




Molecular approach for identification of *Catla catla* using mitochondrial CO1 from Pakistan

Zara Naeem, Samrah Masud , Shoaib Hassan  and Muhammad Naeem 

Institute of Pure and Applied Biology, Bahauddin Zakariya University, Multan, Pakistan

ABSTRACT

DNA barcoding is a rapid, precise, and effective way of species identification. A short and standard target gene marker is used to create sequence profile of identified species. Specific tag or marker is used, which is derived from mitochondrial COI for identification. Effectiveness of this method axes the degree of divergence among species. Identification is necessary for their representation. In the present work, *Catla catla* was used to study by using Cytochrome C Oxidase 1. The genetic distances were computed, and Neighbor Joining tree was constructed based on the Kimura 2 Parameter method. GenBank and BOLD revealed definitive identity matches. Conspecific and congeneric K2P nucleotide divergence was estimated. Evolutionary tree was analyzed clearly by relating their species to phylogenetic tree, as same as species were bunched under same tree node, while species were differently clustered under distinct nodes. These findings conclude that the gene sequence may serve as a milestone for identification and phylogenetic history of related species at molecular level.

ARTICLE HISTORY

Received 21 July 2019
Accepted 9 May 2020

KEYWORDS

Catla catla; COI; DNA barcoding; identification

Introduction

The identification of fish species is one of the major tasks of taxonomy. The identification of fish species is commonly based on the visible morphology and is carried out using different morphological keys (Ward 2009). DNA barcoding provides speed accuracy in species identification with a focus on analysis on small fragment mitochondrial DNA (Muchlisin et al. 2012). Taxonomic issues can also be solved by these molecular studies. DNA barcoding is known to be a source of species identification (Li et al. 2018). Sequencing is the most important for the species of living creatures since it is a great tool used for this variety and also provides comprehensive information (Singh et al. 2010). Identification of the species based on certain sequences of species of mitochondrial DNA helps the precise identification of unknown species (Dawnay et al. 2007). Main barcode goal is assessed by using Cytochrome C oxidase 1 gene to identified species unknown into known species (Kerr et al. 2009). DNA barcoding becomes known as a molecular method for the identification of species. DNA barcoding relies on specific region of the mitochondrial gene being sequenced, amplified, and analyzed comparison. Molecular basis for biological barcode, to identify organism is the central goal of DNA barcoding, used to create a standardized library for DNA based identification of target species (Kerr et al. 2007).

DNA barcoding can correct the field of misidentification, reduces ambiguity for identification of species, exact species identification, and expand taxonomists expertise (Stoeckle et al. 2004). The precise organism identification has been the

realm of taxonomic experts and identifies an organism; DNA based identification system uses standardized molecular techniques of DNA extraction, Polymerase chain reaction, and DNA sequencing that is used for identification of an unknown organism (Seifert et al. 2007).

The aim of the present study was to carry out molecular-based identification of *Catla catla* using mitochondrial CO1 from Pakistan.

Materials and methods

Fish sampling

A total of 15 samples of *C. catla* were collected from River Chenab (Lat: 30°04'31.33"N, Long: 71°11'31.67"E), Punjab, Pakistan. Samples were identified with the help of standard taxonomic key (Mirza and Sandhu 2007) on the basis of morphological characters. *Catla catla* or *Labeo catla* is Synonyms of *Gibelion catla* as confirmed by various studies, such as Hamilton (1822), Jhingran (1966), Bhuiyan (1964), Shaw and Shebbear (1937), Rahman (1974), and Menon (1974), and locally used in Pakistan. DNA extraction was carried out by a modified phenol-chloroform method and stored at -20°C until further analysis. Quantification of extracted DNA was carried out with nanophotometer (IMPLEN) at A₂₆₀/A₂₈₀ nm absorbance. Polymerase chain reaction of extracted DNA for amplification of the identification region was carried out. Polymerase chain reaction was carried out with reaction setting (one cycle of denaturation at 95°C with respective 40

Table 1. Primers sequences used for PCR amplification and sequencing through CO1 identification gene.

Sr. no.	Primer	Sequence (5'→3')	TM (°C)	GC%	Primer size (nt)
1	Forward primer	TCAACCAACCACAAAGACATTGGAAC	64.7	46.15	26
2	Reverse primer	TAGACTTCTGGGTGCCAAAGAATCA	66.3	46.15	26

TM: Melting temperature; nt: Nucleotide (Naeem and Hassan 2019).

Table 2. Kimura 2 parameter distances.

Level of taxonomy	Genetic distance (Kimura 2 parameter percent)				
	Comparison within species (intraspecific)		Comparison between species (interspecific)		
	Genera value (%)	Specie value (%)	Distance variance (%)	Maximum (%)	Average value (%)
Species values	6.72	0.38	0.04	0.65	0.43

cycles of denaturation phase at 94 °C, annealing at 55 °C and extension at 72 °C with one cycle of final extension for 7 min. Sequence of primer used for polymerase chain reaction amplification and sequencing by Cytochrome C oxidase 1 gene is given in Table 1. PCR products were powered by 1.5% gels and displayed on BioRed Gel Doc to observe the quality of the product. All the samples were used to extract DNA, but for further analysis only clear extracted DNA was used for sequence purpose. The obtained sequence was analyzed by BOLD system and blast on NCBI to identify the unknown sequence to the known product was sequenced and analyzed by using BioEdit lign (version 7.0.5.3) following the method of Hall (1999). Pairwise genetic distance was used to calculate Kimura 2 Parameter distance (Hebert et al. 2003). Neighbour-Joining (NJ) tree (Saitou and Nei 1987) was constructed with MEGA 5 software (Tamura et al. 2011). Sequence accession number as following as provided by GenBank (MT373809).

Results

Results of the present study revealed that *C. catla* identification with a total read length of consensus sequence was found 641 base pair. Molecular identification of *C. catla* was carried out, and the total read length of consensus was found 641 base pair. Barcode of life data system match found *C. catla* fish species with 100% similarity index. Results for the study used to calculate the genetic distance by Kimura 2 Parameter analysis inter and intraspecific fish species are shown in Table 2. The NJ tree among species is shown in Figure 1, which was marked for distance analysis from BOLD system.

Discussion

DNA barcoding allows for identification of species and also shared organisms significant interspecies differences DNA barcoding breaks through the over-reliance on the experiences of taxonomists in traditional morphological classification of species identification. The mitochondrial COI gene exhibits high levels of conservation, genetic variability between different species, is usually utilized as a species barcode, and its high efficiency in species identification as marked in the present study.

A short 648 base-pair region in the mtDNA Cytochrome c oxidase 1 gene and its resulting polypeptide CO1 qualifies to serve as a standardized DNA barcode for all animals (Consortium for the Barcode of Life [CBOL] 2009) in

accordance with the above characteristics. The gene has successfully served as a barcode in many different animals such as birds, fish, and insects (Hebert et al. 2004; Ward et al. 2005; Hajibabaei et al. 2006), mainly because it is well conserved, showing low levels of variance, within a species but it is at the same time showing enough divergence between species to allow for differentiation among many different species (Hebert et al. 2003).

Comparing results obtained from database NCBI and BOLD revealed that the identification of species was different. Results were also supported by percentages on similarity basis and phylogenetic tree analysis and limited to correct identification of the species (Ward et al. 2005). 641 base pair sequence of *C. catla* was used for identification region in the present study as 650 base pair CO1 gene sequence was studied by Lohman et al. (2009), and a 649 base pair used to identify the species by Naeem and Hassan (2019). The primary purpose of molecular study is to identify unknown species (Kerr et al. 2009).

Hybridization may also problematic for DNA barcoding species identifications (Hebert et al. 2003, 2004; Mitchell 2008; Ward 2009). Occasionally, a complex relationship between the same species results in failures for the identification of species using DNA barcoding. Marshall et al. (2008) found that there was no species gap within the mitochondrial data to establish a threshold between intra and interspecific variation. The findings of this study add support to previous studies' conclusions that the barcoding gap does not always as phenomenal (Lukhtanov et al. 2009; Ward 2009). The efficiency of species identification through DNA barcoding depends on both interspecific divergence and intraspecific divergence (Bhattacharjee et al. 2012). In this study, the average intraspecific K2P distance was 0.38%, compared with 6.72% for species within genera as it is many-fold higher than the mean intraspecific distance as justified by with the number of freshwater fish in Indonesia (0.15% and 2.53%, respectively) (Muchlisin et al. 2012), Canadian freshwater fish (0.27% and 8.37%, respectively) (Hubert et al. 2008), and Australian marine fish (0.39% and 9.93%, respectively) (Ward et al. 2005); this result corresponds to the DNA barcoding principle that interspecific divergence sufficiently out-scores intraspecific divergence. Comparison (260/280 nm) of 1.6–2 for clear DNA extraction product (Cawthorn et al. 2011) showed that the quality of DNA was found 1.6–1.767. These results of the current study were found in general agreement with the results of Naeem and Hassan (2019). In this study, the average K2P of intraspecific species was 0.2%, compared

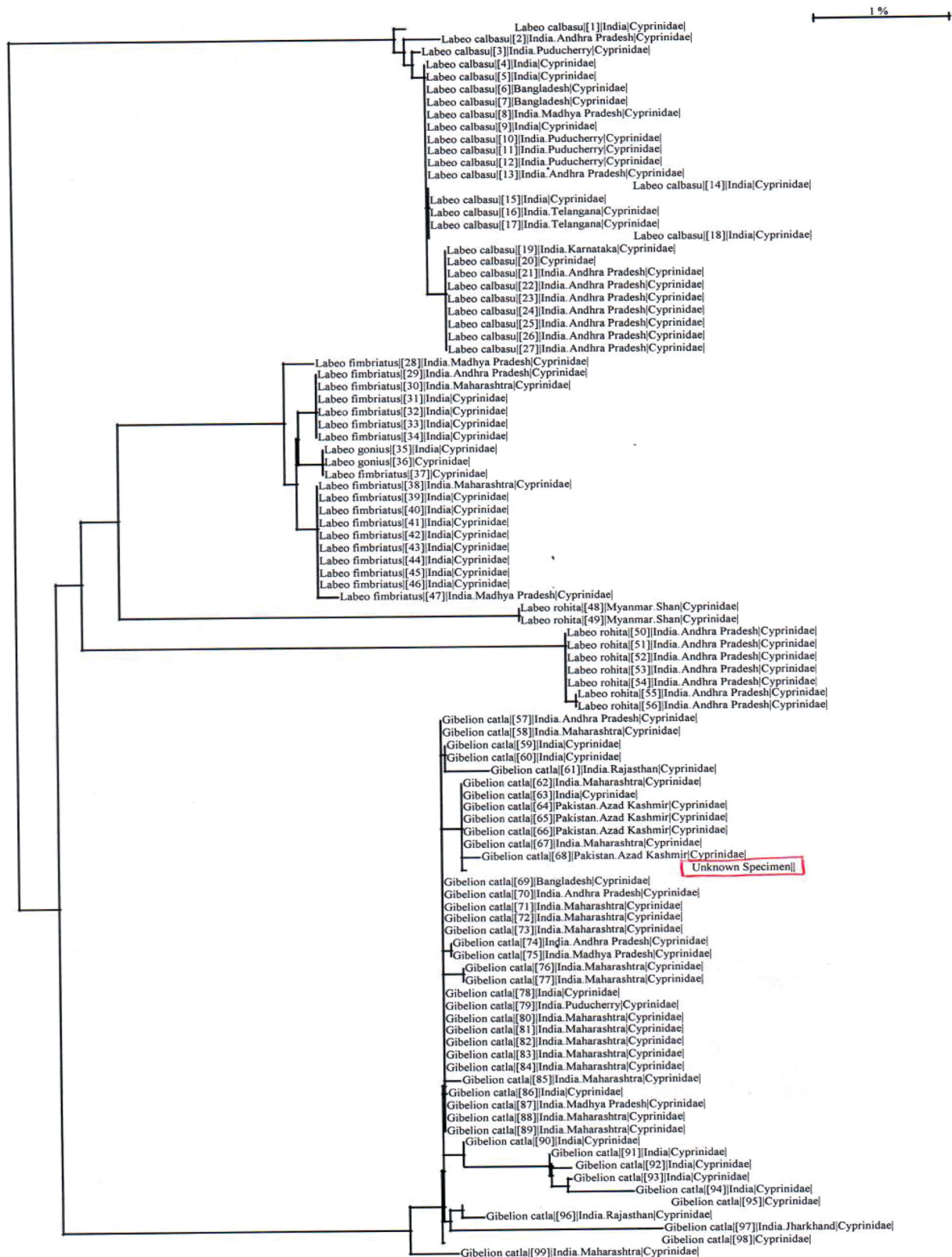


Figure 1. Hierarchical placement of unknown specimen.

to 6.50% for genera. Kenchington et al. (2017) concluded with the principle of DNA indicating that the interspecific separation is sufficient to exchange the differences. In addition, the difference was greater than the 13.9 differences that were mentioned by the marine fish that often met the Atlantic of Canada. In total, the total number of 639 base

pairs were found, the size and the average K2P found to be 0, 1.41, and 0.2%, respectively, within their types, as so 0.11, 1.82, and 0.34% (Naeem and Hassan 2019) available to *C. catla* found in general agreement.

The results of this study reveal that the DNA barcoding has succeeded in identifying the fish species. The DNA-based

description of the genre can be used to assess the fish variety and to assess the production of fish (Takahara et al. 2013). This approach will provide guidance for future fish stocks that are required to be deposited. Once the DNA barcode database is established, the scientific and practical benefits of fishing will be different. As DNA barcoding can distinguish all types of fish and distinguish between the eggs, the larva, and the species, hence, the results will provide more information about the different types of fish to sailors and environmental experts who are responsible for maintenance and sustainable use of fish resources.

Conclusion

DNA barcoding of a functional fish species identification technique compares the different conventional methods. Previously used identification techniques have several limitations that do not work in developmental body stages of fish, processed, fillet, and in case of specimen damaged. DNA barcoding is based on CO1 gene of mitochondrial DNA and has enough variability to differentiate the species. Also, it estimates the nucleic divergence among species, genus, and family. Moreover, DNA barcoding in the field of taxonomy is a useful tool for fish identification.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Samrah Masud  <http://orcid.org/0000-0001-7037-0259>
Shoaib Hassan  <http://orcid.org/0000-0002-1323-3490>
Muhammad Naem  <http://orcid.org/0000-0002-5311-3557>

Data availability

The data that support the finding of this study is openly available in the National Center for Biotechnology Information (NCBI) at <https://www.ncbi.nlm.nih.gov/nuccore/MT373809>, reference number (MT373809).

References

- Bhattacharjee MJ, Laskar BA, Dhar B, Ghosh SK. 2012. Identification and re-evaluation of freshwater catfishes through DNA barcoding. *PLoS One*. 7(11):e49950.
- Bhuiyan AL. 1964. *Fishes of Dacca*. Asiatic Soc. Pakistan, Publ. No. 13, Dacca. p. 31.
- Cawthorn DM, Steinman HA, Witthuhn RC. 2011. Comparative study of different methods for the extraction of DNA from fish species commercially available in South Africa. *Food Cont.* 22(2):231–244.
- Consortium for the Barcode of Life (CBOL). 2009. What is DNA barcoding? <http://www.barcodeoflife.org/content/about/what-dna-barcoding>
- Dawnay N, Oden R, McEwing R, Carvalho GR, Thorpe RS. 2007. Validation of the barcoding gene COI for use in forensic genetic species identification. *Forensic Sci Int*. 173(1):1–6.
- Hall T. 1999. Bioedit: biological sequence alignment editor and analysis program for Window 95/98/NT. *Nucleic Acid Res*. 41:95–98.
- Hamilton B. 1822. *Fishes of the Ganges*. Edinburgh: Archibald Constable and Company; p. 277.
- Hajibabaei M, Smith MA, Janzen DH, Rodriguez JJ, Whitfield JB, Hebert P. 2006. A minimalist barcode can identify a specimen whose DNA is degraded. *Mol Eco Notes*. 6(4):959–964.
- Hebert PDN, Cywinska A, Ball SL, DeWaard JR. 2003. Biological identifications through DNA barcodes. *Proc Biol Sci*. 270(1512):313–322.
- Hebert PDN, Stoeckle MY, Zemplak TS, Francis CM. 2004. Identification of birds through DNA barcodes. *PLoS Biol*. 2(10):e312.
- Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burrigge M, Watkinson D, Dumont P, Curry A, Bentzen P, et al. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One*. 3(6): e2490.
- Jhingran VG. 1966. Synopsis of biological data on catla, *Catla catla* (Hamilton, 1882). *FAO Fish. Synop.* (32): pag.var.
- Kenchington EL, Baillie SM, Kenchington TJ, Bentzen P. 2017. Barcoding Atlantic Canada's mesopelagic and upper bathypelagic marine fishes. *PLoS One*. 12(9):e0185173.
- Kerr KC, Lijtmaer DOA, Barreira AS, Hebert PD, Tubaro PL. 2009. Probing evolutionary patterns through DNA barcodes. *PLoS One*. 4(2):e4379.
- Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PDN. 2007. Comprehensive DNA barcode coverage of North American birds. *Mol Eco Notes*. 7(4):535–543.
- Li Y, Zhang Z, Yang J, Lv G. 2018. Complete chloroplast genome of seven *Fritillaria* species, variable DNA markers identification and phylogenetic relationships within the genus. *PLoS One*. 13(3):e0194613.
- Lohman DJ, Prawiradilaga DM, Meier R. 2009. Improved COI barcoding primers for Southeast Asian perching birds (Aves: Passeriformes). *Mol Eco Resour*. 9(1):37–40.
- Lukhtanov V, Sourakov A, Zakharov EV, Hebert P. 2009. DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification. *Mol Eco Resour*. 9(Suppl 5):1302–1310.
- Marshall DC, Slon K, Cooley JR, Hill KBR, Simon C. 2008. Steady Pliocene Pleistocene diversification and a 2-million-year sympatry threshold in a New Zealand cicada radiation. *Mol Phylogenet Evol*. 48(3): 1054–1066.
- Menon AGK. 1974. A check-list of fishes of the Himalayan and Indo-Gangetic Plains. Barrackpore, India: Central Inland Fisheries Research Institute; p. 23.
- Mirza MR, Sandhu IA. 2007. *Fishes of the Punjab, Pakistan*. Pakistan: Polymer Publications.
- Muchlisin ZA, Fadli N, Siti-Azizah MN. 2012. Genetic variation and taxonomy of *Rasbora* group (Cyprinidae) from Lake Laut Tawar. *J Ichthyol*. 52 (4):284–290.
- Naem M, Hassan S. 2019. Molecular approach for identification of *Labeo bata* based on CO1 gene. *Mitochondrial DNA B Resour*. 4(1):244–246.
- Rahman A. 1974. *Freshwater Fisheries Research Station Bulletin No. 1*. p. 201.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 4:406–425.
- Seifert KA, Samson RA, deWaard JR, Houbraken J, Levesque CA, Moncalvo J-M, Louis-Seize G, Hebert PDN. 2007. Prospects for fungus identification using COI DNA barcodes, with *Penicillium* as a test case study. *PNAS*. 104(10):3901–3906.
- Shaw GE, Shebbeare EO. 1937. The fishes of Northern Bengal. *J Royal Asiatic Soc Bengal Sci*. 8:1–137.
- Singh AK, Pathak A, Lakra WS. 2010. Invasion of an exotic fish common carp (Actinopterygii: Cypriniformes: Cyprinidae) in the Ganga River, India and its impacts. *Acta Ichth et Piscat*. 40(1):11–19.
- Stoeckle M, Waggoner PE, Ausubel JS. 2004. Barcoding life: ten reasons. Consortium for the Barcode of Life, v3.0. Retrieved from: <http://www.barcodeoflife.org/content/barcoding-life-ten-reasons-pamphlet>
- Takahara T, Minamoto T, Doi H. 2013. Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PLoS One*. 8(2):e56584.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol E*. 28:2731–2739.
- Ward RD, Zemplak TS, Innes BH, Last PR, Hebert P. 2005. DNA barcoding Australia's fish species. *Philos Trans R Soc Lond, B, Biol Sci*. 360(1462): 1847–1857. 383: 81–89.
- Ward RD. 2009. DNA barcode divergence among species and genera of birds and fishes. *Mol Eco Res*. 9(4):1077–1085.