishes assessed with DNA barcoding. *Scientific Rep.* 11 (1), 1–12. <https://doi.org/10.1038/s41598-021-96487-1>.
- Sudha, V., Subburaj, J., Mathialagan, R., Thangaraj, M., Sivakumar, R., Chandrasekar, S., 2015. Morphological identification and DNA barcoding study of Labeo calbasu (hamilton, 1822) from lower anicut, Tamil nadu, India. *Indo – Asian J. Multidiscipl. Res.* 1, 355–364. [www.jpsscientificpublications.com](http://www.jpsscientificpublications.com).
- Taillebois, L., Castelin, M., Ovenden, J.R., Bonillo, C., Keith, P., 2013. Contrasting genetic structure among populations of two amphidromous fish species (Sicydiinae) in the Central West Pacific. *PLoS One* 8 (10), e75465.
- Tickner, D., Opperman, J.J., Abell, R., Acreman, M., Arthington, A.H., Bunn, S.E., Young, L., 2020. Bending the curve of global freshwater biodiversity loss: an emergency recovery plan. *BioSci.* 70 (4), 330–342. <https://doi.org/10.1093/biosci/biaa002>.
- Tsoupas, A., Papavasileiou, S., Minoudi, S., Gkagkavouzis, K., Petriki, O., Bobori, D., Triantafyllidis, A., 2022. DNA barcoding identification of Greek freshwater fishes. *PlosOne* 17 (1), e0263118.
- Turan, C., Erguden, D., Gurlek, M., Basusta, N., Turan, F., 2004. Morphometric structuring of the anchovy (*Engraulis encrasicolus* L.) in the black, Aegean and Northeastern Mediterranean seas. *Turk. J. Vet. Anim. Sci.* 28, 865–871. <https://journals.tubitak.gov.tr/veterinary/vol28/iss5/13>.
- Ude, G.N., Igwe, D.O., Brown, C., Jackson, M., Bangura, A., Ozokonkwo-Alor, O., Das, A., 2020. DNA barcoding for identification of fish species from freshwater in Enugu and Anambra States of Nigeria. *Conserv. Gen. Resour.* 12 (4), 643–658. <https://doi.org/10.1007/s12686-020-01155-7>.
- Weigt, L.A., Driskell, A.C., Baldwin, C.C., Ormos, A., 2012. DNA. DNA barcoding fishes. In: *DNA Barcodes*. Humana Press, Totowa, NJ, pp. 109–126. [https://doi.org/10.1007/978-1-61779-591-6\\_6](https://doi.org/10.1007/978-1-61779-591-6_6).
- Yang, C.H., Wu, K.C., Chuang, L.Y., Chang, H.W., 2019. Decision theorybased COI-SNP tagging approach for 126 Scombriformes species tagging. *Front. Gen.* 10, 259. <https://doi.org/10.3389/fgene.2019.00259>.