

**Research Article**

## Use of CO1 gene sequences for computing genetic diversity between *Cirrhinus mrigala* from two different habitats (Farm and River)

Asma Karim\*<sup>1</sup>, Rashid Saif<sup>2,3</sup>, Faisal Salamat Ali<sup>1</sup>, Zohaib Gil<sup>1</sup>, Wajid Ali<sup>1</sup>

<sup>1</sup>Department of Zoology, Govt. College of Science, Wahdat Road, Lahore, Pakistan

<sup>2</sup>Institute of Biotechnology, Gulab Devi Educational Complex, Ferozepur Road, Lahore, Pakistan

<sup>3</sup>Decode Genomics, 264- Q, Johar Town, Lahore, Pakistan

ISSN: 2456-6268

**ARTICLE INFO**

Received: 13 August 2018

Accepted: 05 December 2018

Available online: 30 December 2018

**KEYWORDS**

*Cirrhinus*

*C. mrigala*

CO1 gene

DNA barcoding

**\*CORRESPONDENCE**

drasmakarim@gmail.com

**ABSTRACT**

DNA barcoding technique involving documentation of animal species based on a partial sequence of cytochrome C Oxidase 1 gene from mitochondrial DNA attracted many scientists at the start of 21<sup>st</sup> century. The flaws in morphometric methods for identification of larvae of fishes and other animals, their eggs, and damaged specimen with incomplete morphology upraised the question for an alternate method for identification and grouping of many animal species. Partial sequence of CO1 gene used as barcode has been attested as a valuable tool for identification of fish species as well as supportive in computing evolutionary history and genetic diversity. Current study was conducted to identify the fishes by using the DNA barcoding technique resulting partial sequences of CO1 gene, and then use of these sequences in evaluation of the evolutionary history and genetic diversity of *Cirrhinus mrigala* inhabiting different areas. Short sequences from 5' end of CO1 gene (650 base pair) were amplified, sequenced and analyzed using different soft wares. According to results haplotypes were found which showed genetic variations among fishes from two different habitats. The phylogenetic analysis established a neighboring relationship and a shared ancestor of fish from different habitats. Haplotype diversity of 0.7143 was found in partial sequences of CO1 gene of all experimental fish. From these results it is obvious that fish species sharing identical genus and family but with diverse habitats show genetic variations even though involving a common ancestor.

© 2018 The Authors. Published by [JFLS](http://www.jfls.com). This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

**INTRODUCTION**

The DNA barcoding caught the attention of taxonomists in the beginning of 21<sup>st</sup> century. Scientists have been trying to find out methods for standardized identification of species. In the past various methods were used to identify organisms. These include gel electrophoresis, nucleic acid fingerprinting etc. All these methods had certain limitations because these were applicable to certain organisms and also at particular life stage (Hsieh *et al.*, 2007). Traditional morphometric methods used for fish identification involve features like shape of fish body, different color patterns, scale types and their numbers, position of all fins on the body, type of fin rays, and measurements of various body parts. These features are not found in larvae, processed and damaged fishes. In such conditions genetic identification or DNA barcoding is proved as useful alternate method.

In DNA barcoding technique only small part of DNA sequence is used instead of complete genome. The short mitochondrial DNA sequence is taken from particular gene,

cytochrome C oxidase1 (CO1) gene to create DNA barcode. It is generally considered a reliable, non-expansive and easy molecular identification tool for metazoan taxa (Hebert 2004 and 2005). In DNA bar-coding approximately 700 nucleotides of COI gene at the 5' end are used to identify unknown species (Hebert *et al.*, 2003a, b). He also found that this short sequence of 650-bp region towards 5' end of the COI gene has ability to differentiate a large variety of animal species. This statement was published in a paper titled "biological identifications through DNA barcodes" by Paul Hebert and his team in 2013. They were working at University of Guelph, Canada and their research caught the attention of scientists around the world. DNA barcoding has appeared as a new tool in the field of taxonomy. DNA barcoding system appeared suitable, quick, and accurate species identification method making ecological systems more convenient and reachable (Hebert and Gregory, 2005).

Evolution of organisms on the Earth was supported by certain Global changes in the environment during the Quaternary glacial times (Dynesius and Jansson, 2000; Hewitt, 2000). Many species were successful to survive due to the warmer environment till the end of Pleistocene Period. Before this the temperature was extremely cold, due to this many species were extinct (Hewitt, 2000). The freshwater fishes are better sample for phylo-geographic studies because these depend upon the water routes which are closely related to land (Bernatchez and Wilson, 1988). Fresh water fishes depend on their environment just like obligates present in populations and sub populations.

It is very difficult to find native species whether it is due to interbreeding among them to find out the real rate of gene flow. As with the location change, species are also genetically changed. Pakistan and Azad Jammu and Kashmir contain large amount of geographic diversity of fishes. There are large number of mountains and valleys in these areas. The altitude change is responsible for changes in weather such as temperature and precipitation. The south areas have high summer temperatures and are converted into deserts. The northeastern areas have low temperatures and high precipitations. These physiographic variations show hydrographical characters of country. Pakistan has number of rivers, network of canals and man-made lakes. The variance in fish fauna is due to a transitory zone established between the Pakistani areas and the Oriental, Palearctic and Ethiopian zoogeographical regions. These geographical regions have their impact on distribution of fishes found in different regions of country (Mirza, 1994). Cyprinids are distributed world-wide and their genus is highly diverse that attract the researchers to find the evolutionary history and morphological and genetic diversity with respect to their habitat (Howes, 1991). Due to their high variation and easy adaptation to the environment, there is no consensus knowledge. Unfortunately, due to lack of molecular studies, the cyprinids phylogenetic relationship is unsolved (Cunha ,2002; Gilles, 2001). Management of fish population is very important because of instabilities in local as well as the environmental factors at higher geographical measures (Ostman *et al.*, 2012). The productivity of coastal fish species and their collections are not only subjected to availability of employment, hunting capacities, prey accessibility and predation designs (Ostman *et al.*, 2012; Sundblad *et al.*, 2014; Vetemaa *et al.*, 2010), but also by other factors like eutrophication of the water bodies, over hunting of fishes and stresses built by human (Florin *et al.*, 2013 and Snickars *et al.*, 2015).

With respect to human impact, fish like *C. mrigala* are least affected by fishing pressure but eutrophication of the water bodies can affect their population (Ustups *et al.*, 2013). The Richness of Cyprinids indicates that these species from the carp family probably take benefit from eutrophication (Ådjers *et al.*, 2006; Snickars *et al.*, 2015). Differences in morphological characters of fishes are attributed to difference in environment and habitats. In addition different populations of fishes also show differences on molecular basis which distinguish one population from another. This results in better genetic analysis of fish species (Omer and Abukashwa, 2012). Study of phylogeographical analysis provides an understanding of the history and evolution of different and same species. Phylogeography also helps in the distributions of species in relation to various biotic and abiotic factors

(Devitt, 2006; Riddle and Hafner, 2006; Vallinoto, *et al.*, 2009).

## MATERIALS AND METHODS

Fish sampling was done from River Ravi at Head Balloki region in district Kasur and from one of the fish farm of Lahore. Fish was captured with the help of fisherman from fisheries department with casting net. One of the pectoral fin was removed and transferred immediately to 1.5 ml autoclaved Eppendorf tubes. The fish was then photographed and preserved as voucher specimen.

TIAN amp DNA Extraction Kit was used to extract DNA from fins of fish. Quality of DNA was checked by running it on the agarose gel. For this purpose, 5ul DNA sample along with 3ul loading dye (3X bromophenol blue) was run on 0.8% agarose gel prior to its amplification. Total four DNA samples were conducted at 80V for 40 minutes to determine the success of DNA extraction. DNA bands were seen by passing UV light through the gel in Gel Documentation System (BioDoc- iTM system). The bands in the gel documentation indicated the presence of DNA. All the DNA samples were amplified by using FishF1 and FishR2 primers.

FishF1

5'TCAACCAACCACAAAGACATTGGCAC3'

FishR2

5'ACTTCAGGGTGACCGAAGAATCAGAA3'

Twenty-five µl PCR reaction mixture was conducted for each of the total seven DNA samples with standard recipe. The PCR products with gene of interest were sent for sequencing.

During current study CO1 gene was amplified and sequenced from mitochondrial DNA of fish species from both habitats. These sequences were used to calculate the evolutionary history and genetic diversity of *C. mrigala* from river (n= 6) and *C. mrigala* from farm. For this purpose *C. mrigala* (n=1) from farm already barcoded by Karim, *et al.* (2015) and submitted to Genbank (KP696784) was used. This fish was taken from farm of fisheries research and training institute Manawan, Lahore, Pakistan. *C. mrigala* from Ravi river at head ballaoki region were also barcoded and submitted to Genbank under accession numbers MF468124, MF468127, MF468128, MF468129, MF468130 and KY569370.

## RESULTS AND DISCUSSION

During present work, fresh water fish species from River Ravi and Farm n= 7 were used. All these fish belonged to genus *Cirrhinus*, species *mrigala* and family Cyprinidae. Pectoral fins of fish were considered for the DNA barcodes creation. The Fish F1 and Fish R2 primers were used for amplification of partial CO1 gene which resulted into the barcodes with a normal range of 636 base pairs. There were 209 missing gaps, 291 monomorphic sites, 254 variable sites (polymorphic site) 249 singleton variable sites, and 5 parsimony informative sites found in the barcodes. It was found that all partial CO1 sequences were functional CO1 sequences as no insertions/deletions or stop codons were observed in the analyzed sequences. As these sequences had no stop codons in them, and their average length was 636 which is more than 600bp which confirmed that the NUMT were not sequenced in the present work. NUMT (nuclear DNA sequences that originate from mitochondrial

DNA sequences and are typically less than 600 bp in length) our result is in accordance with the results of Lakra *et al.*, 2011 and Ward *et al.*, 2005.

Base %age of all 4 bases in partial CO1 sequences of all *C. mrigala* from both habitats were calculated (Table 1). Over all excess of Adenine nucleotide was found in all sequences as compared to other nucleotide bases, this pattern is also reported in other cyprinid species as observed by Meyer *et al.*, (2005).

Total GC content comparison in sequences was made between six samples of River *C. mrigala* with 1 sample from Farm. High value of average GC content in *C. mrigala* from River was found as compared to GC content of *C. mrigala* from Farm. The nucleotide composition of fishes from farm and river are shown in table 1. The total GC content of Farm fish was calculated as 46.3% as compared to 53.45% found in River fishes.

The evolutionary history was concluded by using the Maximum Likelihood method based on the Kimura 2-parameter model. The tree with the highest log likelihood (-1500.7860) is shown in Figure 1. The analysis involved 7 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6.

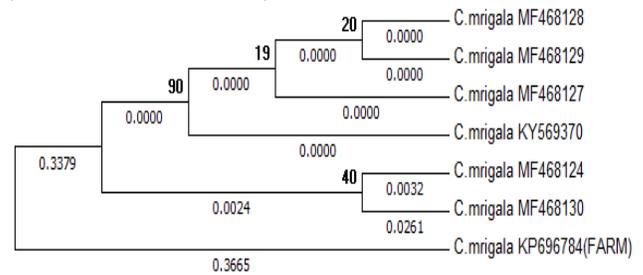
When sequences were analyzed using DnaSP5, haplotypes were found in experimental fish species, showing that all of these form a haplo-group and share same ancestor with a single-nucleotide polymorphism mutation. Four haplotypes were found in all seven sequences. The haplotype diversity is 0.7143.

Transfer of mitochondrial DNA along a maternal lineage dates back thousands of years. Early studies suggest that presence of haplotype numbering as low as 2-6 are enough to express the genetic variation along stretches of DNA (Philip *et al.*, 2013). In our present work 4 haplotypes were found indicating genetic variations among *C. mrigala* from two different habitats. According to another report Haplotype diversity is considered high when it lies in the ranges of 0.75 to 0.92. In another research project study about high Haplotype diversity ranges, 0.98 was also considered as high (Jean *et al.*, 2013). The haplotype diversity in our sequences of fish is 0.7143 which shows that moderate haplotype diversity is present in the sequences. Haplotype numbers found in sequences of fish from Farm and River are shown in table 2. This result of haplotype is also verified by phylogenetic tree (Fig. 1). *C. mrigala* with Genbank accession numbers KY569370, MF468127, MF468128 and MF468129 share same haplotype no. 2, which are clustered together in phylogenetic tree. The sequences with different haplotype number are placed separately.

Saccone *et al.*, (1999) suggested that there is variation in level of GC content of entire genome, within and among major groups of organisms. When Total GC content was studied in mitochondrial DNA, it varied considerably at various taxonomic levels as well as in the same species found in different habitats. The lowest concentration is found in insects and nematodes which range 15–35% while somewhat higher concentration is found in mollusks as it ranges 29–40%. The highest values of GC content is found in birds and mammals (mostly in primates) and teleost fish with ranges 32–46%. Our study on fishes from two different habitats resulted in GC content of 46.5%, which is the range

of teleost fishes. The high percentage of GC content shows DNA stability in them.

Mitochondrial genomes show slight uniformity and show (taxon-dependent) tilts, although in this case, the amount of twist in the coding strand is powerfully linked with overall content of nucleotides G and C content. Fishes and *Xenopus* show a constant GC content percentage in their nuclear genomes. Pattern of base composition vary inside and between species in the course of transformation in a way that the rate of change of bases are not continuous in time or space. It is also found that natural selection also depends on specific pattern of codon or overall GC content (Arne and Adwards, 2000).



**Fig. 1:** Molecular Phylogenetic analysis by Maximum Likelihood method

**Table 1:** Percentage base composition of all samples of *C. mrigala* from River

Samples	Gene bank Accession nos.	%age of base composition				Total length
		T	C	A	G	
<i>C. mrigala</i> Farm	KP696784	28.1	28.4	25.4	18.2	698
<i>C. mrigala</i> River	KY569370	25.4	18.2	28.0	28.3	625.0
	MF468124	24.9	18.5	28.5	28.0	610.0
	MF468127	25.4	18.2	27.9	28.5	625.0
	MF468128	25.8	18.2	27.9	28.2	628.0
	MF468129	25.6	17.9	27.9	28.5	648.0
	MF468130	24.8	18.8	28.9	27.4	616.0
Average		25.3	18.3	28.1	28.2	625.3

**Table 2:** Haplotype data of *C. mrigala*

Species name	No. of haplotypes	Haplotype diversity	Haplotype distribution
<i>C. mrigala</i>	4	0.7143	Hap_1: KP696784(Farm) Hap_2: KY569370, MF468127, MF468128, MF468129 Hap_3: MF468124 Hap_4: MF468130

**Table 3:** Comparison of GC content at the 1st, 2nd and 3rd codon positions of CO1 genes of *C. mrigala* from Farm and River

<i>C. mrigala</i> habitat	No. of fishes	1st	2nd	3rd
River Ravi	6	207.8	210.2	207.3
Farm	1	233.0	232.0	233.0

**Table 4:** Comparison of average base composition of CO1 genes of *C. mrigala* from Farm and River

<i>C. mrigala</i>	No. of fishes	Average base composition				Total length
		T	C	A	G	
River	6	25.3	18.3	28.1	28.2	625.3
Farm	1	28.1	28.4	25.4	18.2	698.0

## REFERENCE

- Ådjers, K., Appelberg, M., Eschbaum, R., Lappalainen, A., Minde, A., Repecka, R., Thoreson, G. 2006. Trends in coastal fish stocks of the Baltic Sea. *Boreal Environ. Res.* 11: 13-25.
- Bernatchez, L. and Wilson C.C. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Mol Ecol.*, 7:431-452.
- Cunha, C., N. Mesquita, T. E. Dowling, A. Gilles & M. M.Coelho, 2002. Phylogenetic relationships of Eurasian and American cyprinids using cytochrome b sequences. *Journal of Fish Biology* 61: 929-944.
- Devitt, T. J. 2006. Phylogeography of the Western Lyresnake (*Trimorphodon biscutatus*): testing aridland biogeographical hypotheses across the Neartic-Neotropical transition. *Mol Ecol.*, 15:4387-4407.
- Dynesius, M. and Jansson, R. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc. Natl. Acad. Sci. USA.*, 97: 9115-9120.
- Florin, A.B., Bergstrom, U., Ustup, D., Lundström, K., Jonsson, P.R., 2013. Effects of a large northern European no-take zone on flatfish populations. *J. Fish Biol.* 83:939-962.
- Gilles, A., Lecointre, G., Miquelis, A., Loerstcher, M., Chappaz, R., Brun, G. 2001. Partial combination applied to phylogeny of European cyprinids using the mitochondrial control region. *Mol. Phylogenet. Evol.*, 19:22-33.
- Hebert, P. D. N., Cywinska, A., Ball, S. L., Deeward, J. R. 2003a. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B., Biol. Sci.*, 270: 313-321.
- Hebert, P. D. N., Ratnasingham, S., De Waard, J. R. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B., Biol. Sci.*, 270: 96-99.
- Hebert, P. and Gregory T.R. 2005. "The Promise of DNA Barcoding For Taxonomy". *Syst Biol.*, 54(5): 852-859.
- Hebert, P.D., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. U.S.A.*, 101: 14812-14817.
- Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907-913.
- Hsieh, H. S., Chai, T.J., Hwang, D. F. 2007. Using the PCR-RFLP method to identify the species of different processed products of billfish meats. *Food Contr.*, 18: 369-374.
- Howes G.J. 1991. Systematics and biogeography: an overview. In: Winfield IJ, Nelson JS (eds), *Cyprinid fishes: systematics, biology and exploitation*. Chapman & Hall, London, pp 1-33
- Jean P.A.H, Lynne v. H, Dean., R. Jerry, G.P, Jones P. L. 2013. High Genetic Diversity in Geographically Remote Populations of Endemic and Widespread Coral Reef Angelfishes (genus: *Centropyge*) 4 *Diversity* ., 5: 39-50.
- Karim, A., Amar, A., Iqbal, A., Rizwan, M., Akhtar, R., Qamar, U., and Shahjahan. 2015. DNA barcoding of fresh water fishes from Pakistan. *Mitoch DNA* early online. Informa UK. Ltd. DOI:103109.pp. 1-4.
- Lakra WS, Verma MS, Goswami M, Lal KK, Mohindra V, Punia P, Gopalakrishnan A, 2011. DNA barcoding Indian marine fishes. *Mol Ecol Resour* 11:60-71.
- Meyer CP, Paulay G. 2005. DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biol* 3:2229-38.
- Mirza, M. R. 1994. Geographical distribution of freshwater fishes in Pakistan: a review. *Punjab Univ. J. Zool.*, 9: 93-108.
- Omer, A., Abukashwa, S., 2012. Morphometric traits and karyotypic features of the African lungfish (Um Koro) *Protopterus annectens annectens* (Owen, 1839) and *Protopterus aethiopicus aethiopicus* (Heckel, 1851) in Sudan. *J. Fish. Poult. Wildl. Sci.* 1 (1)
- Ostman O., Bergenius, Boström, K.M., and Lunneryd, S G. 2012. Do cormorant colonies affect local fish communities in the Baltic Sea? *Canadian Journal of Fisheries and Aquatic Sciences*, 69(6): 1047-1055.
- Philip L.M, Jean-Paul A.H., Lynn V.H., Dean R.J., Geoffrey, P.J. 2013. High Genetic Diversity in Geographically Remote Populations of Endemic and Widespread Coral Reef Angelfishes (genus: *Centropyge*). *Diversity* 2013, 5, 39-50; doi:10.3390/d5010039
- Riddle, B. R. and Hafner, D. J. 2006. A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environ.*, 66: 435-461.
- Saccone, S., Federico, C. and Bernardi, G. 1999. Localization of the gene richest and the gene poorest isochores in interphase nuclei of mammal and birds [https://doi.org/10.1016/S0378-1119\(02\)01038-7](https://doi.org/10.1016/S0378-1119(02)01038-7)
- Snickars, M., Weigel, B., Bonsdorff, E., 2015. Impact of eutrophication and climate change on fish and zoobenthos in coastal waters of the Baltic Sea. *Mar. Biol.*, 162:141-151.
- Sundblad, G., Bergström, U., Sandström, A., Eklov, P., 2014. Nursery habitat availability limits adult stock sizes of predatory coastal fish. *Ices J. Mar. Sci.*, 71:672-680.
- Ustup, D., Mueller-Karulis, B., Bergström, U., Makarchouk, A., Sics, I., 2013. The influence of environmental conditions on early life stages of flounder (*Platichthys flesus*) in the central Baltic Sea. *J. Sea Res.*, 75: 77-84.
- Vallinoto, M., Sequeira, F., Sodre, D., Bernardi, J. A. R., Sampaio, I., Schneider, H. 2009. Phylogeny and biogeography of the *Rhinella marina* species complex (Amphibia, Bufonidae) revisited: implications for Neotropical diversification hypotheses. *Zoologi Script* ., 39: 128-140.
- Vetemaa, M., Eschbaum, R., Albert, A., Saks, L., Verliin, A., Jürgens, K., Kesler, M., Hubel, K., Hannesson, R., Saat, T., 2010. Changes in fish stocks in an Estonian estuary: overfishing by cormorants? *ICES J. Mar. Sci. J. du Cons.*, 67:1972-1979.
- Ward RD, Zemplak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. *Philos Trans R Soc B*, 360:1847-1857.
- Zworykin, D., Budaev, S., Darkov, A., Dzerzhinskii, K., Lyovin, B., Mina, M. 2006. Assessment of the role of chemoreception in the mate choice in barbs of the *Barbus intermedius*; complex from Lake Tana, Ethiopia. *J of Ichth.*, 46: 661-667.