MERIT RESEARCH JOURNALS
www.meritresearchjournals.org

Merit Research Journal of Agricultural Science and Soil Sciences (ISSN: 2350-2274) Vol. 7(5) pp. 071-077, May, 2019
Available online http://meritresearchjournals.org/asss/index.htm
Copyright © 2019 Merit Research Journals
DOI: 10.5281/zenodo.3236481

Original Research Article

DNA barcoding through mitochondrial CO1 gene: Fish Identification, divergence and phylogenetic studies

Dr. Muhammad Naeem

Professor of Zoology, Institute of Pure and Applied Biology, Bahauddin Zakariya University, Multan, 60800, Pakistan

E-mail: muhammadnaeem@bzu.edu.pk

Abstract

DNA barcoding is used to identify particular specie by Cytochrome c oxidase 1 (CO1) mitochondrial DNA sequence. It is also deal with genetic variability between species and its polymorphism. Principally phylogenetic methods based on maximum likelihood and neighbourhood joining tree for variability in sequence and taxonomical classification of species. DNA barcoding is the one means of establishing an effective and precise system for species identification. A standard short target of genetic marker often used to craft sequence contour of identified against unidentified species. Specified marker/tag of mitochondrial cytochrome C oxidase subunit 1 is used for the species identification purpose. Usefulness of following method lay off the divergence degree among species. Compare to previously used identification method through external or visible morphology, barcoding is more reliable. Computation of genetic distances and construction of Neighbor Joining tree based on the Kimura 2 Parameter method were also done systematically. Analysis of evolutionary tree stood clearly by making phylogenetic tree with relationships among their species, alike species bunching under the same tree node while differing species clustered under distinct nodes with no exemptions. Study findings concluded that sequence of Cytochrome C oxidase 1 used as a milestone to provide identification of species at molecular level, its similarity and divergence among the related species.

Keywords: *Oreochromis niloticus*, COI, DNA barcoding, Identification, divergence

INTRODUCTION

DNA barcoding has a significant importance in scientific apparition recently, known as very effective and intensifying species identification tool and has wide reaching popularity. Mitochondrial Cytochrome C oxidase subunit 1 gene recommended identifying species region used to for animals (Hebert et al., 2003). Henceforth, numerous studies shown the sequence diversity of 650 base pair region and nearly have 50 region of the Cytochrome c oxidase subunit 1 gene present well built species level decree for diverse animal group (Schindel et al., 2011), fishes (Laskar et al., 2013, Trivedi et al., 2014). Identification through DNA barcoding comprises short sequence of nucleotide use as standardized tool for identification of specie to which organism belongs (Hebert et al., 2003). Intention is to endow with automatic and simple method for the correct specie identification. Cytochrome c oxydase 1 mitochondrial DNA gene preferred as the barcode for animals and this type of genetic markers showed similar barcoding properties in plants and so. Molecular identification approach of specie has been effectively useful to different type of organisms (Seifert et al., 2007), fisheries (Menezes et al., 2006) and aquaculture (Liu et al., 1998). Analysis of genetic form of data provides set up for conservational priorities

(Bernatchez, 1991) complexities have been arisen in several cases (Elias et al., 2007). DNA barcoding is small nucleotide fragment sequence of mitochondrial genome used for identification and for variation among species. Indeed, the inimitable characters that can makes Cytochrome C oxidase 1 gene as a contender for high throughput for barcoding (Hebert et al., 2003) and also restrict its information contented over in-depth phylogenetic level (Naylor and Brown, 1997).

Genetic form of data is important to assess the flow of gene among various populations in order to maintain genetic diversity. Genetic based technologies are useful for identification of species (Palumbi and Cipriano, 1998). Molecular identification approach through sequencing and amplification of short universal mitochondrial COI for identification and divergence of population (Hebert et al., 2003) .To the best of our knowledge, DNA barcoding through Cytochrome C oxidase I gene for identification, divergence and phylogenetic studies can't be done previously in study region. So study having vast impact and importance in specified region.

METHODOLOGY

Fish sampling

Fish Samples of *Oreochromis niloticus* (Linnaeus, 1758) were collected from River Chenab, Muzafargarh, Punjab, Pakistan. Geographical coordination for collection site (Lat: 30° 04' 31.33" N; Long: 71° 11' 31.67" E) and stored in Fisheries Lab, Institute of Pure and Applied Biology, Bahauddin Zakariya University, Multan, Pakistan.

Species identification

Samples were identified with the help of standard taxonomic key (Mirza and Sandhu, 2007) basis of morphological characters.

DNA isolation by modified phenol chloroform extraction method

DNA was isolated from various tissues (Caudal fin, Pectoral fin, Pelvic fin, Muscle and Gills) were taken for the DNA extraction. The size of each tissue was 50–100 mg with 6-800µl extraction buffers used for grinding and immediately transferred into MCT. Proteinase K was also added, gently mixed through vortex mixer and incubated. After this incubation solution was centrifuged. Supernatant were collected, add phenol: chloroform: isoamyl alcohol (25:24:1), mixed gently and centrifuged at 12000 rpm for 10 minutes. Supernatant was then collected and chloroform: isoamyl alcohol was added in the tube, mixed by gentle inversion for few minutes and centrifuged at 12000 rpm for 10 min, supernatant layer was collected and 0.1% 3M sodium acetate and equal volume of 100% ethanol was added. MCTs were placed at -20°C for 1-2 hrs and centrifuged it again at 1000 rpm for 10min. DNA pellets were washed with 70% ethanol, resuspended in triple distilled water (AquaPro Injection) in appropriate volume and stored at -20°C till further analysis.

Yield and purity of isolated DNA

Impurities in DNA were quantified through the measurement of absorbance at 260/280nm with the help of nanophotometer (IMPLEN). DNA purity was firm by A260/A280 ratios. Reliability of DNA was also observed by using 5µl of DNA with loading dye on .8% gel. Isolated DNA was also quantified by 1 Kb marker (GeneON). Furthermore, polymerase chain reaction (PCR) with CO1 gene of mitochondrial DNA primer (macrogen) was used to amplify DNA.

Polymerase chain reaction

Polymerase chain reaction for amplification of genomic DNA, a conserved region of mitochondrial CO1 gene consisting of 660 base pair fragment were performed by two primers.

To make 25µl Polymerase Chain Reaction solution, 12.5 µl of master mixture (dNTP's, $MgCl_2$, Tag Polymerase; BioShop, Canada Inc.), 6.5 µl water, 1.5 µl each primer (30 nmol), and 2.5 µl DNA template were mixed in PCR tubes. Polymerase chain reaction amplification was carried out in Labcycler (SENSQUEST, Germany), comprising pre-denaturation at 95 °C for 5min, followed by 40 cycles, denaturation at 95°C for 30 sec, annealing at 50°C for 30 sec, and extension at 72°C for 30 sec, with final extension at 72°C for 7 min.

For control region sequencing with set of forward and reverse primer were used as for polymerase chain reaction amplification Table 1. After amplification, the polymerase chain reaction products were run on 2% agarose gel for one hour and best samples were used for sequencing purpose.

Analysis of the data

The align sequence were analyzed by using BioEdit lign sequence version 7.0.5.3 following Hall (1999). Supplementary sequence was downloaded from BOLD and GenBank for validation and for analyses. Pairwise genetical distance within or between the species and genus were calculated by using Kimura 2 Parameter method following Kimura (1980). Neighborhood Joining

Table 1. Primers sequences used for Polymerase chain reaction amplification and sequencing through CO1 Identification gene.

Serial. Number	Primer Name	Primer Sequence (5' →3')	Melting temperature (^{0}C)	GC content	Size of Primer(nt)
1	Forward	TCAACCAACCACAAAGACA TTGGAAC	64.7	46.15	26
2	Reverse	TAGACTTCTGGGTGGCCAA AGAATCA	66.3	46.15	26

Table 2. DNA quantification

Sample No	Concentration ranges	$A^{260/280}$
1-44	452 – 721	1.454-1.731

Table 3. Kimura 2 Parameter inter and intra specie Divergence

	Average Intraspecic	Average interspecie Divergence
SAMN111273	6.277631	3.446341
KU568892.1	0.009172	0.009146
MH026017.1	3.095237	1.634405
KU568892.1	0.000111	0.000113

Table 4. Summary statistics for partitions with its frequency among the divergence of species

Parameter	Mean	Variance	Lower	Upper	Median	PSRF+	Nruns
length[1]	9.343863	22.864273	2.675691	19.785940	8.251208	1.050	2
length[2]	0.004754	0.000017	0.000071	0.013132	0.003740	1.002	2
length[3]	1.738514	1.565293	0.002291	3.833325	1.619900	1.015	2
length[4]	0.007481	0.000021	0.000053	0.015546	0.006945	1.000	2

+Convergence diagnostic (PSRF = Potential Scale Reduction Factor; Gelman and Rubin, 1992) should approach 1.0 as runs converge. NA is reported when deviation of parameter values within all runs is 0 or when a parameter value (a branch length, for instance) is not sampled in all runs.

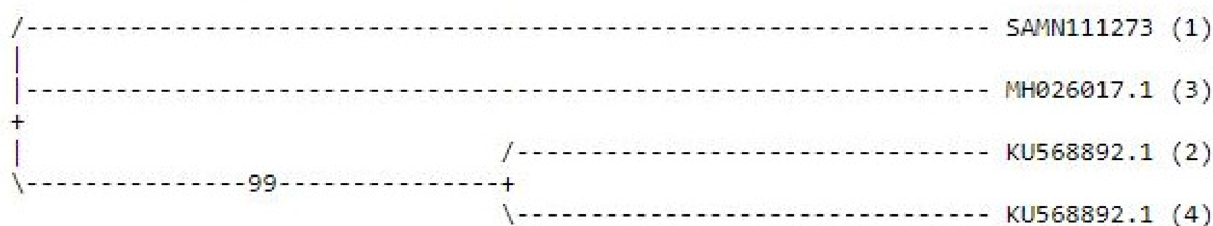
trees (Saitou and Nei) were constructed using the Kimura 2 Parameter model. Kimura 2 Parameter genetic distance and NJ tree were produced through MEGA version 5 software.

Neighbour-joining is a method of constructing a tree from a distance matrix and PhyML alternative phylogenetic method based on maximum likelihood. GenBank accession numbers SAMN111273 (*Oreochromis niloticus*) of present study, KU568892.1 (*Labeo gonius*), MH026017.1 (Labeo *bata*) and KU568892.1 (*Labeo dussumieri*) were used to analyze the results for identification and divergence.

RESULTS

Cytochrome C oxidase 1 was used to amplify the target region to identify specie. Databases revealed ultimate identity with similarity ratio 92.59% in *Oreochromis niloticus*. Length of sequence was noted 660 base pair. Values of concentration its quality ($A_{260/280}$) for quantification purpose also provided in Table 2. Specimen data their assemblies and average percent K2P genetic distance were summarized in Table 3. Genetic distance within specie, between genus, maximum, and average percent K2P genetic distance

Figure 1. Divergence neighborhood joining tree among the following four species.



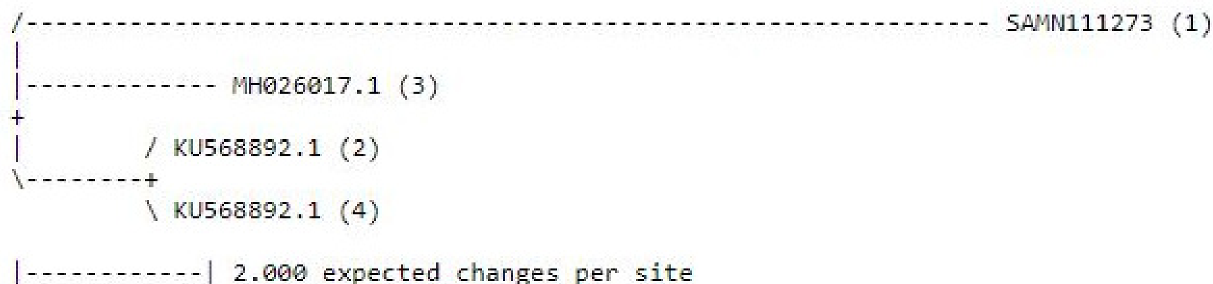


Figure 2. Calculation of Divergence of clade and phylogram in fishes

were found significantly diverge among all species as found in Table 3 and 4.Neighbourhood joining tree shown in Figure 1 and phylogram Figure 2.

Four sequences of were clustered by using Neighbourhood Joining tree, Cytochrome C oxidase subunit I sequence clustered together conferring to the description of specie formed one cluster. Neighbourhood Joining tree was thru as an out-group or as a standard specie. A summary of divergence was shown in Table 3. Output of all the bipartitions with frequency larger or equal in least one run, minimum part frequencies is a parameter to sum command and currently it is set to 0.10. This is followed by a table with statistics for the informative bipartitions, sorted from highest to lowest probability. For each bipartition, the table gives the number of times the partition or split was observed in all runs and the posterior probability of the bipartition, which is the same as the split frequency. If several runs are summarized, this is followed by the minimum split frequency, the maximum frequency, and the standard deviation of frequencies across runs. This is followed by a table summarizing branch lengths, node heights and relaxed clock parameters. The mean, variance, and 95 % credible interval are given for each of these parameters. If several runs are summarized, the potential scale reduction factor (PSRF) is also given. Summary statistics for partitions with frequency found 0.10; Average standard deviation of split frequencies found 0.006591; Maximum standard deviation of split frequencies was 0.006591; average PSRF for parameter values (excluding NA and >10.0) was 1.015 and maximum PSRF for parameter values was found 1.050 (Table 4). So, a significant divergence found in present study also shows divergence in multiple alignment sequence differences among various species as shown in Figure 3.

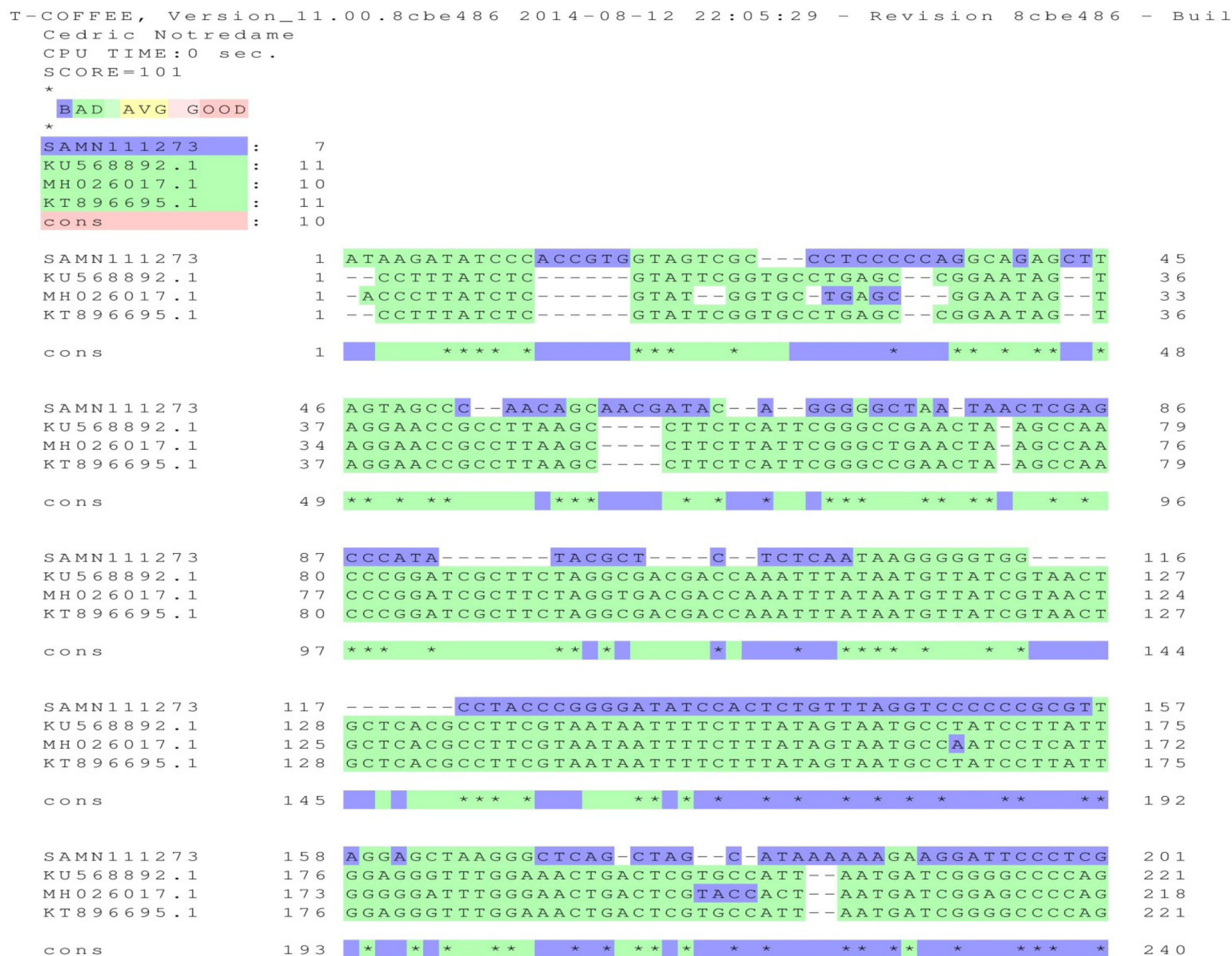


Figure 3. Divergence in multiple alignment sequence

DISCUSSION

Morphological based identification approach has some Limitations so taxonomists used molecular approach for species identification. Efficiency of molecular based approach species identification have intra; interspecific homogeneity and heterogeneity using various methods. Mitochondrial cytochrome C oxidase 1 gene act as striking and highly efficient method for identification of species. DNA barcoding methods used for specie identification and to uncover biological species diversity. Cytochrome oxidase subunit 1 gene were much useful for divergence between species and it suggests that this is greatly applicable throughout the phylogenetically distant groups, concerned regarding the Cytochrome oxidase subunit 1 barcodes from Indian waters that efficiently differentiate the divergent species (Hebert et al., 2003). DNA barcoding to provide an effective method for identification of species by using an array of specific molecular tags Cytochrome C oxidase subunit 1 gene (Pradhan et al., 2015).

Identification of fish species based upon the particular sequence of mitochondrial DNA fragment, which help in the correct differentiation of unidentified species (Dawnay et al., 2007). In this study, mitochondrial cytochrome oxidase subunit I genes was used for molecular identification of freshwater *Oreochromis niloticus*. Nearly 650 base pair of Cytochrome C oxidase subunit 1 gene is used for purpose of specie identification. Present study recorded 660 base pair Co1 length in *Oreochromis niloticus*

The comparison of absorbance at 260/280 nm provide a DNA/protein relationship of 1.6 to 2, for pure samples (Cawthorn et al., 2011, Naeem and Shoaib, 2019). Results of the present study revealed that absorbance values found in ranged, which are in general agreement with study.

Consequent analysis of sequence of DNA barcode

permitted vibrant differentiation among taxonomic status of all examined species. The mean intraspecific K2P genetic distance was shown in Table 3. Generally, the average K2P distance between the same specie and same genus were approximately reliable with the observed range among the freshwater Indonesian fishes (Hubert et al., 2015), freshwater fishes of Canada (Hubert et al., 2008), and Australian freshwater fishes (Ward et al., 2005), Indian freshwater fish than (Larkra et al., 2015), Pakistan (Naeem and Shoaib, 2019). This indicates that the specie analyzed in this study can be well through DNA barcode. The phylogenetic tree generated in this work is as with preceding study of (Yu et al., 2007), for which different genes of mtDNA were used for analyses of target portion of the COI gene. The primer pair used suggests that the COI gene has a potentially high-resolution power (Kan et al., 2007) to discriminate among species and populations of the same species (Jun et al., 2011, Khaliq, 2012). The entire phylogenetic methods (NJ and PhyML) tree derivative from this studied. Similarity of banding outline is the standards to approximation of their variation genetically and diversity of their nucleotide (Clark and Lanigan, 1993). Present study also used to derived NJ tree and so.

The differences appeared significant result emerging from the simulation study and shown when comparing several distance and phylogenies (Ross et al., 2008). Confirm it through that case by comparison through classification statistically. Among phylogenetic and NJ outperforms the maximum-likelihood method (Elias et al., 2007). While algorithms like Phylogram maximum-likelihood might be proficient method to determine roots in phylogenetical trees among separated taxa. Rules have applied for the phylogenetic is relatively rapid and set aside to execute data comparisons (Munch et al., 2008)

Jurado-Rivera et al. (2009) likewise, we consider species with geographical species Structuring genetically across individuals allowed to do barcoding in cases where the Cytochrome c oxidase I locus is useless due to biological peculiarity of mitochondria by Wolbachias 2007. These patterns are confirmed by the study on data as apparent on the phylogenetic tree from Hebert et al. 2004, as the present study also used these all types of method order to countdown the the genetic diversity among species.

In current work used the molecular basis for identification of species rather than the visible external morphology. It was not as familiar practice in Pakistan. This study also enlightens the DNA barcoding work importance as a gateway for the fish identification. Addition to that the method of DNA barcoding can distinguish between species of fishes analysed that would clearly accomplished without any unambiguously identifying individual species.

CONCLUSION

DNA barcoding a functional fish species identification technique compares to the different conventional methods of identification based on the visible external morphology. Previously used identification techniques have several limitations as it don't work in developmental body stages of fish, processed, fillet and in case of specimen damaged. DNA barcoding is based on a Cytochrome C oxidase subunit 1 gene of mitochondrial DNA, have enough variability to differentiate the species. It is also used to estimate divergence among specie, genus, and family. Moreover, DNA barcoding in the field of taxonomy is useful tool for fish identification and diversity because it enhanced focus on data standardization and validation.

Conflict of Interest

Author declares no conflict of interest.

REFERENCES

Bernatchez L (1991). A role for molecular systematic in defining evolutionary significant units in fishes. in: J.L. Nielsen (Ed.), In Evolution and the Aquatic Ecosystem: Defining Unique Units in Population Conservation, American Fisheries Society, Bethesda, Maryland. 114–132.

Cawthorn DM, Steinman HA, Witthuhn RC (2011). Comparative study of different methods for extraction of DNA from fish species commercially available in South Africa. Food Con., 22,231-244.

Clark AG, Lanigan CMS. (1993). Prospects for estimating nucleotide divergence with RAPDs. Mol. Biol. Evol., 10, 1096–1111.

Dawnay N, Ogden R, McEwing R, Carvalho GR, Thorpe RS (2007). Validation of the barcoding gene COI for use in forensic genetic species identification. Forensic Sci. Int., 173,1–6.

Elias M, Hill RI, Willmott KR, Dasmahapatra KK, Brower AV, Mallet, J, Jiggins CD (2007). Limited performance of DNA barcoding in a diverse community of tropical butterflies. Proc. R. Soc. B., 274,2881–9.

Hebert PD, Cywinska A, Ball SL (2003). Biological identifications through DNA barcodes. Proc. Biol. Sci. R. Soc., 270 (1512), 313–321. http://dx.doi.org/10.1098/ rspb.2002.2218.

Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burridge M, Bernatchez L (2008). Identifying Canadian Freshwater Fishes through DNA Barcodes. PLoS ONE 3(6), e2490.

Hubert N, Kadarusman Wibowo A, Busson F, Caruso D, Sulandari S, Nafiqoh N, Pouyaud L, Rüber L, Avarre J, Herder F, Hanner R, Keith P, Hadiaty RK (2015). DNA Barcoding Indonesian freshwater fishes: challenges and prospects. De Gruyter Open., 3, 144–169.

Jun J, Han SH, Jeong TJHC, Lee B, Kwak M (2011). Wildlife forensics using mitochondrial DNA sequences: Species identification based on hairs collected in the field and confiscated tanned Felidae leathers. Gen. & Gen. 33,721–726.

Jurado-Rivera JA, Vogler AP, Reid CA (2009). Petitpierre E and GomezZurita J: DNA barcoding insect-host plant associations. Proc. Biol. Sci., 276,639–48. 31.

Kan XZ, Yang JK, Li XF, Chen I, Lei ZP, Wang M, Qian CJ, Gao H, Yang ZY (2010). Phylogeny of major lineages of galliform birds (Aves: Galliformes) based on complete mitochondrial genomes. Gen. Mol. Res. 9,1625–1633.

Khaliq F (2012). Isolation and sequencing of cytochrome oxidase C subunit 1 (COI) gene of Tor putitora. Thesis, University of Sargodha, Pakistan.

Lakra WS, Singh M, Goswami M, Gopalakrishnan A, Lal KK, Mohindra V, Sarkar UK, Punia PP, Singh KV, Bhatt JP, Ayyappan S (2015). DNA bar- coding Indian freshwater fishes. Mit. DNA., 27(6):1–8.

Laskar BA, Bhattacharjee MJ, Dhar B, Mahadani P, Kundu S, Ghosh SK (2013). Species Dilemma of Northeast Indian Mahseer(Actinopterygii: Cyprinidae): DNA bar., 284.

Liu Z, Li P, Argue BJ, Dunham RA. (1998). Inheritance of RAPD markers in channel catfish (Ictalurus punctatus), blue catfish (I. furcatus) and their backcross hybrids. Anim. Genet., 2 (29), 58–62.

Menezes GM, Sigler MF, Silva HM, Pinho MR (2006). Structure and zonation of demersal fish assemblages off the Azores archipelago (mid-Atlantic). Mar. Ecol. Prog. Ser., 324,241–260.

Munch K, Boomsma W, Huelsenbeck JP, Willerslev E, Nielsen R (2008). Statistical assignment of DNA sequences using Bayesian phylogenetics. Syst Biol., 57,750–7. 32.

Naeem M, Hassan S (2019). Molecular approach for identification of labeo bata. Mit. DNA part B., 416-421.

Naylor G, Brown WM (1997). Structural biology and phylogenetic estimation. Nature., 388 (1997) 527–528.

Palumbi SR, Cipriano F (1998). Species identification using genetic tools: The value of nuclear and mitochondrial gene sequences in whale conservation. J. Hered., 89 (5), 459–464.

Pradhan V, Kamble Y, Ladniya V, Mogul M (2015). An overview of Species Identification by DNA Barcoding. Int.J.Curr.Microbiol. App. Sci., 4(4): 127-140.

Ross HA, Murugan S, Li WLS (2008). Testing the reliability of genetic methods of species identification via simulation. Syst Biol., 57,216–30.

Schindel DE, Stoeckle MY, Milensky CM, Trizna M, Schmidt BK, Gebhard CA, Graves GR. (2011). DNA barcodes of bird species in the National Museum of Natural History, Smithsonian Institution, USA. In: Zoo.Keys., 152,87–91.

Seifert KA, Samson RA, Dewaard JR, Houbraken J, Levesque CA, Moncalvo JM, Louis-Seize G, Hebert PD (2007). Prospects for fungus identification using CO1 DNA barcodes, with Penicillium as a test case. Proc. Natl. Acad. Sci. USA., 104,3901–6.

Trivedi S, Affan R, Alessa AHA (2014). DNA barcoding of Red Sea fishes from Saudi Arabia – the first approach. DNA Bar., 2,17–20. http://dx.doi.org/10.2478/dna-2014-0003.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005). DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B: Biological Sci., 360(1462), 1847–1857.

Yu L, Li QW, Ryder OA, Zhang YP (2007). Analysis of complete mitochondrial genome sequences increases phylogenetic resolution of bears (Ursidae), a mammalian family that experienced rapid speciation. BMC Evo. Bio., 7,198.