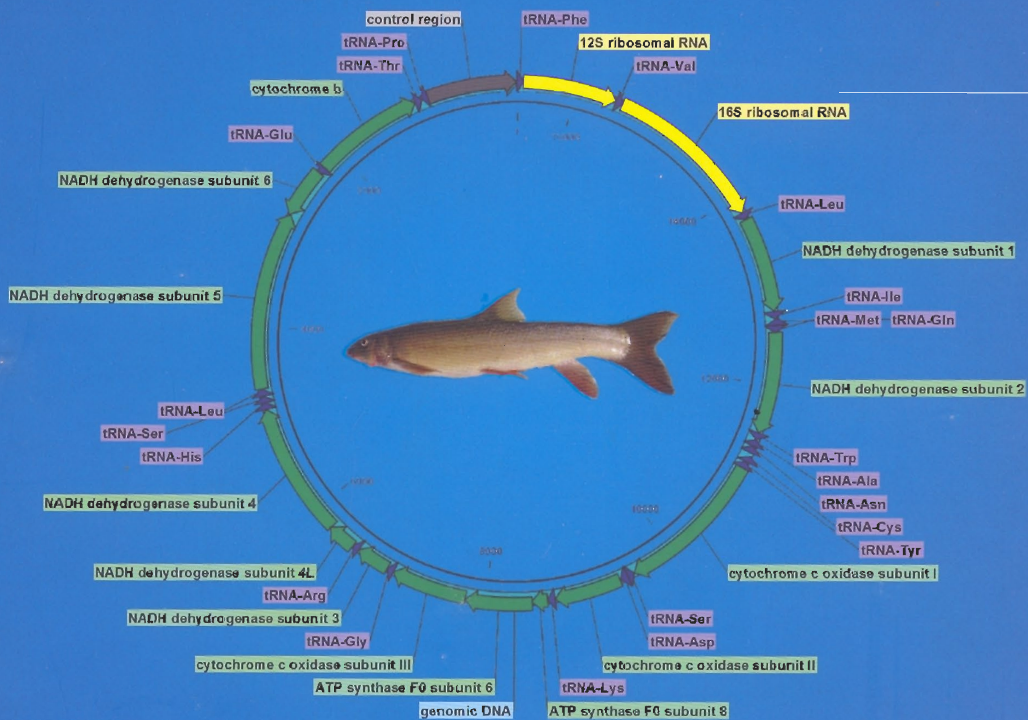
Bulletin No 25

Mitochondrial Genomics of Schizothoracid Fishes



ICAR-Directorate of Coldwater Fisheries Research
(Indian Council of Agricultural Research)
Bhimtal - 263136, Distt. Nainital, Uttarakhand

Mitochondrial Genomics of Schizothoracid Fishes

Compiled and edited by

P. K. Sahoo | Siva C | S. Ali | L. Singh | R. Kumar | A.K. Singh

Bulletin No. 25

2016

ICAR-Directorate of Coldwater Fisheries Research
(Indian Council of Agricultural Research)
Bhimtal – 263136, Distt. Nainital, Uttarakhand

Mitochondrial Genomics of Schizothoracid Fishes

© ICAR- DCFR 2016

Material contained in this bulletin may not be reproduced in any form without the permission of ICAR- DCFR.

Published by:
Dr. A. K. Singh
The Director, ICAR- DCFR, Bhimtal (Nainital), Uttarakhand.

Lasertypeset & Printed by:
M/s Royal Offset Printers, A-89/1, Naraina Industrial Area, Phase-I, New Delhi-110028

Preface

Mitochondrial genomic data generated over past three decades have impacted the scientific fields significantly. Next-generation sequencing (NGS) techniques and sophisticated bioinformatics programs have backed it. It is hard to imagine the area of molecular evolution as it stands today without the contributions from mitochondrial genomics studies. Different mitochondrial genes serves as the best marker for species identification due to its universal primer's robustness and the greater range of the phylogenetic signal. Phylogenetic analysis based on complete mitogenomes yields more reliable phylogenies compared to single or partial mtDNA data based phylogenies. Intraspecies mitochondrial genome comparisons will be vital for future research in labeling the beneficial mitochondrial genes of economictraits. The depiction of complete mitogenomes is an important modern challenge for various aspects of genetics and precursor for further applied cold-water fisheries research. Keeping the well-recognized advantages of the mitochondrial genome in view, the present comparative account has been compiled on six economically important species.

The information presented is based on the sustained efforts of our scientists involved in such projects time to time. All the personnel of the Institute, who contributed in the preparation of this document, deserve appreciation. I hope this document will be baseline information for different researchers to initiate work on mitochondrial genomics of cold water fisheries in the country about which very little is known.

(A.K. Singh)
Director, ICAR-DCFR

Contents

Preface		iii
1	Introduction	1
2	Schizothorax species and it morphological identification	2
3	Sample collection and DNA extraction	5
4	PCR amplifications and DNA sequencing	6
5	Assembly and annotation of the mitochondrial genome	6
6	Phylogeny construction	6
7	Characteristics of whole Mitochondrial Genomes of Six fish species	7
8	Application of mitochondrial genome in Fish species identification	12
9	Phylogenetic Analyses	13
10	Contribution of ICAR-DCFR	15
11	Conclusion	15
12	References	16

INTRODUCTION

In general, most of the animals have their DNA packaged inside the nucleus. Since many years, this nuclear DNA is the primary research interest to the scientific community, but other than this nuclear DNA, the mitochondria with a small amount of its own DNA has also emerged as a strong genetic marker. At present mitochondrial (mt) DNA is the most sequenced type of eukaryotic chromosome (Smith & Keeling, 2015). As of June 2016, there are more than 6462 complete mtDNAs sequences documented in the National Center for Biotechnology Information Databank. The rate of mitochondrial genome sequencing is growing exponentially. In 2014, more than a thousand mtDNAs were deposited in GenBank, which is twice the number from 2012 and almost four times that from 2010 all this informations are indicating the importance of mitochondrial genomics. Mitochondrial genomic data generated over past three decades have impacted the scientific fields significantly. It has given rise to leading hypotheses on evolution and initiated intense debates about the roles of adaptive versus non-adaptive processes in shaping organismal and genomic complexity (Smith, 2015). It is hard to imagine the field of molecular evolution as it stands today without the contributions from mitochondrial genomics studies. Our understanding of eukaryotic life, its origins, its diversity and its complexity have all been shaped by studies of mtDNA.

Next-generation sequencing (NGS) techniques and sophisticated bioinformatics programs have made it quick, easy and cheap to sequence and assemble entire mitochondrial genomes from almost any eukaryotic species. In most instances, we do not even need to purify mitochondria or do tedious mtDNA isolation through cesium chloride density gradient centrifugation before sequencing. A single run of whole genomic DNA on an NGS platform yields enough mtDNA-derived reads to assemble the complete mitochondrial genome. It has been adopted widely as a popular marker because it is exclusively inherited maternally with the rapid rate of evolution when compared to nuclear DNA. It has alower rate of decay relative to nuclear DNA, it can be amplified easily and strongly conserved among vertebrates with little duplication, no introns, and very short intergenic regions (Gissi et al, 2008). This fast rate of mtDNA evolution joined with multiple copies in the cell and maternal

inheritance have made mtDNA an extremely useful tool for high-resolution analyses of recent evolutionary events, level of gene flow, hybrid zones, population structure, founder events, female-mediated gene flow, study of the origin of clonally reproducing species, population-level studies, broad-scale comparative analyses, phylogenetic approaches archeological and forensic study (Meysr, 1993).

It has been suggested that variation can be genetically related or associated with phenotypic plasticity in response to environmental conditions (Murta, 2000). The intra-species comparisons will be vital for future research in determining which mitochondrial genes and/or sequences result in beneficial phenotypes for important physiological traits, such as temperature tolerance, feed conversion, growth and development rate (Saitoh et al, 2006). Thus, the depiction of complete mitogenomes is an important modern challenge for various aspects of genetics and precursor for further applied cold-water fisheries research.

By keeping the well-recognized advantages of mtDNA sequences, the present comparative account has been compiled on six economically important species under two closely related genera *Schizothorax* and *Schizopygae* (Fam: Cyprinidae).

2. Schizothorax species and it morphological identification

The cyprinid fishes of the composite genus *Schizothorax* Heckel, 1838, inhabit the mountain streams usually above an altitude of 670 meters along the Himalayan range. They are of great economic importance as food fishes. The genus *Schizothorax* was established by Heckel (1838), species comes under this genus are identified mainly based on the presence and absence of suctorial mouth, nature of the lip, pattern of colors, standard length, body depth, scale count, number and relative position of fins (Strauss and Bond 1990).

Heckel (1838) grouped ten species of genus *Schizothorax* in three groups, "A", "B","C"; the first group characterized by a sucker at the chin and the latter two groups without it. This situation has led to the description of several genera and uncertain taxonomic history. Undoubtedly much of the confusion arose due to misinterpretation of the type-species of the genus Schizothorax. So it is necessary to identify the fish species very meticulously using the following morphological identification keys provided in following Table 1.

Table 1. Morphological identification key

	Schizothorax richardsonii	*Schizothorax labiatus*	*Schizothorax esocinus*	*Schizopyge niger*	*Schizothorax progastus*	*Schizothorax plagiostomus*
Color	Steel grey, becomes gradually lighter below: belly yellowish white, body often with grey spots. Dorsal and caudal fin greyish white, Pectoral fin and anal fins ochre yellowish	Dark brown with black specks on back, yellowish-white below. Fins pinkish-yellow.	In live condition Silvery grey with numerous dark, small, irregular spots on back and flanks of thebody, Fins silvery grey with similar dark spots, more numerous at their bases.	The fish is much darker in color than other species of the genus. In alcohol, dark brown on theback, lighter on flanks, belly pale yellowish; spots on upper part of body. Fins brownish.	Uniformly silvery often with a few finespots, wFinsdark-edged	In live condition Silvery grey color
Body	Body streamlined, body depth 4.1 to 6.2 times in standard length,	Body elongate, narrow and sub cylindrical, Its depth 4 to 5.1 times in standard length.	Body streamlined, body depth 4 to 5.9 times in standard length	Body elongated and compressed, its depth about 5.5 times in standard length.	Body elongate and subcylindrical, Body depth is 3.8 to 4.9 times in standard length	Elongated subcylindrical body with short
Head	Head length 4 to 5 times in standard length.	Head large and narrow anteriorly, its length 5.4 to 5.6 times in standard length.	Head length 4 to 5 times in standard length.	Head length 4.8 times in standard length.	Head narrow anteriorly, Snout pointed and smooth	Ventral surface of head and anterior part of body flattish, short, somewhat cone-shaped and blunt.
Eye	Eye diameter 4.1 to 5.4 times in head length.	Eye diameter 1.1 to 2 times in head.	Eye diameter 6.5 times in head length.	Eye-diameter about 4.5 times in head.	-	-
Mouth	Mouth inferior, transverse and slightly arched: hard ochre colored cartilaginous covering below lower jaw extend between corners of mouth	Snout pointed and smooth often studded with pores. Mouth subterminal, horizontal and arch- shaped, protractile	Mouth wide, horseshoe-shaped, its cleft is very deep	Mouth moderate, Horseshoe shaped	Mouth subterminal, horizontal and arch shaped, protractile	Blunt and slightly prognathous upper jaw. Mouth inferior, wide and slightly arched
Barbells	Barbels two pairs (maxillary and rostral), usually shorter than eye diameter	Barbles two pairs; both maxillary and rostral barbles are longer than the eye.	Barbells two pairs, Rostral pair about 1.5 times longer than eye diameter, Maxillary pair slightly shorter	Barbels two pairs (maxillary and mandibular).	Barbels two pairs, much longer than eye diameter	Barbless two pairs. Pharyngeal teeth present in three rows.

	Schizothorax richardsonii	***Schizothorax labiatus***	***Schizothorax esocinus***	***Schizopyge niger***	***Schizothorax progastus***	***Schizothorax plagiostomus***
Lips	Lower lip is fleshy and flat. It is covered with a set of raised papillae forming the sucker	Lips thick and fleshy; lower labial fold uninterrupted and tri-lobed, median lobe well-developed and free at its tip.	The lower lip is fleshy; lower labial fold interrupted in themiddle.	Lips marginally sharp and thin; lower labial fold interrupted.	Lips thick and fleshy, Lower labial fold uninterrupted and tri-lobed, medial lobe is significantly small	lips fleshy and continuous, marginally sharply attenuated lower lip papillae and reflected from thejaw, themargin of lower lip sharp, covered with firm and hard horny cartilage; a strip of papillae labial plate at chin present
Dorsal fin	Dorsal fin inserted almost midway between snout tip and base of caudal fin, Dorsal spine strong and serrated behind	Dorsal spine is strong.	Dorsal fin inserted slightly nearer to the base of caudal fin than to snout tip, Dorsal spine equal to head excluding snout. Anal fin, when laid flatdoesn't reach the base of the caudal fin in adult but reaches in juveniles	Dorsal fin inserted nearer to thebase of caudal fin than to snout-tip; Dorsal spine strong, serrated on hind edge. Depth of caudal peduncle 9.9 to 10.2 times in standard length.	Dorsal fin inserted nearer to snout tip or often midway than the caudal fin, Dorsal spine strong and serrated behind	Dorsal fin inserted about opposite to pelvic fins, its last undivided ray osseous, strong and serrated posterior, short than head.
Scales	Scales arevery small, lateral line with 85 to 110	Scales are small; lateral line with 100 to 110 scales.	Scales arevery small, about 104 in lateral line	Scales arevery small, 47 to 99 in thelateral line.	Scales are small, lateral line with 110 to 114 scales	Scales arevery small and elliptical
Fin formula	D iii 8: A iii 5: Pi 15-16: V i 9	D iii 8; A ii 5; P i 17; V i 10	D iii - iv 8: A iii 5: P i 18-19: V i 9	D ii 7; A iii 5; P i 17; V i 9	D iii 8: A iii 5: P i 16: V i 10	
Distribution	Along the Himalaya from Jammu and Kashmir to Assam: Sikkim:	Kashmir Valley, Indus R, Ladakh;	Indus river, its tributaries in ladakh and Kashmir valley	Kashmir Valley.	Jammu and Kashmir valley, Ganga river in Uttar Pradesh, Brahmaputra river in Assam	Tributaries throughout Himalaya
Maximum size	60 cm	30 cm	20 cm	2.7 kg	-	41.5 cm TL male

(Silas, 1960), (Jayaram KC, 1999), (Talwar PK & Jhingran G, 1991).

S.richardsonii

S. esocinus

S. labiatus

S. niger

S. progastus

S. plagiostomus

3. Sample collection and DNA extraction

- Specimens of six species were collected from the different states of India (Table 2.) and identified based on the key presented in Table 1.
- The caudal fin tissues were dissected out and preserved in 90% ethanol.
- Voucher specimens were deposited at a reference laboratory (corresponding author's address).
- The genomic DNA was extracted from fin tissue using Phenol: Chloroform: Iso-amyl alcohol method (Sambrook and Russell, 2001).

Table 2. Summary of sampling sites

SpeciesName	**Collection locality**	**GPS Coordinates**
S.richardsonii	Kosi river, Ratighat region Uttarakhand.	29° 27.488’N 79°28.812’E
S. plagiostomus	Alaknanda river, Srinagar, Uttarakhand	30°15’N 78°55’E
S. progastus	Alaknanda river, Auri Garhwal, Uttarakhand.	30° 15’ N 78° 55’ E
S. labiatus	Lidder stream, Pahalgam, Jammu and Kashmir	34° 0’ 45”N 75° 18’ 56”E
S. esocinus	Dal Lake Srinagar, Jammu and Kashmir	34° 7’ N 74o 52’E
S. niger	Dal lake, Srinagar, Jammu and Kashmir	34° 7’ N 74o 52’E

4. PCR amplifications and DNA sequencing

- The mtDNA was selectively amplified from the genomic DNA using REPLI-g Mitochondrial DNA Kit (Qiagen, Venlo, NL) following the manufacturer's protocol.
- Complete mtDNA was amplified as concatenated sequences using selectively amplified mtDNA template and seventeen primer pairs derived from the literature (Cheng et al, 2012; Mabuchi et al, 2006; Wang et al, 2000).

The PCR was performed in a total reaction volume of 50 μl containing 1× PCR Buffer, 200 nM of each primer, 400 μM dNTP, 1 μl template (as per kit) and 1 U Taq DNA Polymerase (Invitrogen, Waltham, USA). The reaction conditions were initial denaturation at 94 °C for 4 min; 35 cycles of denaturation at 94 °C (45 s), annealing at 48 °C–50 °C (45 s) and extension at 72 °C (90 s) followed by a final extension at 72 °C for 10min. Annealing temperatures were varied within the above-listed ranges to optimize the efficiency of different primers. List of primers to be incorporated. Contiguous, overlapping fragments of the entire mitochondrial genome was sequenced using Big Dye Terminator cycle sequencing chemistry v.3.1 in ABI 3730xl automatic sequencer (Life Technologies, Carlsbad, CA).

5. Assembly and annotation of the mitochondrial genome.

- The chromatographs were proofread and assembled into contigs using CLC Genomics Workbench v.7.5.2 (CLC Bio, Aarhus, Denmark).
- Contigs so obtained were assembled on the complete mitochondrial genome of the closely related reference congeneric species.
- In addition, CLC Genomics Workbench v.7.5.2 was used for the prediction of open reading frame (ORFs), codon usage and the preparation of graphical representation of the annotated mitochondrial genome.
- The ORF Finder tool of the workbench was set to find ATG or GTG as start codons and non-triplet 3' ends of the protein-coding genes immediately adjacent to the beginning of downstream features and reading T or TA as truncated stop codons.
- Transfer RNA analysis was conducted using tRNAscan-SE v.1.21 (Schattner et al, 2005) with a cut-off score of 0.1.
- Origin of replication of the light strand (OL) and the heavy strand (OH) and conserved blocks in the non-coding control region was identified by multiple sequence alignments with annotated sequences of other species as well as by secondary structure search.
- The ribosomal RNA secondary structure was predicted using mfold web server (Zuker, 2003). Then the mtDNA sequences of six species were deposited in NCBI GenBank under the accession number exposed below.

6. Phylogeny construction

- The phylogenetic tree was constructed using mitogenome sequences of 29 species,retrieved from NCBI GenBank.

- The relationship of six species with its congeneric, related species and its taxonomic positions were delineated using the sequences of 29 species from 4 different *genera Schizothorax, Ptycobarbus, Schizopygopsis.*
- The 12 concatenated protein coding genes (excluding ND6) encoded on heavy strand were aligned using CLC Genomics v.7.5.2 with default parameters.
- The best-fitted model was selected based on hierarchical likelihood ratio test (hLRT) and Bayesian information criteria (BIC) parameter and finally, general time reversible (GTR) model was selected for phylogeny construction. UPGMA dendrogram was constructed based on maximum likelihood phylogeny (1000 replicates; maximum log likelihood = −69452.92).

7. Characteristics of whole Mitochondrial Genomes of Six fish species

The complete nucleotide sequences of mtDNA were determined for six fish species in this study. The total lengths of the mtDNA varied between 16,575 bp to 16,592 bp (Sahoo et al, 2013; Sahoo et al, 2014; Goel et al, 2013; Goel et al, 2014a; Goel et al, 2014b; Goel et al, 2014c). It is very close to the mahseer species such as *T. tor* (16554 bp) (Kumar et al, 2015),*T. putitora* (16,576 bp) (Sati et al, 2014) and *T. sinensis* (16,579 bp) in length. The base compositions in these mtDNAs (Table 3) were skewed similarly to those of other vertebrate mtDNAs (Asakawa et al, 1991), with more A–T base pairs than G–C base pairs and more A - C contents in the light-strand sequences than in the heavy-strand sequences.

Table 3. The mtDNAs base compositions of Schizothorax species

Name of the species	Accession number	A (%)	G (%)	C (%)	T (%)
Schizothorax richardsonii	KC790369	29.9	17.7	26.9	25.5
Schizothorax plagiostomus	KF928796	A30	17.4	26.8	25.8
Schizothorax progastus	KF739399	29.8	17.8	27.0	25.4
Schizopyge niger	KF600712	29.9%,	17.7	27.1	25.3
Schizothorax labiatus	KF739398	29.8	17.7	27.0	25.5
Schizothorax esocinus	KF600713	29.8	17.8	27.0	25.4

The mitogenome of our study species consists of 37 genes (2 rRNAs, 13 protein coding, and 22 tRNAs) and two non-coding regions, and the gene organization and its order are similar to other vertebrates (Anderson et al, 1981). This gene content is common to all other fishes mtDNAs so far sequenced. The details of gene order (Fig. 1) and size of mtDNA of six fishes are shown in Table 4. Most of these genes are encoded on the H-strand, except for the ND6 gene and eight tRNA genes (tRNAGln, tRNAAla, tRNAAsn, tRNACys, tRNATyr, tRNASer, tRNAGlu and tRNAPro), which are encoded on the L-strand (Fig. 1).

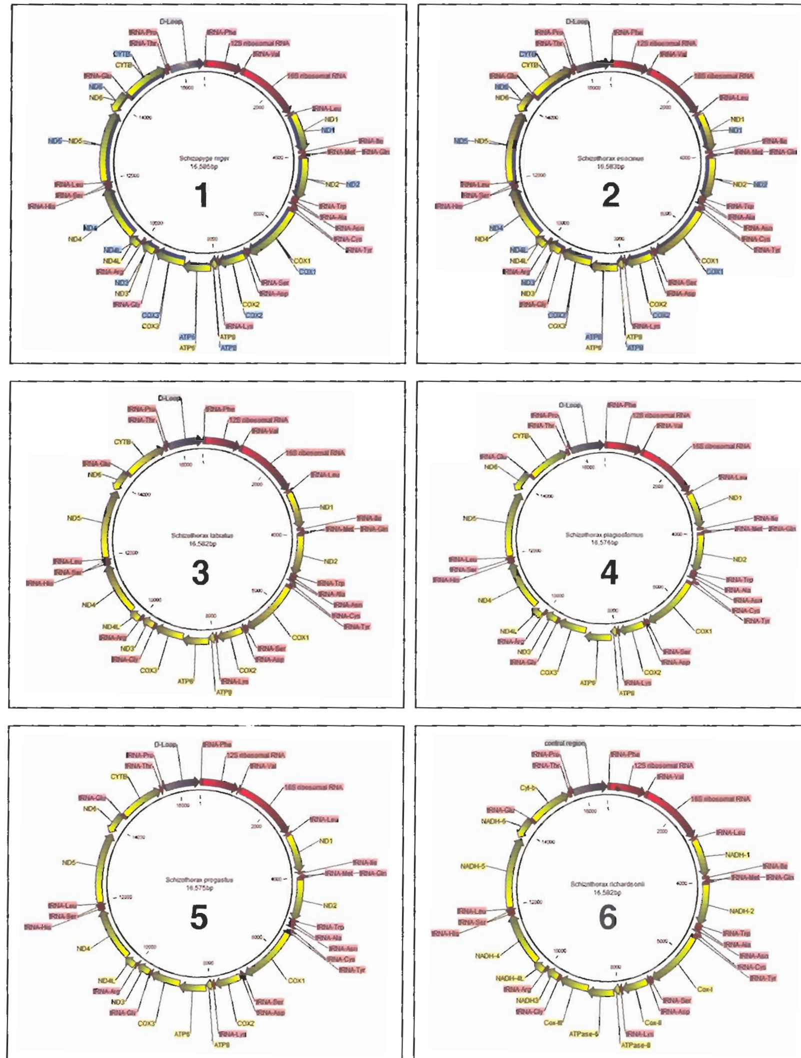


Fig.1. Complete mitochondrial genome organization and gene arrangement of six fishes.
(1. *Schizopyge niger*, 2. *S. esocinus*, 3. *S. labiatus*, 4. *S. plagiostomus*, 5. *S. progastus*, 6. *S. richardsoni*)

Thirteen protein coding genes were identified; all are similar among them in length and identical in the gene arrangements, ranging from 291 bp (ND 4 L) to 1824 bp (ND 5) accounting for 69.2% of its total mtDNA genome length. The total nucleotide length of the protein coding genes for all the six fishes was 11,409 bp. This is higher than *Neolissochilus hexagonolepis, Tor putitora, Tor tor* and *Tor sinensis* (11,391 bp) in length. Most of the protein coding genes are inferred to use ATG as start codon, whereas COI start with GTG (Table4). Ten genes terminated with TAA or TAG stop codon. The remaining six genes (ND2, COII, COIII, ND3, ND4 and Cytb) had an incomplete stop Codon "T" whereby post-transcriptional polyadenylation provides the two adenosine residues required for generating the TAA stop codon (Jondeung et al, 2007).

Twenty-two discrete tRNA genes ranging from 67 to 79 bp (Table 5) interspersed between rRNA and protein-coding genes were identified (Fig. 1). They can be folded into the standard secondary structure (Kumazawa and Nishida 1993) if G–U wobble pairs and occasional mismatches are allowed.

The control region contains several regions that are constrained in primary sequence or secondary structure to regulate replication and transcription. It is characterized by the displacement loop (D-loop), a stretch of DNA that is complementary to the light L-strand, the D-loop strand displaces the H strand. Table 5 shows the length of CR sequences for each fish species Control region of our six fish species ranges 920–935 bp in length and none among the six species exhibited identical length.

The alignment of mtDNA genomes of six *Schizothorax* species allowed the identification of the most variable mtDNA regions (Fig. 2). The average number of nucleotide differences (dxy) among sequencesof the representative species of the *Schizothorax* genus, showed informative sites for some portions of the mtDNA genome; specifically the NADH dehydrogenase genes ND6 showed high dxy value 0.056666, followed by the ND1, ND2 & CytB (dxy = 0.05466), ND4 (dxy = 0.04733), D loop (dxy =0.044) , ND5 & ND3 (dxy = 0.04066), ATP6 (dxy = 0.04), COX I (dxy = 0.02733), COX III (dxy = 0.02466), COX II (dxy =0.024), ND4L (dxy = 0.01933), ATP8 (dxy = 0.016). The genes with lowest dxyvalues were the ribosomal genes 16S (dxy =0.00933), 12s (dxy =0.008) (Fig. 2).

Table 4. Mitochondrial PCG's of Schizothorax species

S. No	Gene	Strand	Size of the PCG (bp)						Codon		Amino acid
			Schizothorax progastus	*Schizothorax plagiostomus*	*Schizothorax labiatus*	*Schizothorax esocinus*	*Schizothorax richardsonii*	*Schizopyge niger*	Start	Stop	
1	ATP 8	H	165	165	165	165	165	165	ATG	TAG	54
2	ND 4L	H	297	297	297	297	297	297	ATG	TAA	98
3	ND 3	H	349	349	349	349	349	349	ATG	TAG	116
4	ND 6	L	522	522	522	522	522	522	ATG	TAA	173
5	ATP 6	H	683	683	683	683	683	683	ATG	TAA	227
6	CO II	H	691	691	691	691	691	691	ATG	T--	230
7	CO III	H	785	785	785	785	785	785	ATG	TAA	269
8	ND 1	H	975	975	975	975	975	975	ATG	TAA	324
9	ND 2	H	1045	1045	1045	1045	1045	1045	ATG	TAG	348
10	Cyt b	H	1141	1141	1141	1141	1141	1141	ATG	T--	380
11	ND 4	H	1381	1381	1381	1381	1381	1381	ATG	T--	460
12	CO I	H	1551	1551	1551	1551	1551	1551	GTG	TAA	516
13	ND 5	H	1824	1824	1824	1824	1824	1824	ATG	TAG	607

Table 5. Mitochondrial tRNA, rRNA and control region details of Schizothorax species

Gene	Strand	*Schizothorax progastus*		*Schizothorax pagiostomus*		*Schizothorax labiatus*		*Schizothorax esocinus*		*Schizopyge niger*		*Schizothorax richardsonii*	
		Size (bp)	*Position*	*Size (bp)*	*Position*	*Size (bp)*	*Position*	*Size (bp)*	*Position*	*Size (bp)*	*Position*	*Size (bp)*	*Position*
tRNAPhe	H	69	1-69	69	1-69	69	1-69	69	1-69	69	1-69	69	1-69
tRNAVal	H	72	1027-1098	72	1026-1097	72	1027-1098	72	1025-1096	72	1026-1097	72	1030-1101
tRNAleu	H	75	2777-2851	75	2772-2846	75	2777-2851	74	2772-2845	75	2778-2852	75	2780-2854
tRNAIle	H	72	3833-3904	72	3828-3899	72	3833-3904	72	3827-3898	72	3834-3905	72	3836-3907
tRNAGln	L	71	3903-3973	71	3898-3968	71	3903-3973	71	3897-3967	71	3904-3974	71	3906-3976
tRNAMet	H	69	3976-4044	69	3971-4039	69	3976-4044	70	3970-4039	69	3977-4045	69	3979-4047
tRNATrp	H	71	5090-5160	71	5085-5155	71	5090-5160	71	5085-5155	71	5091-5161	71	5093-5163
tRNAAla	L	69	5163-5231	69	5158-5226	69	5163-5231	69	5158-5226	69	5164-5232	69	5166-5234
tRNAAsn	L	73	5233-5305	73	5228-5300	73	5233-5305	73	5228-5300	73	5234-5306	73	5236-5308
tRNACys	L	68	5340-5407	67	5334-5400	69	5340-5408	68	5335-5402	67	5341-5407	67	5343-5409
tRNATyr	L	71	5407-5477	71	5400-5470	71	5408-5478	71	5402-5472	71	5407-5477	71	5409-5479
tRNASer	L	71	7030-7100	71	7023-7093	71	7031-7101	71	7025-7095	71	7030-7100	71	7032-7102
tRNAAsp	H	79	7104-7182	79	7097-7175	79	7105-7183	80	7099-7178	79	7104-7182	79	7106-7184

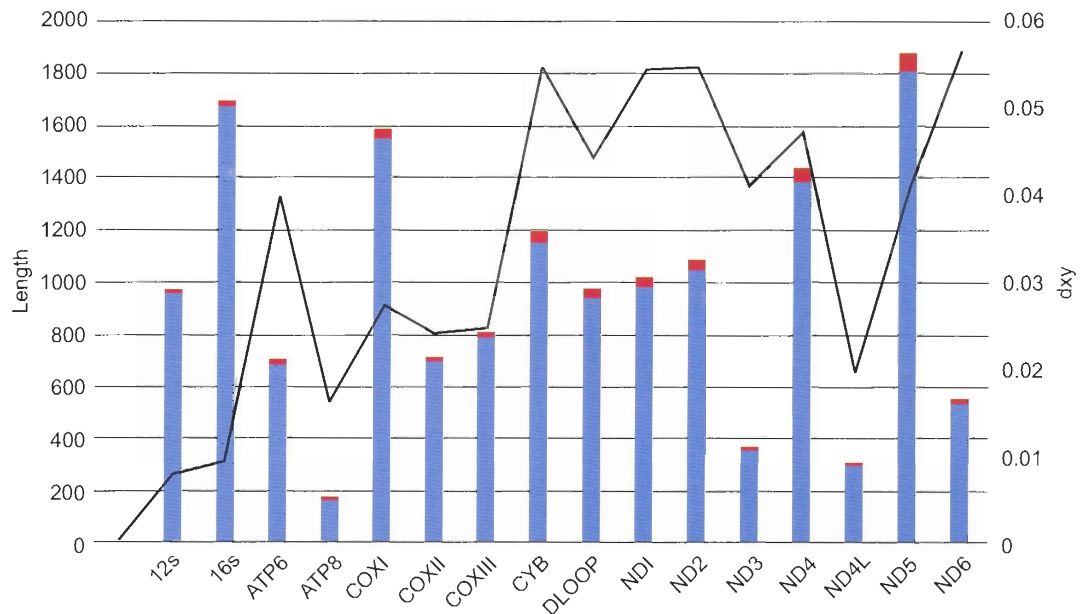


Fig 2. Bar graph of the length (number of bases) of the main mitochondrial regions and number of polymorphic sites within each region (orange) showing the average number of nucleotide differences between six species (dxy).

Till now, different authors have used different portions of the mitochondrial genome, as *Cytb, ATP6,* or the *D-loop* to assess the relations within the different marine fish genus (Chow et al, 2006), all of them with different proportion of informative sites. To understand the evolutionary relationship of different schizothorax fish species, Syed et al. 2014 examined the sequence information of mitochondrial D-loop and stated that D-loop region is highly variable. However, this study allowed the identification of NADH dehydrogenase genes ND6, ND1, ND2 & Cyt-B as the most variable mtDNA regions than the D-loop region.

8. Application of mitochondrial genome in Fish species identification

Based on morphological key characters, only a few taxonomists can critically identify more than 0.01% of the traditionally assessed species (Hammond 1992; Hawk-sworth & Kalin-Arroyo 1995), a community of 15000 taxonomists will be required, in infinity, to identify life if our reliance on morphological diagnosis is to be sustained. This morphological diagnosis has many limitations like,

1. Both phenotypic plasticity and genetic variability in the characters employed for species recognition can lead to incorrect identifications.
2. It overlooks morphologically cryptic taxa, which are common in many groups.
3. Since morphological keys are often effective only for a particular life stage or gender, many individuals cannot be identified.

4. The use of identification keys often demands such a high level of expertise that misdiagnoses are common.

The difficulty to identify species can lead to an incomplete survey of the true fauna present in a given area (Tinti et al, 2003; Schander and Willassen 2005) or to misunderstand ecological interactions (Bucciarelli et al, 2002). To overcome the above complications genomic approaches to taxon diagnosis exploit diversity among DNA sequences to identify organisms (Kurtzman1994; Wilson 1995). Over the last decades,mitochondrial gene cytochrome oxidase I (COI) serves as the best marker for species identification due to its universal primer's robustness and the greater range of phylogenetic signal than any other mitochondrial gene. Apart from this the evolution of this gene is rapid enough to allow the discrimination of not only closely allied species, but also phylogeographic groups within a single species (Cox & Hebert 2001; Wares & Cunningham 2001), and provides deeper phylogenetic insights than other marker (Simmons & Weller 2001).

Other than COI the mitochondrial 12S rRNA and/or 16S rRNA genes have been used as molecular markers to shrimp, and fish species using species-specific primers that amplify the 12S rRNA or 16S rRNA gene regions from mtDNA. The mitochondrial 12S rRNA gene served in an earlier study to differentiate snail species according to a PCR-RFLP approach. Researchers have used species-specific mitochondrial 12S rRNA and/or 16S rRNA gene primers to identify species of 26 shrimp, fish, poultry, fruit flies, and snakes (from venom), among others Yang et al, 2013.

The phylogenetic relationships and biogeography of *Schizothorax* fish species are unclear. It is often difficult to decide whether these are different species, different phenotypes of single species or an intermediate situation between the two extremes (Raina and Petr, 1999). Which are characterized by the presence and absence of a suctorial mouth size, shape and nature of the lip. Sometimes adaptation characters in various respects of their morphology may mislead the morphological characters to misidentification, due to this these fishes have had a rather long uncertain taxonomic history. Some species like *Schizothorax sinuats* and *Schizothorax nasus* have been treated as synonymous with *Schizothorax plagiostomous. Schizothorax planifrons* has been treated as a synonym of *Schizothorax niger. Schizothorax longipinnis, Schizothorax hugely* and *Schizothorax micopogon* has been treated as synonymous with *Schizothorax curvifrons* (Silas, 1960). The morphology of Kashmiri schizothracine fishes is often confusing (Raina and Petr 1999; Balkhi 2005). So it is necessary to identify the fish species accurately.

In the present study, we have presented the sequence information of all the three essential genes for six study species in Table 5. These data will be helpful to overcome the different complications in taxon diagnosis to identify organisms.

9. Phylogenetic Analyses

From the evolutionary viewpoint, mtDNAs are "small genomes" that coevolves at their own rate with the organism in which they are lodged. Thus, mtDNA sequences are widely used to construct phylogenetic trees. The most popular sequences in phylogenetic study are the

cytochrome b (Cyt b) and cytochrome oxidase 1 (COI) genes, which are utilized for the species-and-family level analysis (Kartavtsev and Lee, 2006). However, such single gene approaches provide inadequate phylogenetic data when applied to higher taxa, as a result of insufficient information capacity and homoplasy effects (Miya et al, 2001). Phylogenetic analysis based on a complete mitogenome yields more reliable phylogenies than that based on single or partial mtDNA data (Zardoya and Meyer, 1996). It delivers molecular resolution that is sufficient to distinguish patterns that have arisen over thousands an extended period.

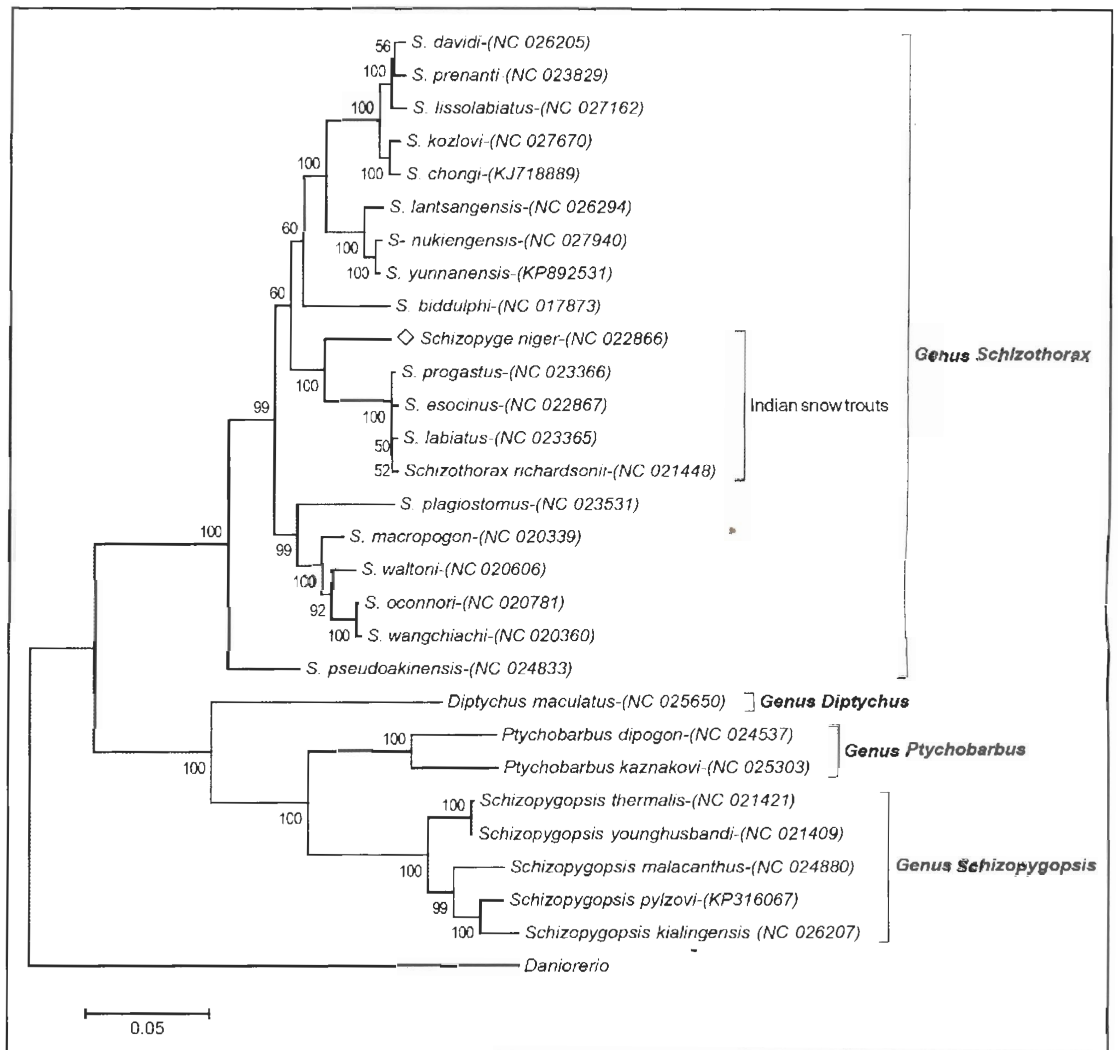


Fig 3. Maximum likelihood phylogenetic tree (UPGMA)

Fig 3. depicts the phylogenetic tree constructed by Maximum Likelihood method based on the General Reversible Mitochondrial model with the parameters provided in parenthesis (substitution model= General time reversible +G+T; Bootstrap= 1000; Uniform rate using all

the sites; log likelihood= -110310.8163). The same tree topology was obtained by the Neighbor joining method with strong posterior probability values for each nodal relationship. It shows the position of different genus *Schizothorax, Ptycobarbus, Schizopygopsis,* and *Diptychus.* The Maximum Likelihood tree (Fig. 3) indicates that S. *progastus, S. esocinus, S. labiatus, S. richardsonii* (generally known as Indian snow trout) have the same ancestral origin and closely related among themselves, while *S. plagiostomus forms* different clade suggesting that it may have diverged during evolution. The snow trout species majorly found in India have the common ancestral origin and further diverged into two Genus viz. *Schizothorax* and *Schizopyge.* So the present phylogeny supports the classification of *Schizopyge niger* in separate Genus from Schizothorax.

10. Contribution of ICAR-DCFR

- Apart from the six snow trout species, ICAR- DCFR has also developed the complete mitochondrial genome of threatened chocolate mahseer; *Neolissochilus hexagonolepis* (Sahoo et al, 2015), an endemic cold-water fish; Golden Mahseer, *Tor putitora* (Sati et al, 2014), Tor mahseer, Tor tor (Kumar et al, 2016) and revealed the organization of different genes in the genome.
- The genetic variations between seven geographically isolated populations of *T. putitora* have been elucidated using Cyt b (Cytochrome b) and ATPase6/8 gene sequences of mitochondrial DNA (Sati et al, 2015).
- Five different species of Mahseer (*Tor putitora, Tor tor, Tor khudree, Tor chelynoides* and *Neolissochilus hexagonolopis*) was identifiedand predicted the phylogenetic relationship among them using Cytochrome Oxidase I (COI) DNA barcodes (Sati et al, 2013).
- The high level of genetic differentiation within and between the populations of *Barilius bendelisis* was revealed by the cytochrome b gene sequences and the suitability of partial Cytochrome b gene (307 bp) sequence in determining the genetic diversity in *B. bendelisis* populations have been validated successfully (Seema et al, 2011).

11. Conclusion

Although the number of species with complete mitochondrial genomes is increasing rapidly, there are only few species such as rainbow trout (Brown et al, 2006b), Atlantic salmon (Hurst et al, 1999) and Medaka fish (Temminck & Schlegel 1846), that have multiple complete mitochondrial genomes sequences, to allow any intra specific mitochondrial genome comparisons. These Intra species comparisons will be vital for future research in determining, which mitochondrial genes and/or sequences result in beneficial phenotypes for important physiological traits, such as temperature tolerance, feed conversion, growth and development rate. There are few intra specific analyses have been performed using complete mitochondrial genomes. Using genetic manipulations to create families with identical nuclear backgrounds with differing mitochondrial backgrounds researchers have shown developmental rate and growth differences based solely on mitochondrial genomes. Developmental rate differences have been observed in common agricultural mushrooms and rainbow trout (De La Bastide et al, 1997). Additional animal studies have shown beneficial yield trait differences in poultry and cattle linked to mitochondrial genome variation indicating the potential for further enhancement of candidate species in agriculture and aquaculture (Brown et al, 2006).

Thus, the description of complete mitogenomes of six schizothoracid fish species will serve as a base source for such intra specific analyses for future research in determining which mitochondrial genes and/ or sequences result in beneficial phenotypes, inter specific studies in revealing interrelationships within families, orders. It will also be helpful in further taxon diagnosis to identify Schizothorax species.

12. References

1. Anderson S, Bankier AT, Barrell BG, de Bruijin MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH,. Smith AJH, Standen R, Young IG. (1981). Sequence and organization of the human mitochondrial genome. Nature. (290): 457-465.

2. Asakawa S, Kumazawa Y, Araki T, Himeno H, Miura K, Watanabe K. (1991). Strand-specific nucleotide composition biasin echinoderm and vertebrate mitochondrial genomes. J Mol Evol. 32:511–520.

3. Brown KH, Drew RE, Weber LA, Thorgaard GH. (2006). Intraspecific variation in the rainbow trout mitochondrial DNA genome. Comparative Biochemistry and Physiology D, Genomics and Proteomics. 1: 219-226.

4. Brown KH. (2008). Fish mitochondrial genomics: sequence, inheritance, and functional variation. J Fish Biol. 72:355-374.

5. Bucciarelli G, Golani D, Bernardi G. (2002). Genetic cryptic species as biological invaders: the case of a Lessepsian fish migrant, the hardy head silverside *Atherinomorus lacunosus.* J Exp Mar Biol Ecol. 273:143-149.

6. Cheng YZ, Xu TJ, Jin XX, Tang D, Wei T, Sun YY, Meng FQ.(2012). Universal primers for amplification of the completemitochondrial control region in marine fish species. Mol Biol. 46:27-30.

7. Chirag Goel, Jyoti Sati, Ashoktaru Barat, R. S. Patiyal, S. Ali, and Prabhati Kumari Sahoo. (2014). Complete mitochondrial genome organization of *Schizothorax richardsonii* (Gray, 1832). (2014). Mitochondrial DNA25(3): 171–172.

8. Chirag Goel, Prabhati Kumari Sahoo, and Ashoktaru Barat. Complete mitochondrial genome organization of *Schizothorax* progastus (McClelland, 1839). (2015). Mitochondrial DNA 26(6): 906–907.

9. Chirag Goel, Prabhati Kumari Sahoo, Ashoktaru Barat.(2016). Complete mitochondrial genome organization of *Schizothorax plagiostomus* (Heckel, 1838) Mitochondrial DNA Part 27(1).

10. Chirag Goel, Prabhati Kumari Sahoo, Farooz A. Bhat, M. H. Balkhi, Ashoktaru Barat. (2015). Complete mitochondrial genome organization of *Schizopyge niger* (Heckel, 1838). Mitochondrial DNA 26 (5).

11. Cox AJ & Hebert PDN. (2001). Colonization, extinction and phylogeographic pattern in gina freshwater crustacean. Mol. Ecol. 10: 371-386.

12. De La Bastide, P. Y., Sonnenberg, A. S. M., Van Griensven, L. J. L. D., Anderson JB, Horgen, PA. (1997).Mitochondrial haplotype influences mycelial growth of Agaricus bisporus heterokaryons. Appl and Envi Microbiol. 63, 3426-343.

13. Gissi C, Iannelli F, Pesole G. (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity. 101:301-320.

14. Hammond P. (1992). Species inventory. In: Global biodiversity: status of the earth's living resources (ed. B. Groombridge). Chapman and Hall, London. 17-39.

15. Hawksworth D. L and Kalin-Arroyo MT. (1995). Magnitude and distribution of biodiversity. In: Global biodiversity assessment (ed. V.H. Heywood). Cambridge University Press, London. 107 -19.

16. Iffat F. (2002). Mullets of Korangi Creek, Karachi. Rec Zool Surv Pak. 14:11-18

17. Jayaram, K.C. (1999). The fresh water fishes of the Indian region.New Delhi, India: Narendra publishing house.140-148.

18. Jyoti Sati , Seema Sah , Himani Pandey , Shahnawaz Ali , Prabhati Kumari Sahoo , Veena Pande and Ashoktaru Barat. (2013). Phylogenetic relationship and molecular identification of five Indian Mahseer species using COI sequence. J Env Biol. 34 (5), 933-939.

19. Jyoti Sati, Rohit Kumar, Prabhati Kumari Sahoo, Rabindar S. Patiyal, Shahnawaz Ali, and Ashoktaru Barat. (2015). Genetic characterization of Golden mahseer (*Tor putitora*) populations using mitochondrial DNA markers. Mitochondrial DNA. 26(1): 68–74

20. Kartavtsev Y. P & Lee JS. (2006). Analysis of nucleotide diversity at genes Cytb andCo-1 on population, species, and genera levels. Applicability of DNA and allozyme data in genetics of speciation. Russ. J. Genet. 42: 341-362.

21. Kurtzman CP. (1994). Molecular taxonomy of the yeasts. Yeast. 10:1727-1740.

22. Mabuchi K, Miya M, Senou H, Suzuki T, Nishida M. (2006). Complete mitochondrial DNA Sequence of the Lake Biwa wild strain of common carp *(Cyprinus carpio L)*: Further evidence for an ancient origin. Aquaculture. 257:68–77.

23. Meyer A. (1993). Evolution of mitochondrial DNA in fishes. In: Hochachka, P.W., Mommsen, T.P. (Eds.), Biochemistry and Molecular Biology of Fishes. Elsevier Science Publishers, New York:1-38.

24. Miya M, Kawaguchi A, Nishida M. (2001). Mitogenomic exploration of higher teleostean phylogenies: a case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. Mol Biol Evol. 18:1993-2009.

25. Murta AG. (2000). Morphological variation of horse mackerel *(Trachurus trachurus)* in the Iberian and North Africa Atlantic: implications for stock identification. ICES J Mar Sci. 57: 1240- 1248.

26. Prabhati Kumari Sahoo, Ashoktaru Barat,Farooz A. Bhat, M. H. Balkhi, Chirag Goel. (2015). Complete mitochondrial genome organization of Schizothorax esocinus (Heckel, 1838). (2015). Mitochondrial DNA 26(5): 706–707.

27. Prabhati Kumari Sahoo, Chirag Goel, Farooz A. Bhat, M. H. Balkhi, Ashoktaru Barat. (2015). Complete mitochondrial genome organization of *Schizothorax labiatus* (McClelland, 1842). Mitochondrial DNA, Early Online: 1–2.

28. Raina HS and Petr T. (1999). Coldwater Fish and Fisheries in the Indian Himalayas: Lakes and Reservoirs. FAO fish. Tech. papers, version T-385, FAO.
29. Rohit Kumar, Chirag Goel, Prabhati Kumari Sahoo, Atul K Singh, Ashoktaru Barat.(2016). Complete mitochondrial genome organization of Tor tor (Hamilton, 1822) Mitochondrial DNA. 27(4).
30. Sahoo PK, Goel C, Kumar R, Dhama N, Ali S, Sarma D, Nanda P, Barat A.(2015). The complete mitochondrial genome of threatened chocolate mahseer (*Neolissochilus hexagonolepis*) and its phylogeny.Gene. 570(2):299-303.
31. Saitoh K, Sado T, Mayden RL, Hanzawa N, Nakamura K, Nishida M, Miya M. (2006). Mitogenomic evolution and interrelationships of the Cypriniformes (Actinopterygii: Ostariophysi): the first evidence toward resolution of higher-level relationships of the world's largest freshwater fish clade based on 59 whole mitogenome sequences. J. Mol. Evol. 63:826–841.
32. Sambrook J & Russell D. (2001). Molecular Cloning: a Laboratory Manual, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
33. Sati J, Goel C, Kumar R, Ali, S, Patiyal RS, Singh VK, Sahoo PK, Barat A. (2014). Completemitochondrial genome organization of *Tor putitora*. Mitochondrial DNA 25 (4): 278–279.
34. Schander C & Willassen E. (2005). What can biological barcoding do for marine biology?. Mar Biol Res. 1:79-83.
35. Schattner P, Brooks AN, Lowe TM. (2005). The tRNAscan-SE, snoscan and snoGPS webservers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. 33 (2):686-689.
36. Seema Sah, Ashoktaru Barat, Veena Pande, Jyoti Sati and Chirag Goel. (2011). Population Structure of Indian Hill Trout (*Barilius bendelisis*) Inferred from Variation in Mitochondrial Dna Sequences.Advan. Biol. Res. 5 (2): 93-98.
37. Silas EG. (1960). Fishes from Kashmir Valley. J. Bombay Natural History Soc. 57: 66-67.
38. Simmons RB &Weller SJ. (2001). Utility and evolution of cytochrome b in insects. Mol. Phylogenet. Evol. 20:196– 210.
39. Smith DR & Keeling PJ. (2015). Mitochondrial and plastid genome architecture: reoccurring themes but significant differences at the extremes. Proc. Natl. Acad. Sci. U.S.A. 112:10177-10184.
40. Smith DR. (2016). The past, present, and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief. Funct. Genomics. 15: 47-54
41. Strauss RE, Bond CE. (1990). Taxonomic methods: morphology. In: Schreck CB, Moyle PB (eds) Methods for fish biology. American Fisheries Society, Maryland.109–140.
42. Talwar PK, Jhingran GA. (1991). Inland fisheries of India and adjacentcountries. New Delhi, India: Oxford and IBH Publishing Company Pvt Limited. (1): 410–12.

43. Tinti F, Ungaro N, Pasolini P. (2003). Development of molecular and morphological markers to improve species-specific monitoring and systematics of Northeast Atlantic and Mediterranean skates (Rajıformes). J Exp Mar Biol Ecol. 288:149–165.

44. Wang HY, Tsai MP, Tu MC, Lee SC. (2000). Universal primers for amplification of the complete mitochondrial 12S rRNA gene invertebrate. Zool Stud. 39(6):1–6.

45. Wares JP & Cunningham CW. (2001). Phylogeography and historical ecology of the North Atlantic intertidal. Evolution. 12:2455–2469.

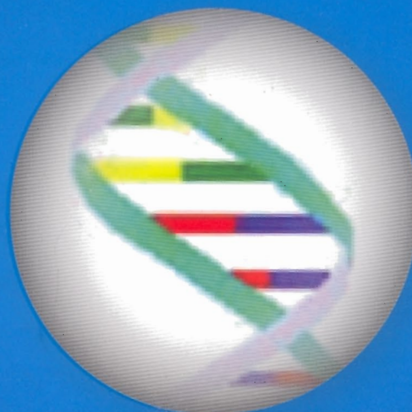
46. Wilson KH. (1995).Molecular biology as a tool for taxonomy. Clin Infect. Dis.20 (Suppl.):192 –208.

47. Zardoya R, Meyer A. (1996). Phylogenetic performance of mitochondrial protein coding genes in resolving relationships among vertebrates. Mol. Biol. Evol. 13: 933–942.

48. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31(13): 3406–3415.

हर कदम, हर डगर
किसानों का हमसफर
भारतीय कृषि अनुसंधान परिषद

Agrisearch with a human touch



ICAR-Directorate of Coldwater Fisheries Research
(Indian Council of Agricultural Research)
Bhimtal - 263136, Distt. Nainital, Uttarakhand
Website: www.dcfr.res.in
E-mail: director@dcfr.res.in; dcfrin@rediffmail.com
Phone: 05942-247280, 247279; Fax: 05942-247693