The Journal of Animal & Plant Sciences, 25 (3-Supp. 2) 2015 Special Issue Page: 261-265
ISSN: 1018-7081

# CYTOCHROME OXIDASE SUBUNIT 1 (COI) GENE BASED PHYLOGENETIC ANALYSIS OF MAHSEER (*TOR PUTITORA*) OF PAKISTAN

F. Khaliq[1], N. Shafi[2], A. Rehman[3] and S. Janjua[3,*]

[1]Department of Zoology, Institute of Natural and Management Sciences, Rawalpindi, Pakistan,
[2]University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan,
[3]Bioresource Research Centre (BRC), Islamabad, Pakistan.
*Corresponding Author's Email: safiajanjua@hotmail.com

## ABSTRACT

Mahseer *(Tor putitora)* is a migratory freshwater fish inhabiting warm water rivers/streams which are distributed throughout South and South East Asia. The extreme inter-species similarities resulted in confusing morphological features causing problem to identify individual species, in order to clear this ambiguity scientist used many molecular markers including SSR, RAPD, RFLP and AFLP etc. However, due to problems of reproducibility and result interpretation with these markers scientists switched to new dimension of DNA barcoding using cytochrome oxidase C subunit 1 (CO1), a mitochondrial gene sequence analysis. Samples of *Tor putitora* were collected from different rivers of Pakistan and Azad Jammu Kashmir. The mitochondrial CO1 gene was amplified using gene specific primers. The PCR product was sequenced using Sanger sequencing method and phylogenetic analysis was done utilizing the variant found in COI gene of Mahseers obtained from different water sources. The COI gene sequence not only confirms the species but also separate the populations of species.

**Keywords**: Mahseer of Pakistan, Cytochrome Oxidase Subunit 1 (COI) Gene, *Tor putitora*, phylogenetic analysis.

## INTRODUCTION

Like applicability of identification keys to all life stages, absence of universal standard characters across taxa, and dwindling global pool of expertise (Hebert *et al*., 2003). Barcoding provides accurate and relatively simple species identification based on the nucleotide sequence of usually one short DNA fragment, therefore is widely used for species identification of invertebrate (Costa *et al*., 2007; Mikkelsen *et al*., 2007) and vertebrate (Hebert *et al*., 2004; Hajibabaei *et al*., 2006) species. In morphology based taxonomy, the boundary between intra specific morphological variation and inter specific morphological similarities may be blurred. Additional drawbacks of morphology-based taxonomy include inapplicability of identification keys to all life stages, lack of universal standard characters across taxa and requirement of high levels of expertise among a dwindling pool of taxonomists. The approach of molecular taxonomy i.e. DNA barcoding solves many of these issues and assisted many taxonomist in identifying different species (Asgharian *et al*., 2011). DNA barcoding was introduced to overcome some of the shortcomings of classical taxonomy.

Fish of the genus *Tor* (commonly called as "mahseers") are highly sought due to their high commercial and recreational values (Ng, 2004). Twenty species are currently recognized within this genus, occurring throughout Asia, from the trans-Himalayan region to the Mekong River Basin to Malaysia and Indonesia. Over-exploitation of the natural stocks and environmental degradation has resulted in a rapid decline of mahseers in the wild (Ogale, 2002).

In Pakistan, Poonch River is found to be the most promising area for Mahseer (*T. putitora*) (Anjum, 2013). This river has highest and stable population of Mahseer in the country. The second promising area for Mahseer is the Lower Swat River in the Malakand Agency and Mardan District. In river Hingol (Baluchistan), Mirza (1994) reported the presence of *T. putitora* along with other species of Cyprinidae.

The present study uses the classical taxonomy for identification of species of one of the important fish *T. putitora* and validated its identification with advance molecular taxonomy; COI gene and further use the sequence data to study the genetic variation between different selected populations.

## MATERIALS AND METHODS

**Chemicals and Reagents:** All the chemicals and reagents used in the study were purchased from Sigma else mentioned along the chemical.

**Sample Collection:** Twenty two fish samples were collected from different potential sites (Figure 1). Six samples were from Population A: Swat River (KPK) and Population B: Indus River (KPK) each while five samples were from Population C: Poonch River (AJK) and Population D: Hingol River (Baluchistan) populations each. The collected samples were identified on the basis of identification key provided by Mirza (2004). The fish samples were transported in ice boxes to the lab where

muscles tissues were taken out and stored at -20 ºC until further experimentation.

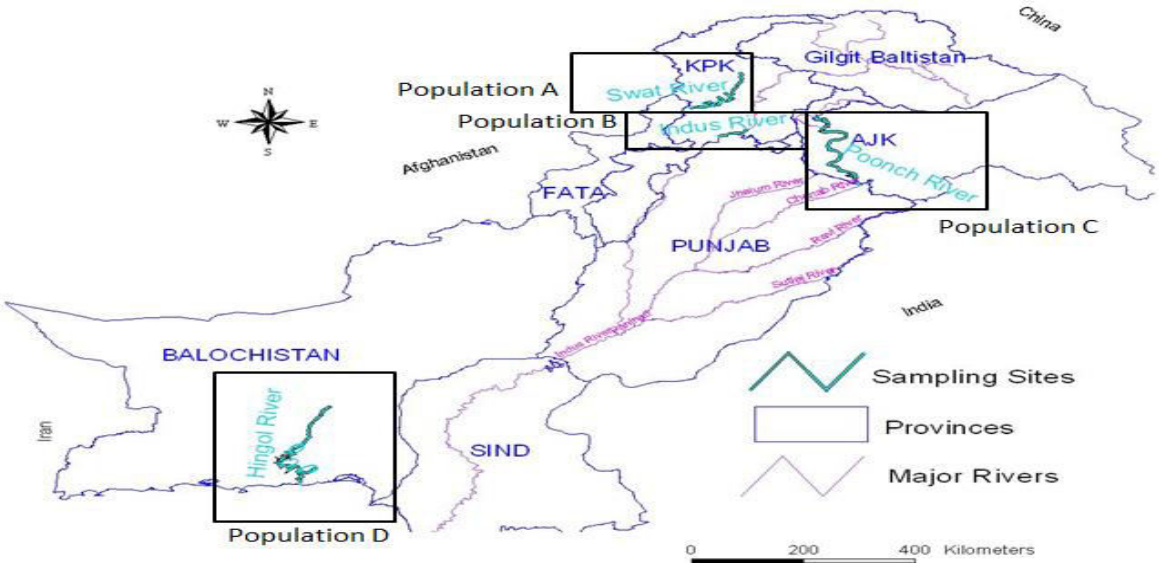


**Figure 1: Map showing the four sampling sites.**

**DNA Extraction:** The DNA was extracted from frozen muscle tissue by Phenol-Chloroform protocol as described by Sambrook and Russell (2001) with slight modifications. Briefly, 500 µL lysis buffer (50 mM Tris, 50 mM EDTA, 200 mM NaCl, 1% SDS) was added into 100 mg tissue maintained at 37 ºC for half hour for thawing of frozen samples. Then 25 µL proteinase K and 25 µL DTT was added and mixture was incubated at 56 ºC for 3 hrs followed by centrifugation (SIGMA 1-14) at 10, 000 rpm for 10 min and supernatant was transferred to fresh tubes. 300 µL phenols was added to supernatant, mixed vigorously, and centrifuged at 12, 000 rpm for 10 min and aqueous phase was transferred to fresh tube. Equal volumes of Phenol: Chloroform: Isoamyl alcohol (20: 24: 1) mixture was added to aqueous phase, and centrifuged again under aforementioned conditions. Again to the aqueous phase equal volume of chloroform and isoamyl alcohol (24:1) mixture was added and centrifuged at 12,000 rpm for 10 min. Aqueous phase was transferred to fresh sterilized tube and for DNA precipitation equal volume of isopropanol with 25 µL Na- acetate (3 M) was added and incubated at –20 ºC for 30 min. Then DNA was pellet down by centrifugation at 10, 000 rpm for 3 - 5 min. 70 % ethanol was used for washing of pellet and then 40 µL of RNase / TE buffer was added to DNA pellet and incubated at 37 ºC for an hour. Extracted DNA sample was stored at -20 ºC after confirmation by 1% agarose gel electrophoresis.

**Polymerase Chain Reaction (PCR):** We amplified 650 bp COI gene fragment using primer pair of Ward *et al*. (2005):

FishF1 (5'-TCAACCAACCACAAAGACATTGGCAC-3') and
FishR1 (5'-TAGACTTCTGGGTGGCCAAAGAATCA-3').

The PCR was carried out (Thermo scientific, PCR Sprint thermocycler) in volume of 25 µL having 50–75 ng of genomic DNA, 1X Taq buffer (10 mM Tris-HCl, pH 9.0; 50 mM KCl; 0.01% triton X-100), 2.0 mM $MgCl_2$, 0.2 mM of each dNTP, 5 pmol of each primer and 1.5 units of Taq DNA polymerase. We followed standard cycle conditions with annealing at 54 ºC for a min. We confirmed PCR product by electrophoresis on 1.5% agarose gel, and amplicon purified using GeneJET™ PCR purification Kit (#K0702, Fermentas, USA). PCR product was sequenced using Sanger sequencing method by utilizing commercial sequencing facility (MACROGEN Inc, Seoul, Korea).

**Phylogenetic Analysis:** We performed phylogenetic analysis of COI gene sequence data using MEGA5 computer program (Tamura *et al*., 2011). The evolutionary history was obtained by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 4 nucleotide sequences (each of around 650 bp), representative of each population.. Codon positions included were $1^{st}$ +$2^{nd}$ +$3^{rd}$ +Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 648 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura *et al*., 2011).

## RESULTS AND DISCUSSION

The Figure 2 shows the gel picture of 650 bp amplified gene fragment of COI. We analyzed genetic differences between different populations based upon 650 bp COI gene sequences (Table 1, Figure 3). We suggest no genetic difference between three populations, *viz*. Swat River, KPK, Indus River, KPK and River Poonch, AJK. However 0.5% difference was observed between these three populations and River Hingol, Baluchistan. Alignment of four population is shown in figure 4. This difference is justified by river system of Pakistan. Fish samples were collected from the

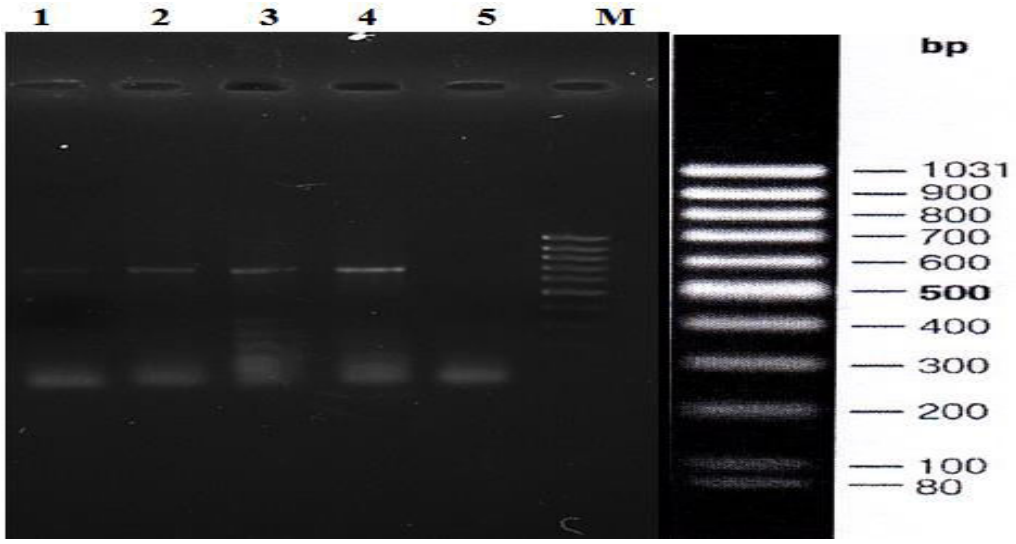


**Figure 2: Agarose gel of PCR product of 650 bp fragment of COI gene amplified in four populations. Lane 1: Swat River, Lane 2: Indus River, Lane 3: Poonch River, Lane 4: Hingol River, Lane 5: Negative Control and Lane M: DNA Ladder.**

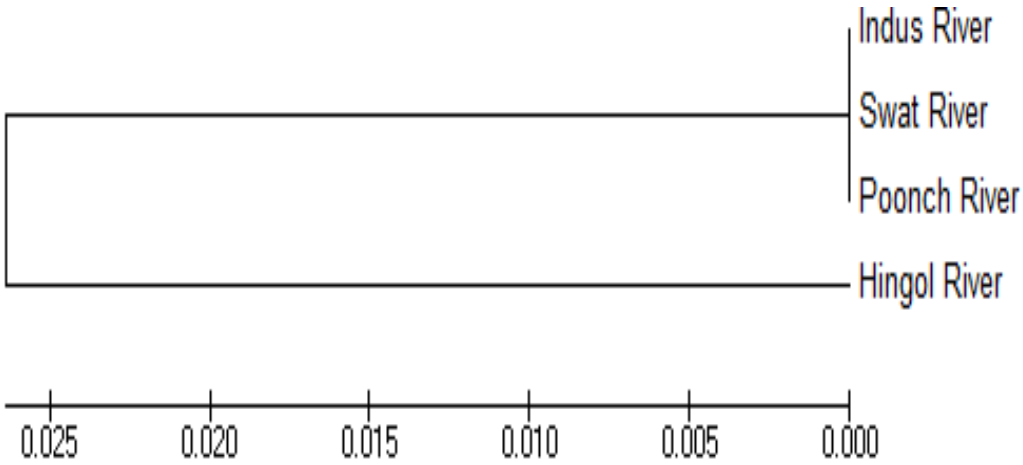


**Figure 3: Phylogenetic tree generated by MEGA5. Indus, Swat and Poonch River populations form single cluster while Hingol River population falls is separate cluster.**

**Table 1. Genetic distances between different *Tor putitora* populations based upon sequence data of COI gene.**

| | Swat | IndusS | Poonch | Hingol |
|---|---|---|---|---|
| **Swat** | - | - | - | - |
| **Indus** | 0.000 | - | - | - |
| **Poonch** | 0.000 | 0.000 | - | - |
| **Hingol** | 0.005 | 0.005 | 0.005 | - |

upper side of the River Indus, which starts off from the glacial Tibetan mountains, and near Attock, Swat and Kabul rivers join the Indus. While River Poonch, Jhelum, Chenab, Ravi and Sutlej all converge at the Indus at Panj Nadh in the south (The Trend-line of the Himalayas, 1936). This river sharing of the mighty Indus river system ensures commonality of aquatic fauna, thus gives rise.

**Figure 4: Clustal alignments of four sequnces, the varialbe sites are shaded in white.**

To the genetic flow without much hindrances (The Trend-line of the Himalayas, 1936). However, ecological barriers as well as river diversions, channelizing of water courses and construction of dams etc, these obstructions also provide chance for adaptations to some aquatic species and several genetic separation/diversity.

Baluchistan coastal drainage system consists of a number of relatively small and shallow rivers, namely Hub, Porali, Hingol, Basol and Dasht. All these rivers emerge from southwestern hills and independently fall into the Arabian Sea. According to report of Mirza (1994) the fresh and cold water of the Hingol River has *Tor putitora*.

The four populations were selected three from upper side of Indus and its confluent rivers Swat and Poonch. Fourth population was collected from Hingole River. The genetic diversity estimated on the basis of COI gene is vindicated for the reason that these populations are separated since Pleistocene epoch, due to Himalayan barrier, from Indus fauna (Le Fort, 1996; Rowley, 1996; Harrison *et al*., 2012).

DNA barcoding is useful tool in identification of different species of animals (Hebert et al., 2004; Ward 2009). In the present result it is also indicating that COI gene sequence can be used to separate the population of a species. According to Ward (2009) the sample showing no divergence or differs by only one or two base pairs, it is most likely the samples are of same species. Our

sample shows divergence of only 0.5 percent which is due to three transversion sites observed in the sequence. In conclusion the four populations have the same species of *Tor putitora* despite of 0.5% genetic divergence which is due to separation of the populations 30-40 million years ago.

## REFERENCES

Anjum, A. (2013). A new species and first host record of *Rhabdochona* Ralliet, 1916 (Nematoda: Thelazoidea) in fresh water fish *Tor Putitora* from Poonch river of J & K state, India. IJAR, 3: 110-115.

Asgharian, H., H. H. Sahafi, A. A. Ardalan, S. Shekarriz and E. Elahi (2011). Cytochrome c oxidase subunit 1 barcode data of fish of the Nayband National Park in the Persian Gulf and analysis using meta-data flag several cryptic species. Mol. Ecol. Resour. 11: 461–472.

Costa, F. O., J. R. deWaard, J. Boutillier, S. Ratnasingham, R. T. Dooh, M. Hajibabaei and P. D. Herbert (2007). Biological identifications through DNA barcodes: the case of the Crustacea. Canadian J. Fisheries Aquatic Sci. 64: 272- 295.

Hajibabaei, M., M. A. Smith, D. H. Janzen, J. J. Rodriguez, J. B. Whitfield and P. D. N. Hebert (2006). A minimalist barcode can identify a specimen whose DNA is degraded. J. Mol. Ecol. Notes, 6: 959– 964.

Harrison, T. M., M. Grove, O. M. Lovera and E. J. Catlos (2012). A model for the origin of Himalayan anatexis and inverted metamorphism. J. Geohys. Res. 103: 27017-27032.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen and W. Hallwachs (2004). Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. PNAS. 101 (41): 14812– 14817.

Hebert, P. D. N., A. Cywinska, S. L. Ball and J. R. deWaard (2003). Biological identifications through DNA barcodes. Proc. R. Soc. Lond. 270: 313– 332.

Le Fort, P. (1996). Evolution of Himalaya. In: Yin, A. and Harrison T. M. Tectonic Evolution of Asia. New York Cambridge university press. pp 95-109.

Mikkelsen, N. T., C. Schander and E. Willassen (2007). Local scale DNA barcoding of bivalves (Mollusca): a case study. J. Zoologica Scripta, 36 (5): 455- 463.

Mirza, M. R. (1994). Geographical distribution of freshwater fishes in Pakistan: a review. Punjab Univ. J. Zool. 9: 93-108

Mirza, M. R. (2004). Freshwater Fishes in Pakistan, 2nd edition. Urdu Science Board, Lahore, Pakistan.

NG, C. K. (2004). King of the rivers: mahseer in Malaysia and the region. Inter sea fishery. SDN BHD, Kuala Lumpur Malaysia, pp.170.

Ogale, S. N. (2002). Mahseer breeding and conservation and possibilities of commercial culture. The Indian experience. In T. Petr and D.B. Swar (eds.) Cold Water Fisheries in the Trans-Himalayan Countries. FAO Fish. Tech. Pap. 431.

Rowley, D. B. (1996). Age of Initiation of collision between India and Asia; A review of stratigraphic data. Earth and Plantery science letters. 145:1-13.

Sambrook and Russell (2001). Molecular Cloning: A Laboratory Manual (3rd ed.). Cold Spring Harbor Laboratory Press.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei. and S. Kumar (2011). MEGA 5: Molecular evolutionary genetics analysis using Maximum likelihood, Evolutionary distance and Maximum Parsimony methods. J. Mol. Biol. Evol. 28: 2731- 2739.

Tamura, K. and M. Nei (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. J. Mol. Biol. Evol. 10 (3): 512-526.

The Trend-line of the Himalayas (1936). Himalayan Oxford Journal, 8: p. 63.

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. Mol. Ecol. Resour. 9: 1077–1085.

Ward, R. D., T. S. Zemlak, B. H. Innes, P. R. Last and P. D. N. Hebert (2005). Barcoding Australia's fish species. Phil. Trans. R. Soc. B. 360: 1847−1857.