Bangladesh J. Zool. **48**(1): 1-19, 2020 ISSN: 0304-9027 (print)
: 2408-8455 (online)

MOLECULAR DIVERSITY OF FRESHWATER FISHES OF BANGLADESH ASSESSED BY DNA BARCODING

Md. Sagir Ahmed[1]*, Sujan Kumar Datta[2] and Ayesha Akhter Zhilik[1]

Department of Zoology, University of Dhaka, Dhaka1000, Bangladesh

ABSTRACT: This study represents the comprehensive molecular identification of freshwater fishes of Bangladesh based on a fragment of the cytochrome c oxidase subunit I (COI) gene in the mitochondrial genome. A total of 315 mitochondrial COI barcode sequences were obtained from 153 species,114 genera, 49 families and 16 orders of fishes. The mean length of the sequences was 652 base pairs. For all the samples, %G was significantly lower compared to the other nucleotides and %GC was lower compared to %AT (p-value < 0.05). Also, a significantly lower %GC content was observed in second and third codon position compared to the first one in all the samples ($1^{st}>2^{nd}>3^{rd}$, p-value< 0.05). The average K2P distances within species, genera, families and orders were 0.38%, 7.02%, 12.75% and 18.68%, respectively. The mean interspecific distance was 18-fold higher than the mean intraspecific distance. The K2P neighbor-joining (NJ) trees based on the sequences generally clustered species according to their taxonomic position. A total of 12 species were newly recorded in Bangladesh. High efficiency in species identification were demonstrated in the present study by DNA barcoding, and concluded that COI sequencing can be used as an authentic identification marker for freshwater fish species.

Key words: Freshwater fishes, COI, DNA Barcoding, Genetic Diversity, Phylogeny

INTRODUCTION

Bangladesh is a deltaic country that emerged on the confluence of the three mighty river systems, Padma-Ganges, Jamuna-Brahmaputra and the Surma-Meghna. This unique geophysical condition constitutes about 46,99,387 ha of diverse inland water areas comprises more than 250 native freshwater fish species (IUCN 2015). Inland fisheries play an important role in the economy of the country in terms of nutrition, income, employment and foreign exchange earnings. More than 11 percent of the total population is engaged in this sector in full time and part time basis for their livelihoods. And fish itself contributes about 60% of animal protein in the daily dietary requirement of 160 million population of the country (DoF 2018). Inland capture fishery production remarkably declined in the past few decades. Currently, inland capture fishery contributes only 29% of the country's total fish production which was 63% in

*Author for correspondence: <sagir@du.ac.bd>, [2]Department of Zoology, Jagannath University, Dhaka1100, Bangladesh

 DOI: https://doi.org/10.3329/bjz.v48i1.47872

1983-84 (DoF 2018). This dramatic change in production was due to the continuous habitat destruction, unregulated fishing and aquatic pollution from industrial, domestic and agricultural sources. That is also leading to at least 64 species of freshwater fishes at different categories of threatened with extinction (IUCN 2015).

Nonetheless, description and information of freshwater fishes of Bangladesh are scattered throughout a wide range of publications. The first complete description of freshwater fishes of Bangladesh was compiled by Rahman (1989) which corroborated 260 species including migratory and estuarine species. Siddiqui *et al.* (2007) described 251 species which is in fact reproduce the previous works.

Published books on diversity of Bangladeshi fishes (Bhuyian 1964, Rahman 2005, Shafi and Quddus 1982, IUCN 2000, IUCN 2015) evident that the information is inconsistent and taxonomy used is not updated.

The accurate identification of fish species is a pivotal component to protect the extant ichthyofaunal biodiversity and to perform regular assessments of local fish faunas for conservation planning. As an alternative to the traditional species identification based on morphological characters, Hebert *et al.* (2003) has been suggested partial cytochrome c oxidase subunit I (COI) sequences (DNA barcodes) for standardized and routine species identification. The DNA-based barcoding method has been proven to be a valuable molecular tool for species identification and it is accessible to non-specialists (Hebert *et al.* 2003, Fre´zal and Leblois 2008, Leray *et al.* 2015). This barcoding technique has been successfully identified ichthyofauna in many geographic regions, such as Australia (Ward *et al.* 2005), Canada (Hubert *et al.* 2008), India (Lakra *et al.* 2015), China (Zhang 2011, Wang *et al.* 2018), Japan (Zhang and Hanner 2011), Portugal (Costa *et al.* 2012), Germany (Knebelsberger *et al.* 2015), Taiwan (Bingpeng *et al.* 2018, Chang *et al.* 2017) and Vietnam (Thu *et al.* 2019), thus enriched the barcode reference library. New specimens and products can confidently be identified by comparing their DNA barcode sequences against this barcode reference library. Although barcodes for almost two-third (19,000) of all described fish species are already available (BOLD 2020), the permanent addition of new barcode data is essential to increase the taxonomic resolution.

Considering the economic importance of inland fishery and the expected richness of the fish fauna and in the absence of an expert-based taxonomy, the first attempt for barcoding of freshwater fishes of Bangladesh was initiated in 2014 and consequently partially published barcode data of 81 small indigenous fish species (SIS) by Ahmed *et al.* (2019). This paper deals with the molecular characterization of morphologically identified freshwater species of Bangladesh using partial COI gene sequence.

MATERIAL AND METHODS

Study area and specimen collection: Fish samples were collected from rivers, haor, baor, beels, floodplain, fish landing centers, fish markets or from the local fishermen during July 2014 to June 2018. Personal fishing was also conducted to collect some rare and non-commercial fish species whenever necessary. Photographs of all the fishes were taken immediately and taxonomic identification of specimens were done following previous reports (Talwar and Jhingran 1991, Rahman 2005, Siddiqui *et al.* 2007). Immediately after collecting the specimens, tissue samples were excised and stored in 90% ethanol. Voucher specimens were fixed with 10% formalin and then transferred to 70% ethanol solution for preservation. Voucher specimens were transported to Dhaka and deposited in the Professor Kazi Zaker Hussain Museum at the Department of Zoology, University of Dhaka.

DNA barcoding: Genomic DNA was extracted from the muscle tissue samples by the standard Proteinase-K/Phenol-Chloroform-isoamyl alcohol method (Green and Sambrook 2012, Ahmed *et al.* 2019). The quality and quantity of the extracted DNA was measured using Nanodrop™ spectrophotometer. Approximately 658bp was amplified from the 5′ region of the MT-COI gene using the following primers: FishF1 5′TCAACCAACCACAAAGACATTGGCAC3′ and FishR1 5′ TAGACTTCTGGGTGGCCAAAGAATCA3′ only when failed to amplified the FishF2 5′TCGACTAATCATAAAGATATCGGCAC3′ and FishR2 5′ACTTCAGGGTGACCGAAGAATCAGAA3′ were used (Ward *et al.* 2005). For this, 25 µl PCR reaction mixtures were prepared which included 17.25–18.75 µl of ultrapure water, 2.5 µl of 10× PCR buffer, 1.25 µl of $MgCl_2$ (50 mM), 0.25 µl of each primer (0.01 mM), 0.125 µl of each dNTP (0.05 mM), 1 µl (0.625 U) of Taq polymerase, and 0.5–2.0 µl of DNA template. Amplifications were performed using ABI thermal cycler (Thermo Fisher Scientific). The thermal regime consists of an initial step of 2 min at 95°C followed by 35 cycles of 0.5 min at 94°C, 0.5 min at 54°C, and 1 min at 72°C, followed in turn by 10 min at 72°C and then held at 4°C. PCR products were visualized on 1% agarose gel. The PCR products were purified using PureLink™ PCR purification kit and sequenced from First BASE Laboratories, Sdn Bhd, Malaysia. All sequences were translated into amino acids to confirm the effectiveness of the sequences and to detect the presence of nuclear DNA pseudo genes, insertions, deletions, or stop codons. Sequences were checked and aligned using Sequencer v5.4.6 and were submitted to GenBank with referred accession numbers. All the data including taxonomic characteristics and GenBank accession numbers were tagged with the voucher specimens preserved at the Professor Kazi Zaker Husain Museum of Department of Zoology, University of Dhaka.

***Bioinformatic and statistical analyses*:** Bioinformatic analyses of the sequences were performed using CLC Workbench v7.7.1, Mega X, Clustal Omega, and T-Coffee. Base compositions were analyzed using CLC Workbench v7.7.1 and Mega X. Genetic distance and sequence divergences were calculated using the Kimura two parameter (K2P) distance model (Kimura 1980). Neighbor-joining (NJ) trees of K2P distances were created to provide a graphic representation of divergence pattern between species (Saitou and Nei 1987). Bootstrapping was performed in MEGA X (Tamura and Nei 1993) with 1000 replications (Felsenstein 1985). Necessary statistical analyses were performed in Excel 2013 and RStudio.

RESULTS

A total of 550 tissue samples were collected, during the study period among which 315 COI sequences were obtained (Table 1). Based on morphological and molecular identifications, these samples represented 153 species of 114 genera, 49 families and 16 orders (Table 1). Among the collected species, Order Cypriniformes was recorded as the most diversified fish group in terms of both number of species and individuals observed followed by Perciformes and Siluriformes (Table1). Among the 153 species, 12 fish species were newly recorded in Bangladesh. The average length of all barcode sequences was 652 bp ranging from 352 to 696 bp where 95% were longer than 600bp. No stop codon, insertion, or deletion was observed in any of the obtained sequences.

Table 1. List of freshwater fish species barcoded along with their GenBank (GB) accession numbers

SL No.	Order	Family	Species	No. of individual	GB Accession Number
1	Osteoglossiformes	Notopteridae	*Chitala chitala*	1	MF140393
2			*Notopterus notopterus*	2	KT346361 KT364757
3	Elopiformes	Megalopidae	*Megalops cyprinoides*	2	MN171367 MN171368
4	Anguilliformes	Ophichthidae	*Pisodonophis boro*	1	MG969529
5	Clupeiformes	Clupeidae	*Corica soborna*	2	KX455892 KY124368
6			*Gonialosa manmina*	1	MH087054
7			*Tenualosa ilisha*	2	KX657721 MH230965
8			*Anodontostoma chacunda*	2	MK878431 MH429338
9		Engraulidae	*Setipinna phasa*	2	MN083101 MH429325
10			*Coilia ramcarati*	3	MK926759 MN083109 MH311288
11			*Coilia dussumieri*	5	MN083117 MK988524 MN171355 MN200458 MH230984
12		Pristigasteridae	*Ilisha melastoma*	2	MN200469 MN200470
13			*Pellona ditchela*	1	MN083106

SL No.	Order	Family	Species	No. of individual	GB Accession Number
14	Gonorynchiformes	Chanidae	*Chanos chanos*	1	MN083123
15	Cypriniformes	Nemacheilidae	*Acanthocobitis botia*	2	KT762380 MN013423
16			*Acanthocobitis zonalternans*	1	KT762362
17			*Schistura fasciolata**	1	KY124367
18			*Paracanthocobitis zonalternans*	2	MN200466 MN200467
19		Cyprinidae	*Amblypharyngodon mola*	2	KT364774 MH087039
20			*Aspidoparia jaya*	2	MG969527 MG969532
21			*Barbonymus gonionotus*	1	KX657718
22			*Barilius barna*	1	KY124376
23			*Cirrhinus cirrhosus*	1	KT353104
24			*Cirrhinus reba*	3	KX455893 MG969514 MN083095
25			*Crossocheilus latius*	2	MG969525 MG969531
26			*Ctenopharyngodon idella*	1	KX657712
27			*Cyprinus carpio*	3	KX657710 KX657711 MN234111
28			*Danio rerio*	1	MF170952
29			*Devario aequipinnatus*	3	KT364769 KY124372 KY124375
30			*Esomus danricus*	1	KT364776
31			*Garra nasuta**	1	KY124363
32			*Garra sp.*	1	MF190550
33			*Gibelion catla*	1	MG969520
34			*Hypophthalmichthys molitrix*	3	KX657713 MH087046 MF140395
35			*Labeo bata*	5	KT353105 MG969515 MN083093 MN083094 MH087029
36			*Labeo calbasu*	1	KT364767
37			*Labeo gonius*	2	KX455894 MN200474
38			*Labeo rohita*	4	MG969513 MG969519 MH087049 MF170947
39			*Chela cachius*	1	KT353102
40			*Laubuca laubuca*	1	KT353103
41			*Neolissochilus hexastichus*	1	KT364770
42			*Oreichthys cosuatis*	2	KX455909 MN013419
43			*Osteobrama cotio*	4	KT762359 MN200463 MN200464 MN200465
44			*Pethia conchonius*	4	MK988520 MK988542 KY124379 KY124380
45			*Pethia gelius*	2	MN200473 KT364772
46			*Pethia guganio*	1	KT762360

SL No.	Order	Family	Species	No. of individual	GB Accession Number
47		Cyprinidae	*Pethia phutunio*	1	KT353106
48			*Pethia ticto*	1	MN083131
49			*Puntius chola*	3	KT364771 MN171353 MN171354
50			*Systomus sarana*	3	KT364773 MH087036
51					MN171373
52			*Puntius sophore*	1	KX455895
53			*Puntius terio*	2	KX455896 MN200455
54			*Raiamas bola*	1	KY124369
			Rasbora daniconius	4	KT364777 MG280610 MN013420 MN200472
55			*Rasbora rasbora*	1	MK995091
56			*Salmophasia bacaila*	6	MN171372 KT364775 MG550117 MH087030 MG550117 MN234103
57			*Salmostoma phulo*	1	KT364758
58			*Securicula gora*	2	MG969526 MG969533
59			*Tor putitora*	2	KT762361 KT762379
60		Botiidae	*Botia Dario*	5	MH087038 MH087045 MN171346 MN171347 KY124374
61			*Botia lohachata*	3	KX455912 MH087044 MN083135
62			*Botia rostrata*	1	KY124362
63		Cobitidae	*Canthophrys gongota*	2	KX455897 MH087035
64			*Lepidocephalichthys annandalei*	2	KY124364 MF140396
65			*Lepidocephalichthys guntea*	7	KT364759 KT364778 MN013421 MN171348 MN171349 MN171350 MN171351 KY124365
66			*Pangio pangia*	1	MF170949
67		Psilorhynchidae	*Psilorhynchus balitora*	1	KY124373
68			*Psilorhynchus sucatio*	1	MF170951
69	Siluriformes	Ailiidae	*Ailia coila*	3	KT364761 KT364782 MN083152
70		Ariidae	*Osteogeneiosus militaris*	3	MH429317 MH429348 MH230983
71		Amblycipitidae	*Amblyceps mangois*	1	KT762370
72		Bagridae	*Batasio convexirostrum**	1	KY124366
73			*Mystus bleekeri*	2	KT364779 MN083144
74			*Mystus cavasius*	3	KT762365 KX657719 MN083157
75			*Mystus tengara*	3	KT762366 MK988521 MN083145
76		Bagridae	*Mystus vittatus*	1	KT364780
77			*Mystus gulio*	4	KX455898 KX455905 MN083111MK995086

SL No.	Order	Family	Species	No. of individual	GB Accession Number
78			*Hemibagrus menoda*	2	KT762363 MG969522
79			*Rama chandramara*	1	KT762367
80			*Rita rita*	2	KT364781 KT762374
81			*Sperata aor*	1	KT762381
82			*Sperata seenghala*	3	KT364786 KT762382 MN171374
83		Plotosidae	*Plotosus canius*	3	KX657716 MK995093 MN171370
84		Chacidae	*Chaca chaca*	2	KX455900 MN083143
85		Clariidae	*Clarias batrachus*	1	KT762385
86			*Clarias gariepinus*	1	KX657715
87		Schilbeidae	*Clupisoma prateri**	5	KT364783 KT762369 KX455899 MG969517 MN200476
88			*Clupisoma garua*	1	KX455904
89			*Eutropiichthys vacha*	1	KT364762
90			*Neotropius atherinoides*	2	KT364763.1 KT364784
91		Sisoridae	*Gagata cenia*	2	KT762384 MG969536
92			*Gagata gagata*	2	KT364785 MG969523
93			*Bagarius bagarius*	4	KT762371 KX455910 MG969530 MN200478
94			*Glyptothorax indicus*	1	MH087037
95		Erethistidae	*Hara jerdoni*	1	KT762372
96			*Erethistes pusillus*	1	MG969534
97		Heteropneustidae	*Heteropneustes fossilis*	5	KT364787 MG969521 MN083153 MN083154 MN083155
98		Olyridae	*Olyra longicaudata*	2	MF176156 KT762373
99		Siluridae	*Ompok bimaculatus*	3	KT762368 MH087040 MN083156
100			*Ompok pabda*	3	KT364760 KT762383 MN200457
101			*Ompok pabo*	1	KX455911
102			*Wallago attu*	2	KX657717 MH087042
103		Pangasiidae	*Pangasianodon hypophthalmus*	1	MF373123
104	Batrachoidiformes	Batrachoididae	*Batrachomoeus trispinosus**	3	MN234104 MN234105 MN234107
105	Cyprinodontiformes	Poeciliidae	*Poecilia reticulata*	1	MN083133
106	Syngnathiformes	Syngnathidae	*Microphis deocata*	1	KT762375
107	Synbranchiformes	Mastacembelidae	*Macrognathus aral*	6	MK995083 MN083138 MN083148 MN083149 KT762377 MF170946
108		Mastacembelidae	*Macrognathus pancalus*	4	KT762378 MH087034 MN200459 MN200460

SL No.	Order	Family	Species	No. of individual	GB Accession Number
109			*Mastacembelus armatus*	1	KT762364
110		Synbranchidae	*Monopterus cuchia*	1	MG969535
111	Scorpaeniformes	Platycephalidae	*Platycephalus indicus*	1	MH429330
112	Perciformes	Gobiidae	*Awaous grammepomus*	1	MK988544
113			*Awaous sp.*	1	MN083092
114			*Favonigobius gymnauchen**	2	MN083121 MK995095
115			*Pseudapocryptes elongatus*	3	MK926762 MK988530 MN013430
116			*Glossogobius giuris*	4	KT364791 MH087041 MK926756 MH429327
117			*Tridentiger barbatus**	1	MN083132
118			*Trypauchen vagina*	1	MK926755
119			*Scartelaos histophorus*	4	MH087031 MK926760 MK988529 MN234102
120			*Stigmatogobius sadanundio*	1	MK995090
121			*Odontamblyopus rubicundus*	2	MH882462 MH882463
122			*Oligolepis acutipennis**	1	MK9885324
123			*Boleophthalmus boddarti*	2	MH429333 MN083126
124			*Acentrogobius nebulosus**	1	MN083110
125			*Parapocryptes serperaster**	1	MN083127
126		Anabantidae	*Anabas cobojius*	1	KY124377
127			*Anabas testudineus*	3	KX455903 MN083163 MN083164
128		Badidae	*Badis badis*	1	KT364764
129			*Badis chattagongis*	4	KX455902 KX455906 KY124378 KY124371
130			*Badis tuivaiei**	1	KY124370
131		Ambassidae	*Chanda nama*	4	KT364788 MH087050 MN083146 MN083147
132			*Parambassis lala*	*2*	KT364789
133		Eleotridae	*Eleotris fusca*	5	MK926753 MN083150 MN083151 MN013422 MF170948
134			*Butis butis*	1	MH827972
135			*Butis koilomatodon**	1	MN171371
136		Polynemidae	*Polynemus paradiseus*	5	MH087032 MH311275 MH311276 MH311282 MH230971
137		Latidae	*Lates calcarifer*	3	MN171369 MG969518 MH087052
138		Gerreidae	*Gerres filamentosus*	1	MK988532
139		Nandidae	*Nandus nandus*	3	KT762376 MN083160 MN083161
140		Osphronemidae	*Trichogaster chuna*	1	MH087047

SL No.	Order	Family	Species	No. of individual	GB Accession Number
141			*Trichogaster fasciata*	1	MH087051
142			*Trichopsis vittata*	3	KT364765 KT364768 KT364792
143		Mugilidae	*Rhinomugil corsula*	5	KT364790 MG969528 MN083165 MN171356 MN171357
144			*Mugil cephalus*	1	MK988536
145		Cichlidae	*Oreochromis niloticus*	1	KX657714
146		Channidae	*Channa marulius*	2	KX808573 MG969516
147			*Channa gachua*	1	KT364793
148			*Channa punctata*	1	KT762386
149			*Channa striata*	1	KT762387
150	Beloniformes	Adrianichthyidae	*Oryzias javanicus**	1	MF170950
151		Belonidae	*Xenentodon cancila*	1	MH087053
152		Zenarchopteridae	*Zenarchopterus ectuntio*	1	MK988518
153	Tetraodontiformes	Tetraodontidae	*Leiodon cutcutia*	4	MK926757 MN200461 MN200462 MF140394

**Species of new records*

Table 2. Genetic divergence (%K2P distance) of freshwater fishes within various taxonomic levels

Level	Sample size	Mean	Minimum	Maximum	SE
Species	145	0.38	0.00	2.93	0.01
Genus	105	7.02	0.00*	28.67	0.02
Family	44	12.75	0.57	22.39	0.03
Order	13	18.68	5.95	26.18	0.04

**Single sequence*

The lack of stop codons in these sequences indicates that they are functional mitochondrial COI sequences, together with the fact that each of the amplified sequence was about 658 bp in length. Hence, it suggests that Nuclear DNA Sequences Originating from Mitochondrial DNA Sequences (NUMTs) were not sequenced, as vertebrate NUMTS are typically less than 600 bp (Zhang and Hewitt 1996).

The sequence analysis indicated that the average nucleotide frequencies to be A: 25.20%, T:29.80%, G:17.70% and C: 27.30%(Fig. 1).The base composition analysis for the COI sequence showed that the average T content was the highest and the average G content was the lowest; the AT content (55.0%) was higher than

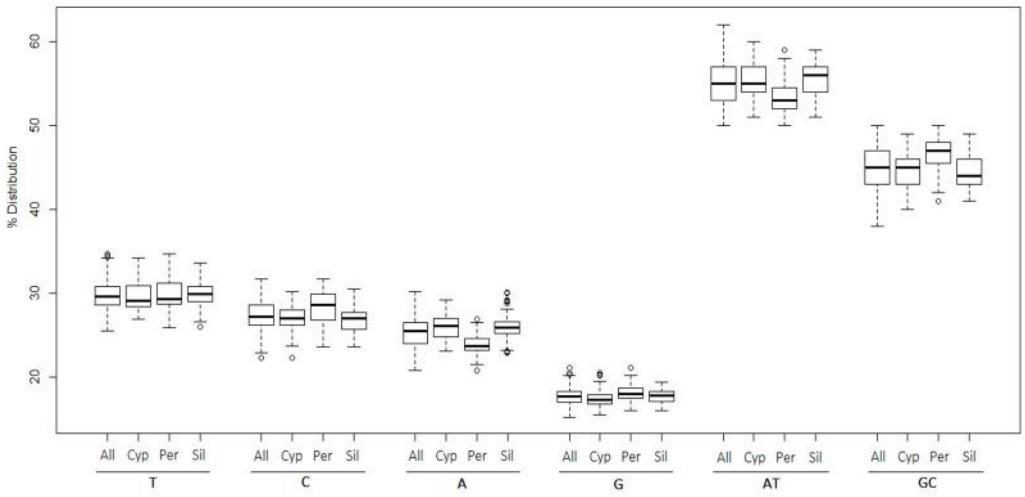

Fig. 1. Composition of the nucleotides in the sequenced COI region of freshwater fishes of Bangladesh.

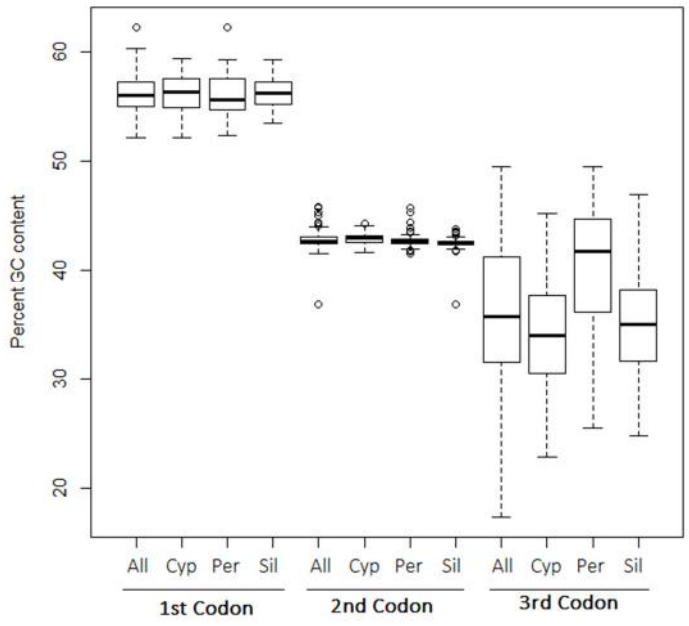

Fig. 2. Percent (%) GC content at different codon positions in the sequenced COI region of freshwater fishes of Bangladesh.

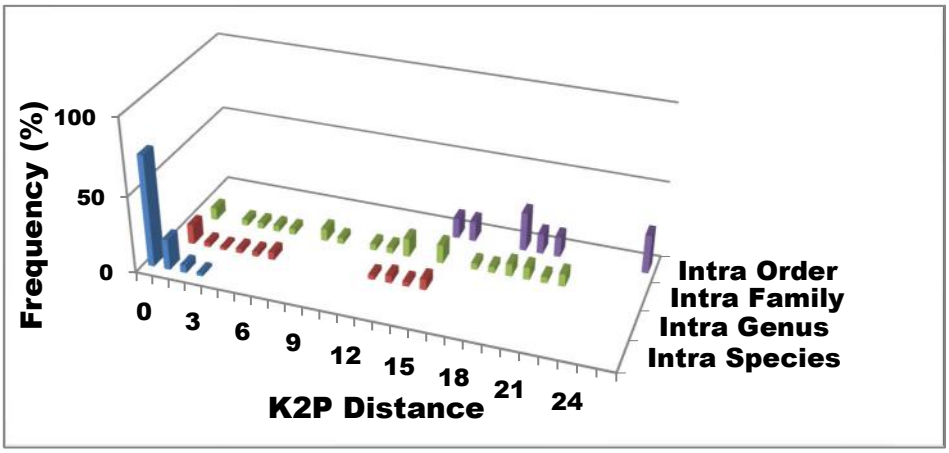

Fig. 3. Distribution of K2P distances (percentage) within different taxonomic categories.

the GC content (45.0%). The GC contents at the first, second and third codon positions for all fish were 56.09%, 42.78% and 36.02% respectively (Fig. 2). At the first codon position, the usage of T (19.10%) was the lowest, and the usages of the other bases were C: 25.40%, A: 24.80% and G: 30.60%. At the second codon position, the content of T (41.80%) was highest, and the contents of the other bases were C: 28.10%, A: 15.50% and G: 14.70%. At the third codon position, the base usage was T (28.40%), C (28.20%), A (35.60%) and G (7.80%); the G content was the lowest, exhibiting a clear pattern of anti-G bias. The average genetic distance within species, genus, family and order were 0.38±0.01%, 7.02±0.02%, 12.75±0.03% and 18.68±0.04%, respectively (Table 2). The NJ tree of 315 generated sequences including 153 species is presented in Figure 4.

Order Cypriniformes: This order includes many of the most important forage and food fish. A total of 112 samples were sequenced belonging to five families, 37 genera and 54 species. The overall mean nucleotide base frequencies observed for these sequences were- T: 29.70%, C: 27.0%, A: 25.80% and G: 17.50%. The AT content (55.50%) was higher than the GC content (44.50%). The GC contents at the first, second and third codon positions were 49.10%, 48.20% and 44.40% respectively. The NJ tree clearly distinguished all the species (Fig. 4).

Order Perciformes: A total of 72 samples were sequenced belonging to 12 families, 27 genera and 33 species. The overall mean nucleotide base frequencies observed for these sequences were T: 29.90%, C: 28.20%, A: 23.80% and G: 18.10%. The AT content (53.70%) was higher than the GC content (46.30%). The GC contents at the first, second and third codon positions were 56.70%, 42.70%

and 39.50%, respectively. In the NJ tree, most of the specimens belonging to the same species were clustered together bolstering the prior taxonomic assignment based on morphology (Fig. 4).

Order Siluriformes: A total of 76 samples were sequenced belonging to 14 families, 25 genera and 36 species. The overall mean nucleotide base frequencies observed for these sequences were- T: 29.70%, C: 26.90%, A:25.60% and G:17.70%. The AT content (55.30%) was higher than the GC content (44.70%). The GC contents at the first, second and third codon positions were 56.70%, 42.70% and 38.10% respectively.

DISCUSSION

DNA barcoding could be considered as a global bio-scanner for rapid and authentic identification of organisms using the partial sequence of mitochondrial COI gene. Barcoding has clearly discriminated freshwater fish species from around the globe including Australia, Canada, India, Thailand, Germany (Ward *et al.* 2005, Hubert *et al.* 2008, Lakra *et al.* 2015, Panprommin *et al.* 2019, Knebelsberger *et al.* 2014). Here, we have complied the COI sequence profile of freshwater fishes collected from the different inland waters of Bangladesh and confirmed the efficacy of barcoding to identify these species. Barcodes were generated for153species of belonging to 114 genera and 49 families and 16 orders (Table 1). We observed no insertions/ deletions or codon stops after translating the nucleotide sequences, supporting the view that all of the amplified sequences denote functional mitochondrial COI sequences. Moreover, average length of the amplified sequences was larger than 650bp, the limit typically observed for nuclear DNA sequences originating from mtDNA (NUMTs) (Gunbin *et al.* 2017). All of these species were differentiable based on the individual COI barcodes. Hence, this study has strongly validated the efficiency of COI barcodes for identifying fish species.

The base composition analysis of the COI sequences revealed that AT content (55.0%) to be higher than GC content (45.0%), similar to the patterns were observed in Australian (Ward *et al.* 2005), Canadian (Steinke *et al.* 2009) and Cuban fish species (Lara *et al.* 2010). The GC contents in the first, second and third codon positions were 56.09%, 42.78% and 36.02%, respectively (Fig. 2). At the first codon position, the usage of G (19.00%) was the lowest, and the usages of the other bases were 23.9%, 32.6% and 24.00% for C, A and T, respectively. At the second codon position, the content of T (31.00%) was highest, and the contents of the other bases were 25.7%, 23.1% and 20.0 for C, A and G, respectively. At the third codon position, the base usage was- T: 33.00%, C:32.0%,

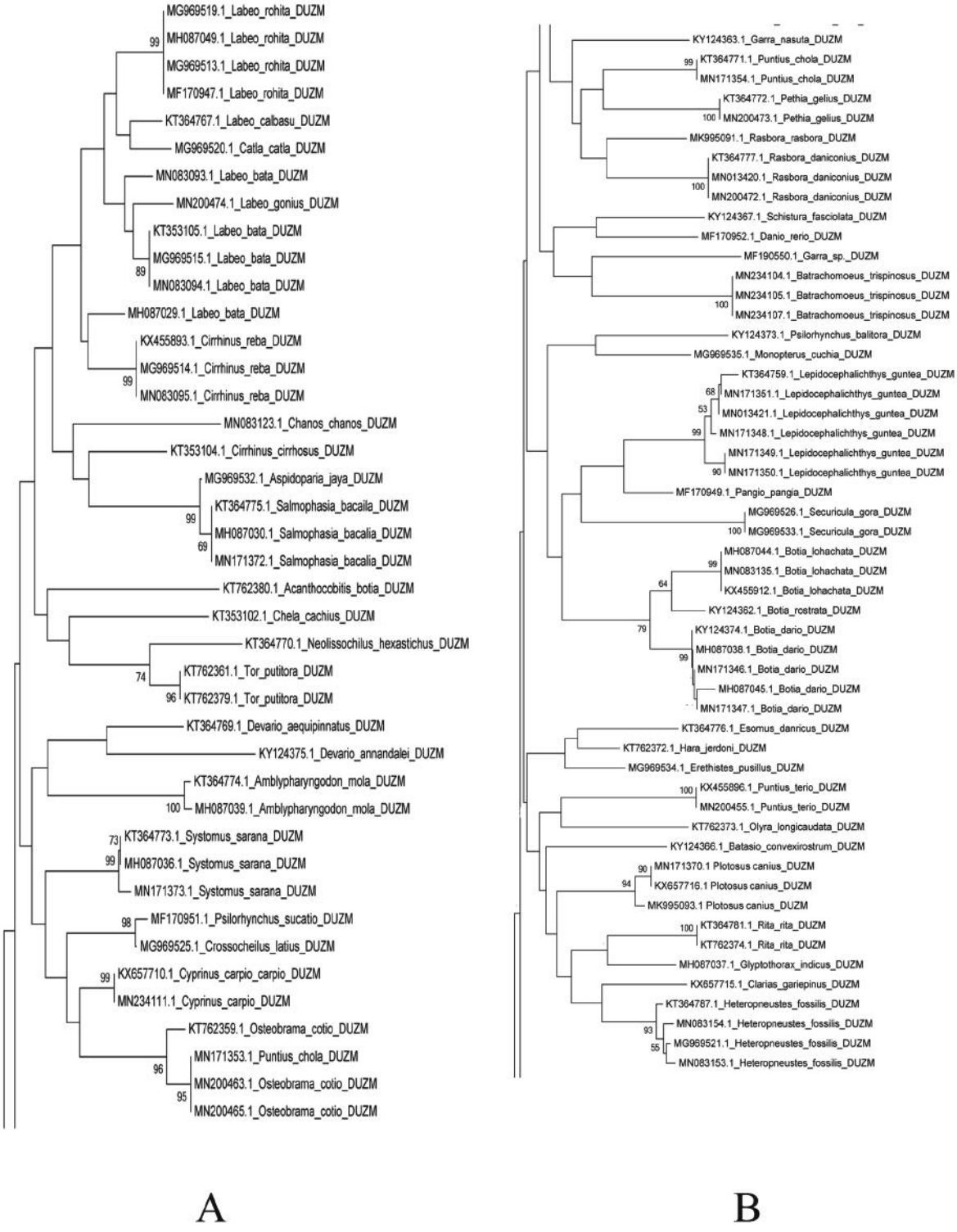

Fig. 4. Neighbor-joining (NJ) tree of freshwater fishes using K2P distances.

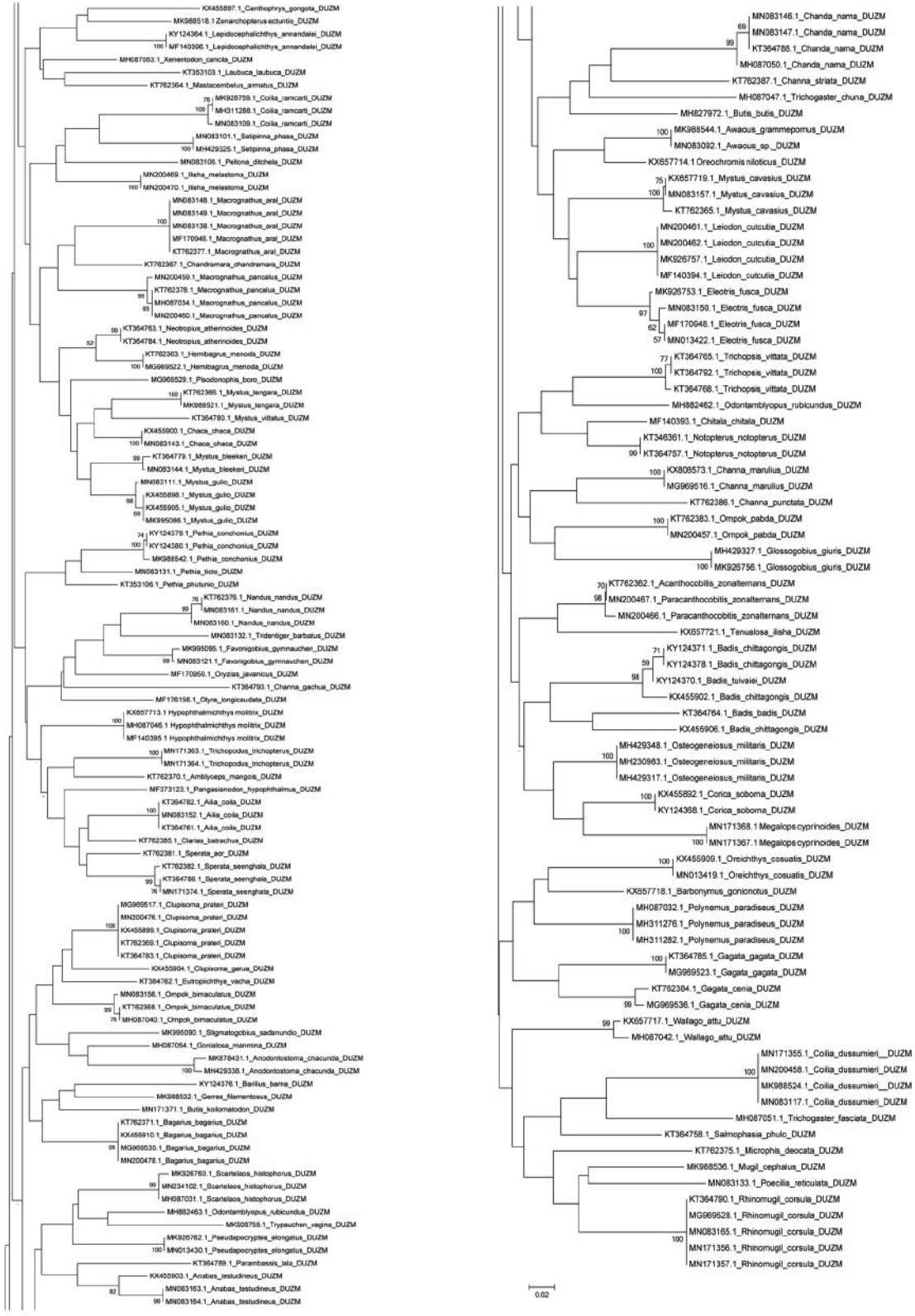

Fig. 4. Neighbor-joining (NJ) tree of freshwater fishes using K2P distances (Contd.)

A:20.3% and G:14.3%. There was a significantly higher overall GC content in the 153 species and this difference was attributable to the GC content at the 2nd and the 3rd codon base. The pattern of %GC content at different codons for all the fishes was invariably 1st>2nd>3rd (p<0.005) (Fig. 2).

Kimura 2-parameter distance values of 7.02 ±0.02%, 12.75 ± 0.03% and 18.68 ± 0.04% were obtained for within genus, within family and within order respectively (Table 2, Fig. 3). The pairwise genetic distance values were increasing at higher taxonomic levels as expected and consistent with the previous studies that supports the significant change in genetic divergence at the species boundaries (Hubert *et al.* 2008, Lakra *et al.* 2015, Ahmed *et al.* 2019). In this study, the average K2P distance within species was 0.38%, compared with 7.02% for within genera. The mean interspecific distance was found to be 18-fold higher than the mean intraspecific distance. More than 20-fold difference was observed in the freshwater fishes commonly encountered in the Australian, Canadian and Indian freshwater fishes (Ward *et al.* 2005, Hubert *et al.* 2008, Lakra *et al.* 2015). This finding corresponds to the DNA barcoding principle that interspecific divergence sufficiently outscores intraspecific divergence.

The accuracy of species identification through DNA barcoding mostly depends on both interspecific and intraspecific divergence. In our study, the average genetic distance within species was found 0.38±0.01%. Mean intraspecific genetic distance was calculated as <1% in previous studies; Hubert *et al.* (2008) found 0.30% Lakra *et al.* (2015) 0.25 % (0-0.82%) for the freshwater fishes. Phylogenetic relationships of barcoded species were shown in NJ tree (Fig. 4). Each species was associated with a specific DNA barcode cluster and the relationship among these species was clearly revealed. Closer species in terms of genetic divergence, were clustered at the same nodes and the distance between the terminal branches of the NJ tree widened as they got more distinct.

In some cases, deviations were observed. For example, out of five sequences of *Labeo bata* two form different cluster, one (MH087029.1) with *Cirrhinus reba* and another (MN083093.1) showed more close relationship with *Labeo gonius* (Fig. 4). We know that *L. bata* is widely used as an aquaculture species and its induced breeding performed almost all hatcheries of the country. Not only this, its fry and fingerlings are released in the natural habitat. So, it not unlikely to be its hybrid population in natural habitats. Moreover, among the three sequences of *Anabas testudineus*, one sequence deviated from the rest with 83% bootstrap value and we suspect this an exotic variety (Thai koi) which is now available everywhere in the country. *Channa orientalis* was documented in Bangladesh (Siddiqui *et al.* 2007, IUCN 2015) which is an endemic species of Sri Lanka, and it often misidentified as *C. gachua* (Tanomtong *et al.* 2014). The major

morphological difference between the two species is that *C. gachua* has ventral fins and *C. orientalis* lack of ventral fin. We could not find any specimen of *C. orientalis*, and this led us to believe that *C. orientalis* not existing in Bangladesh and all the older reports are pertaining to *C. gachua* (Ahmed *et al.* 2018, Conte-Grand *et al.* 2017). *Macrognathus* of Bangladesh having been referred by virtually all authors to *M. aculeatus* (Rahman 2005, Siddiqui *et al.* 2007). *M. aculeatus* body usually with 14-17 oblique dark bars whereas, *M. aral* body typically with two or more broad pale longitudinal bands of varying width extending its entire length, never with oblique bars (Roberts 1980). Our five generated sequences of collected species showed high similarity with GB pre-existing sequences and we thus confirm the presence of *M. aral* instead of *M. aculeatus*. Loaches are another diverse group of fishes with very confusing characteristics. At least three new records (*Garra nasuta, Schistura fasciolata, Botia rostrata*) have been confirmed in this study and we presume that some new species/records are still to be explored under the genus *Lepidocephalichthys, Botia*, or *Garra*. A taxonomic revision on taxa of this group is urgently needed for their biodiversity conservation.

The present study revealed that DNA barcoding has been successful in identifying the freshwater fishes of Bangladesh. We have barcoded 153 species of freshwater fishes and these barcode data confirms the 12 new records from Bangladesh. When traditional morpho taxonomy does not work, this molecular tool is effective for species identification, particularly with specimens that are damaged, incomplete, or morphologically distinct stages. Nevertheless, DNA barcoding also has its limitations too. Therefore, DNA barcoding can serve as a complementary tool for species identification, but it cannot replace the traditional morpho-taxonomy. Through this study, a reliable DNA barcode reference library for Bangladeshi freshwater fish was established, which could be used to assign fish species by screening sequences against it in the future. This could enhance to achieve better monitoring, conservation, and management of fisheries in this overexploited country.

Acknowledgements: We acknowledge the financial support from the USAID Trust Fund and the Ministry of Agriculture, Government of the Peoples Republic of Bangladesh, as a Competitive Research Grant (Grant no. CRG 471, 2017-2018) to the first author through the National Agricultural Technology Program-Phase II Project (NATP-2) from Bangladesh Agricultural Research Council (BARC), Dhaka, Bangladesh. We are also thankful to Zarif Hossain, Nusrat Jahan Sanzida, J.B.M. Aysha Akter, Nishat Jahan Chowdhury, Tonmoy Saha, Sumaiya Salam, Tamim Afrin, Anindita Barua and Noor Aida Arfin for their support during this research.

LITERATURE CITED

AHMED, M.S., CHOWDHURY, M.M.K. and NAHAR, L. 2019. Molecular characterization of small indigenous fish species (SIS) of Bangladesh through DNA barcodes. *Gene*. **684**: 53-57.

AHMED, M.S., DINA, S.R., NAHAR, L., ISLAM, N.N. and REZA, A.H. 2018. Molecular characterization of *Channa* species from Bangladesh based on Cytochrome c Oxidase Subunit I (COI) gene. *FishTaxa*. **3**(4): 87-93.

BHUIYAN, A. L. 1964. *Fishes of Dacca*. Asiatic Society of Pakistan, Dacca, 148pp.

BINGPENG, X., HESHAN, L., ZHILAN, Z., CHUNGUANG, W., YANGUO, W. and JIANJUN, W. 2018. DNA barcoding for identification of fish species in the Taiwan Strait. *PLOS One.* ***13***(6): 13(6): e0198109. doi.org/10.1371/journal.pone.0198109.

BOLD 2020. The Barcode of Life Data System (http://www.boldsystems.org).

CHANG, C.H., SHAO, K.T., LIN, H.Y., CHIU, Y.C., LEE, M.Y., LIU, S. H. and LIN, P. L. 2017. DNA barcodes of the native ray-finned fishes in Taiwan. *Mol. Ecol. Resour.* **17**(4): 796-805.

CONTE-GRAND C., BRITZ R., DAHANUKAR N., RAGHAVAN R., PETHIYAGODA R. and TAN H.H. 2017. Barcoding snakeheads (Teleostei, Channidae) revisited: Discovering greater species diversity and resolving perpetuated taxonomic confusions. *PLOS One*. **12**(9): e0184017.

COSTA, F.O., LANDI, M., MARTINS, R., COSTA, M.H., COSTA, M.E., CARNEIRO, M. and CARVALHO, G.R. 2012. A ranking system for reference libraries of DNA barcodes: application to marine fish species from Portugal. *PLOS One*. **7**(4): e35858. doi:10.1371/journal.pone.0035858

DoF (Department of Fisheries). 2018. *Yearbook of Fisheries Statistics of Bangladesh, 2017-18*. Fisheries Resources Survey System (FRSS), Department of Fisheries. Bangladesh: Ministry of Fisheries, 2018. Volume **35**:129pp.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. **39**(4): 783-791.

FRE´ZAL, L. and LEBLOIS, R. 2008. Four years of DNA barcoding: current advances and prospects. *Infect. Genet. and Evol.***8**(5):727–36.

GREEN, M. R and SAMBROOK, J. 2012. *A laboratory manual.* Cold Spring Harbor. 178pp.

GUNBIN, K., PESHKIN, L., POPADIN, K., ANNIS, S., ACKERMANN, R.R. and KHRAPKO, K. 2017. Data on the time of integration of the human mitochondrial pseudogenes (NUMTs) into the nuclear genome. *Data in brief*. **13**: 536-544.

HEBERT, P.D., CYWINSKA, A., BALL, S.L. 2003. Biological identifications through DNA barcodes. *Proc. Royal Soc. B.* ***270***(1512): 313–21.

HUBERT, N., HANNER, R., HOLM, E., MANDRAK, N.E., TAYLOR, E., BURRIDGE, M. and ZHANG, J. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLOS One*. **3**(6): 1-9.e2490 doi:10.1371/journal.pone.0002490

IUCN Bangladesh. 2000. *Red Book of Threatened Fishes of Bangladesh.* IUCN- The World Conservation Union. Xii+116PP.

IUCN Bangladesh. 2015. *Red List of Bangladesh Volume 5: Freshwater Fishes.* IUCN, International Union for Conservation of Nature, Bangladesh Country Office, Dhaka, Bangladesh. xvi+360.

KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111–120.

KNEBELSBERGER, T., DUNZ, A. R., NEUMANN, D. and GEIGER, M. F. 2015. Molecular diversity of Germany's freshwater fishes and lampreys assessed by DNA barcoding. *Mol. Ecol. Resour.* **15**(3): 562-572.

LAKRA, W.S., SINGH, M., GOSWAMI, M., GOPALAKRISHNAN, A., LAL, K.K., MOHINDRA, V., SARKAR, U.K., PUNIA, P.P., SINGH, K.V., BHATT, J.P. and AYYAPPAN, S.2015. DNA barcoding Indian freshwater fishes. *Mitochondrial DNA Part A,* **27**(6): 4510-4517. DOI: 10.3109/19401736.2015.1101540

LARA, A., PONCE DE LEÓN, J.L., RODRIGUEZ, R., CASANE, D., COTE, G., BERNATCHEZ, L., and GARCÍA-MACHADO, E.R.I.K. 2010. DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts. *Mol. Ecol. Resour.* **10**(3): 421-430.

LERAY, M., KNOWLTON, N. 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proc. Natl. Acad. Sci.* **112**(7):2076–81.

PANPROMMIN, D., SOONTORNPRASIT, K., TUNCHAROEN, S., PITHAKPOL, S. and KEEREELANG, J. 2019. DNA barcodes for the identification of species diversity in fish from Kwan Phayao, Thailand. *J. Asia-Pac Biodiv.* **12**(3):382-389.

RAHMAN A.K.A. 1989. *Freshwater Fishes of Bangladesh,*1st ed., Zool. Soc. Bangladesh, Dhaka, Bangladesh.364pp.

RAHMAN A.K.A. 2005. *Freshwater Fishes of Bangladesh,* 2nd ed., Zool. Soc. Bangladesh, Dhaka, Bangladesh. xviii+394pp.

ROBERTS, T. R. 1980. A Revision of the Asian Mastacembelid Fish Genus. American Society of Ichthyologists and Herpetologists (ASIH). *Copeia.***1980**(3): 385-391

SAITOU, N. AND NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. and Evol.* **4**(4):406-425.

SHAFI, M., and QUDDUS, M.M.A. 1982. *Bangladesher Mothsho Shampad* (In Bengali). Bangla Academy, Dhaka. 444pp.

SIDDIQUI, K.U., ISLAM, M.A., KABIR, S.M.H., AHMAD, M., AHMED, A.T.A., RAHMAN, A.K.A., HAQUE, E.U., AHMED, Z.U., BEGUM, Z.N.T., HASSAN, M.A., KHONDKER, M. AND RAHMAN, M.M. (eds.) 2009. *Encyclopedia of Flora and Fauna of Bangladesh, Vol. 23 Freshwater Fishes.* Asiatic Society of Bangladesh, Dhaka.300pp.

STEINKE, D., ZEMLAK, T.S., BOUTILLIER, J.A., and HEBERT, P.D. 2009. DNA barcoding of Pacific Canada's fishes. *Mar.Biol.* **156**(12): 2641-2647.

TALWAR, P.K. and JHINGRAN, A.G. 1991. *Inland fisheries of India and adjacent countries.* Vol. **I & II**: 1-1158.

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

TANOMTONG, A., SUPIWONG W., JEARRANAIPREPAME P., KHAKHONG S., KONGPIRONCHUEN C., GETLEKHA N. 2014. A new natural autotetraploid and chromosomal characteristics of dwarf snakehead fish, *Channa gachua* (Perciformes, Channidae) in Thailand. *Cytologi.***79**(1): 15-27.

THU, P. T., HUANG, W. C., CHOU, T. K., VAN QUAN, N., VAN CHIEN, P., LI, F. and LIAO, T. Y. 2019. DNA barcoding of coastal ray-finned fishes in Vietnam. *PLOS One.* **14**(9): e0222631. https://doi.org/10.1371/journal.pone.0222631

WANG, L., WU, Z., LIU, M., LIU, W., ZHAO, W., LIU, H., and YOU, F.2018. DNA barcoding of marine fish species from Rongcheng Bay, China. *Peer J.* **6**: e5013.

WARD, R.D., ZEMLAK, T.S., INNES, B.H., LAST, P.R. and HEBERT, P.D. 2005. DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. B.* **360**: 1847-1857.

ZHANG, D. X., and HEWITT, G. M. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol. Evol.* **11**(6): 247-251.

ZHANG, J. 2011. Species identification of marine fishes in China with DNA barcoding. *Evid.-Based Complementary Altern. Med.* 2011(978253):10 pp.doi:10.1155/2011/978253

ZHANG, J. B. AND HANNER, R. 2011. DNA barcoding is a useful tool for the identification of marine fishes from Japan. *Biochem. Syst. and Ecol.,* **39**(1): 31-42

(Manuscript received on 15 May, 2020 revised on 29 May, 2020)