Introduction

The reba carp, *Cirrhinus reba* (Hamilton, 1822) is an important freshwater labeonine fish of the family Cyprinidae and the order Cypriniformes. The species is geographically distributed over the Indian sub-continent including Bangladesh, India, Nepal, and Pakistan [1]. In Bangladesh, it is locally known as reba carp, tatkini, raik or bhagna bata and the natural habitats of the fish include the rivers, small creeks, natural depressions, and floodplains [2]. Siltation, conversion of wetlands, drying-up, and fragmentation of habitats, aquatic pollution, over fishing and climate changes have posed an alarming threat to the existence of many fish species including *C. reba*. Consequently, among 54 threatened freshwater fishes, *C. reba* is categorized as a near threatened (NT) species in Bangladesh [3]. In India, *C. reba* has been reported as vulnerable species based on International Union for Conservation of Nature (IUCN) criteria [4]. A few reports on the induced-breeding [5], stock identification [2], cryogenic preservation of sperm [1] have potentiated for the hatchery-based selective breeding, identification of the population...
status and the conservation of *C. reba*.

The mitochondrial genome is composed of a circular, double-stranded DNA with semi-autonomous central dogma machinery and situated in the mitochondria of each vertebrate cell outside the nuclear genome [6]. It possesses 37 genes in specific order that code for 13 polypeptides necessary for the oxidative phosphorylation (OxPhos) system [7]. The mitochondrial DNA of most metazoan species is predominantly inherited maternally [8]. The clonal inheritance, coupled with a substitution rate is typically 5–10 times more than that of nuclear DNA [9]. Mitochondrial DNA is an essential raw material as molecular marker for the investigation of population genetics [10], phylogenetic and evolutionary histories [11], genetic barcoding, and biodiversity of species [12], genetic disorder or mitochondrial diseases [13].

In the present study, we characterized the complete mitochondrial genome of *C. reba* using next-generation sequencing (NGS) with Illumina MiSeq platform. Sequence of the full mitochondrial genome would be an essential source to design specific molecular markers, for instance, single nucleotide polymorphisms studying population genetic structure. Moreover, detailed decoding of the mitochondrial genome would be a valuable insight for resolving any conservation and management strategies of the culture promising but NT fish.

**Methods**

**Specimen collection, genomic, and mitochondrial DNA isolation**

The specimen was collected from Khulna, Bangladesh (22°50′44.3″N, 89°32′27.7″E) in March 2017. The specimen is stored at the Department of Biotechnology, Sher-e-Bangla Agricultural University, Sher-e-Bangla Nagar, Dhaka 1207, Bangladesh. The genomic DNA was extracted from the fish tissue using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer’s manual. The mitochondrial partial COI sequence was amplified for the species confirmation using the BCL-BCH primers [14]. The specimen was confirmed based on both morphological characteristics and its COI sequence with 100% identity to the GenBank database (GenBank no. AP013325). The mitochondrial DNA was extracted with a commercially available mitochondrial DNA isolation kit (Abcam, Cambridge, UK) followed by the manufacturer’s protocol and DNA concentration was checked by qubit fluorometer.

**Library construction and NGS**

In advance to the library preparation, the purified mitochondrial DNA was fragmented as 300–350 bp by Covaris M220 Focused-Ultrasonicator (Covaris Inc., Woburn, MA, USA). A library was prepared by TruSeq RNA library preparation kit V2 (Illumina, San Diego, CA, USA) which needed a couple of steps including end repair reaction, adenylation at 3′ ends, adaptor ligation, and enrichment of DNA fragments by PCR. To have better quality DNA, the enriched DNA library templates were purified using DNA purification kit (RBC Bioscience, Jerusalem, Israel) and quantified using qubit fluorometer. Size, purity, and quality of the constructed DNA libraries were further validated by the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Finally, the NGS was performed using the Illumina MiSeq platform 2 x 300 bp pair ends (Illumina).

**Sequence assembly and annotation of the mitochondrial genome**

The MiSeq raw reads were assembled using the Geneious Prime 2020.0.3 software [15] to construct a complete mitogenome of *Cirrhinus reba*. The gene order, sequences, and sizes of each of the 13 protein-coding genes and 2 ribosomal RNA genes were organized with the ORF finder program (https://www.ncbi.nlm.nih.gov/orffinder/) following a reference mitogenome of *Cirrhinus reba* (GenBank no. AP013325). All transfer RNA genes were identified, the positions of their anticodons were predicted by using the ARWEN software [16].

**Gene mapping and phylogenetic tree construction**

The circular gene map of *C. reba* was drawn by OGDRAW software (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html). The phylogenetic tree was constructed by MEGA7 program using the Minimum Evolution (ME) algorithm [17] performing with 1,000 bootstrap replications.

**Results and Discussion**

**Genome organization**

The complete mitochondrial genome of *C. reba* (GenBank no. MN862482) is a closed circular molecule of 16,597 bp in size. It possesses the typical combination of 37 genes including 13 protein-coding genes, 22 tRNA genes, two genes for ribosomal RNA subunits (12SrRNA and 16SrRNA), and two non-coding regions (control region, D-loop, and origin of light strand, OL) (Table 1, Fig. 1). Twenty-eight genes were located on the heavy (H) strand, while the ND6 gene and eight tRNA genes were transcribed from the light (L) strand (Fig. 1).

The overall nucleotide composition of *C. reba* mitogenome was found biased toward A + T contents (58.93%, A = 32.91% and T = 26.02%) over the G + C contents (41.07%, G = 15.16% and C = 26.02%).
Table 1. Organizational characteristics of the complete mitochondrial genome of *Cirrhinus reba*

| Gene/element | Strand | Nucleotide position | Size (bp) | Codon | Anti-codon | Intragenic nucleotides | A + T (%) |
|--------------|--------|---------------------|----------|-------|------------|------------------------|-----------|
| tRNA^Phe^   | H      | 1 - 70              | 70       | -     | -          | GAA                    | 52.9      |
| tRNA^12S^RN| H      | 71 - 1,024          | 954      | -     | -          | TAC                    | 51.0      |
| tRNA^Val^   | H      | 1,025 - 1,096       | 72       | -     | -          | TAG                    | 50.0      |
| tRNA^16S^RN| H      | 1,097 - 1,867       | 776      | -     | -          | TAA                    | 57.6      |
| tRNA^Leu^   | H      | 2,784 - 2,860       | 77       | -     | -          | TAA                    | 48.1      |
| ND1         | H      | 2,861 - 3,835       | 975      | ATG   | TAA        |                        | 58.6      |
| tRNA^12S^RN| H      | 3,839 - 3,912       | 74       | -     | -          | GAT                    | 54.1      |
| tRNA^Glu^   | L      | 3,910 - 3,980       | 71       | -     | -          | TAG                    | 60.6      |
| tRNA^Met^   | H      | 3,982 - 4,050       | 69       | -     | -          | CAT                    | 42.0      |
| ND2         | H      | 4,051 - 5,096       | 1,046    | ATG   | TAA        |                        | 59.6      |
| tRNA^Pro^   | H      | 5,096 - 5,167       | 72       | -     | -          | TCA                    | 63.9      |
| tRNA^Ala^   | L      | 5,169 - 5,237       | 69       | -     | -          | TGC                    | 69.6      |
| tRNA^Glu^   | L      | 5,239 - 5,311       | 73       | -     | -          | GTG                    | 52.1      |
| OL          | -      | 5,312 - 5,345       | 34       | -     | -          |                        | 52.9      |
| tRNA^Cys^   | L      | 5,346 - 5,412       | 67       | -     | -          | GCA                    | 52.2      |
| tRNA^Asp^   | L      | 5,414 - 5,483       | 70       | -     | -          | GTA                    | 48.6      |
| COX1        | H      | 5,485 - 7,035       | 1,551    | GTG   | TAA        |                        | 58.3      |
| tRNA^Ser^   | L      | 7,037 - 7,105       | 69       | -     | -          | TGA                    | 53.6      |
| tRNA^Asp^   | L      | 7,110 - 7,181       | 72       | -     | -          | GTC                    | 62.5      |
| COX2        | H      | 7,197 - 7,887       | 691      | ATG   | T--        | -1                     | 59.6      |
| tRNA^Lys^   | H      | 7,887 - 7,964       | 78       | -     | -          | TTT                    | 55.1      |
| ATPB        | H      | 7,965 - 8,129       | 165      | ATG   | TAG        | -7                     | 61.8      |
| ATP6        | H      | 8,123 - 8,806       | 684      | ATG   | TAA        | -1                     | 61.4      |
| COX3        | H      | 8,806 - 9,591       | 786      | ATG   | TAA        |                        | 55.9      |
| tRNA^Glu^   | H      | 9,592 - 9,663       | 72       | -     | -          | TCC                    | 66.7      |
| ND3         | H      | 9,664 - 10,012      | 349      | ATG   | T--        | 0                      | 59.3      |
| tRNA^Arg^   | H      | 10,013 - 10,082     | 70       | -     | -          | TCG                    | 51.4      |
| ND4L        | H      | 10,083 - 10,379     | 297      | ATG   | TAA        | -7                     | 55.6      |
| ND4         | H      | 10,373 - 11,753     | 1,381    | ATG   | T--        | 0                      | 59.1      |
| tRNA^His^   | H      | 11,754 - 11,822     | 69       | -     | -          | GTG                    | 69.6      |
| tRNA^Ser^   | H      | 11,822 - 11,892     | 71       | -     | -          | GCT                    | 59.2      |
| tRNA^Glu^   | H      | 11,892 - 11,966     | 75       | -     | -          | TAG                    | 60.0      |
| ND5         | H      | 11,969 - 13,792     | 1,824    | ATG   | TAA        | -4                     | 62.2      |
| ND6         | L      | 13,789 - 14,310     | 522      | ATG   | TAA        | -1                     | 57.7      |
| tRNA^His^   | L      | 14,310 - 14,380     | 71       | -     | -          | TTC                    | 62.0      |
| CYTB        | H      | 14,385 - 15,525     | 1,141    | ATG   | T--        | 0                      | 60.1      |
| tRNA^Trp^   | H      | 15,526 - 15,598     | 73       | -     | -          | TGN                    | 47.9      |
| tRNA^Pro^   | L      | 15,595 - 15,667     | 73       | -     | -          | TGG                    | 57.5      |
| Control region (D-loop) | - | 15,668 - 16,597 | 930 | - | - | - | 67.6 |

= 25.91%) indicating a strong anti-guanine bias commonly observed in fishes [18]. This pattern was almost the same as the Cyprinid species, which has 15.70% of G content [19]. A total of 31 bp overlapping region was found in 11 different locations of the *C. reba* mitogenome. Besides, 11 intergenic spacers were present in the mitogenome involving a total size of 34 bp. The longest spacer included a nucleotide sequence of 15 bp which located between the tRNA Asp and COX2 genes. As shown in Table 1, the positive (+) numbers indicate intergenic space/gap between genes, the negative (–) numbers indicate overlapping between genes, and the zero (0) revealed either overlap or space does not exist between genes in the *C. reba* mitogenome.
**Protein-coding genes**

The mitochondrial genome of *C. reba* is consisted of 13 canonical protein-coding genes (PCGs) of 11,412 bp in length, accounted for 68.76% of the total mitogenome. With exception to the reading frame of ND6 gene, which was oriented clock-wise direction on the light strand, reading frames of all other PCGs were oriented counter-clock wise direction on the heavy strand (H). The shortest PCG was ATP8 (165 bp), whereas the longest one was ND5 (1,824 bp). Except for the COX1 gene with an unusual and alternative start codon, GTG (Valine), other twelve PCGs started with the typical translation initiation codon, ATG (Methionine). The open reading frame of ATP8 used TAG as stop codon, whereas, the remaining seven PCGs (*ND1, COX1, ATP6, COX3, ND4L, ND5, and ND6*) terminated with the typical stop codon (TAA), as in other vertebrates [20]. Four PCGs, COX2, ND3, ND4, CYTB ended with an incomplete termination codon (T--), while the ND2 gene ended with “TA-” (Table 1). These incomplete stop codons are assumed to be completed by using the posttranscriptional polyadenylation, poly-A tail [21]. These kinds of variations in the termination codons are not uncommon among vertebrate mitochondrial genes and previously reported in other bony fishes [22].

Analysis of the sequences of 13 PCGs showed that overall percentage of A and T contents (58.93%) also reflected in the codon usage of PCGs where the first, second and third position of the codons estimated the A plus T contents as 49.6%, 58.9%, and 69%, respectively (Table 2). The G content (15%) of all PCGs presents obvious anti-guanine characteristics as similar to other bony fishes [23]. Here, we observed that the 2nd and 3rd position of the codons were dominated by T (40.4%) and A (46.7%) contents, respectively over the C and G contents (Table 2). A very strong bias against the guanine contents (4.9%) was also demonstrated at the wobble or third position of codons across all 13 PCGs.

**Ribosomal RNA and transfer RNA genes**

The mitogenome of *C. reba* possesses two genes encoding two ribosomal RNA subunits, a small (12S) and a large (16S) which were typical as in other mitogenomes. Both the ribosomal RNA genes were located on the H-strand and consisted of 15.91% (2,641 bp) of the total circular mitogenome. The 12S rRNA (946 bp in length) located between the tRNA<sup>Val</sup> and tRNA<sup>Val</sup>, whereas...
The 16S rRNA was 1,695 bp in length and was located between tRNA\textsubscript{Val} and tRNA\textsubscript{Leu}. The A + T content (55.24%) was higher than the G plus C content (44.76%) in the ribosomal RNAs where the nucleotide, adenine was more prevalent (35.25%) followed by cytosine (24.50%) as similar to the previous reports in other bony fishes\textsuperscript{[24]}.

The 22 tRNA genes were identified in the mitochondrial of \textit{C. reba} which demonstrated a varied size of 67–78 bp in length, estimating a total length of 1,580 bp (~9.6% of the total mitogenome) (Table 1). Among all tRNA genes, eight tRNAs were located on the L-strand, whereas the remaining fourteen tRNA genes were on the H-strand (Fig. 1). The highest content of A and T was observed in tRNA\textsubscript{Ala} and tRNA\textsubscript{His} (69.6%), whereas that of the lowest content was observed in tRNA\textsubscript{Met} (42%) which was consisted of previously published reports in other teleosts\textsuperscript{[24]}. All these 22 tRNA genes were predicted to capable of folding typical cloverleaf secondary structures and showed great similarity as it is in other vertebrate mitogenomes\textsuperscript{[23]}.

Non-coding regions

Like other animal mitochondrial genome, \textit{C. reba} contained no introns and possessed two non-coding regions, an OL and a control region or displacement loop (D-loop) region. The OL region, consisted of 34 nucleotides was situated between tRNA\textsuperscript{Asn} and tRNA\textsuperscript{Cys} and was oriented on the L-strand in a cluster of five tRNA genes (WANCY region). D-loop region of \textit{C. reba} consisted of 930 nucleotides which represented 5.6% of the total mitogenome. The control region was overpresented by A plus T content (67.6%). The AT-rich D-loop contains promoters and an origin of replication of mtDNA which are essential for transcription and replication of mtDNA, respectively\textsuperscript{[25]}. The D-loop region is very flexible to size variation.

### Phylogenetic relationship analysis

The phylogenetic relationship of \textit{C. reba} with other labeonine fishes was constructed using 13 complete mitochondrial genomes of the subfamily Labeoninae by MEGA 7.0 program. The mitogenome of our species of interest, \textit{C. reba} (GenBank no. MN862482) showed 99.6% sequence identity with different haplotype of \textit{C. reba} (GenBank no. AP013325, deposited from Japan), followed by 90.18% identity with \textit{Labeo bata} (AP011198). Slight variations between two haplotypes of \textit{C. reba} might reflect different populations of the same species. However, two haplotypes of \textit{C. reba} was clearly isolated from other 11 species of labeonine fishes and formed a separate clade (Fig. 2). Two close relatives of \textit{C. reba}, \textit{L. pangusia}, and \textit{C. cirrhosis} have already been considered globally as

---

### Table 2. Base composition of PCGs of \textit{Cirrhinus reba} mitochondrial genome

| PCGs | T   | C   | A   | G   | A + T | Total |
|------|-----|-----|-----|-----|-------|-------|
| ND1  | 27.2| 27.7| 31.4| 13.7| 58.6  | 975   |
| ND2  | 24.0| 29.3| 35.6| 11.1| 59.6  | 1,046 |
| COX1 | 30.0| 24.8| 28.3| 16.8| 58.3  | 1,551 |
| COX2 | 27.5| 25.3| 32.1| 15.1| 59.6  | 691   |
| ATP8 | 26.7| 26.7| 35.2| 11.5| 61.8  | 165   |
| ATP6 | 30.0| 25.7| 31.4| 12.9| 61.4  | 684   |
| COX3 | 25.8| 28.4| 30.0| 15.8| 55.9  | 786   |
| ND3  | 30.4| 27.2| 28.9| 13.5| 59.3  | 349   |
| ND4L | 28.6| 29.3| 26.9| 15.2| 55.6  | 297   |
| ND4  | 26.5| 27.9| 32.6| 13.0| 59.1  | 1,381 |
| ND5  | 27.5| 26.0| 34.7| 11.8| 62.2  | 1,824 |
| ND6  | 44.1| 11.1| 13.6| 31.2| 57.7  | 522   |
| CYTB | 28.0| 26.2| 32.2| 13.7| 60.1  | 1,141 |

Position in the codon

|        | 1st | 2nd | 3rd | Across all PCGs |
|--------|-----|-----|-----|-----------------|
|        | 24.2| 40.4| 22.3| 28.9            |
|        | 24.3| 27.0| 26.1| 25.8            |
|        | 25.4| 18.6| 46.7| 30.2            |
|        | 26.1| 14.0| 4.9 | 15.0            |
|        | 49.6| 58.9| 69.0| 59.2            |
|        | 3,007| 3,803| 3,802| 11,412         |

Reverse complementary sequence was used for \textit{ND6} gene which was encoded on the L-strand. PCG, protein-coding gene.
NT and vulnerable species, respectively [3]. Therefore, we alarmingly recommend further studies relating to conservation of the labeonine fish including reba carp.

ORCID

Mohammad Nazrul Islam: https://orcid.org/0000-0001-5692-3566
Shirin Sultana: https://orcid.org/0000-0002-0607-1497
Md. Jobaidul Alam: https://orcid.org/0000-0002-3594-8147

Authors’ Contribution

Conceptualization: MNI, MJA. Data curation: MNI, MJA. Formal analysis: MNI, MJA. Funding acquisition: MNI. Methodology: MNI, MJA. Writing - original draft: MNI, SS. Writing - review & editing: SS, MJA.

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

References

1. Sultana S, Hasan MM, Hossain MS, Islam MR, Salimullah M, Sarder MR, et al. Cryogenic preservation of critically endangered Cirrhinus reba fish sperm. Int J Fish Aquat Stud 2017;5:334-339.
2. Ethin R, Hossain MS, Roy A, Rutegwa M. Stock identification of minor carp, Cirrhinus reba, Hamilton 1822 through landmark-based morphometric and meristic variations. Fish Aquat Sci 2019;22:12.
3. International Union for Conservation of Nature. Bangladesh: Red Book of Threatened Fishes of Bangladesh. Gland: International Union for Conservation of Nature, 2015.
4. Sarkar UK, Negi RS, Deepak PK, Singh SP, Srivastava SM, Roy D. Captive breeding of vulnerable Indian carp Cirrhinus reba with Ovaprim for conservation of wild populations. Aquac Asia 2004;9:5-7.
5. Rahman MA, Zaher M, Azimuddin KM. Development of fingerling production techniques in nursery ponds for the critically endangered reba carp, Cirrhinus ariza (Hamilton, 1807). Turk J Fish Aquat Sci 2009;9:165-172.
6. Artuso L, Romano A, Verri T, Domenichini A, Argenton F, Santorelli FM, et al. Mitochondrial DNA metabolism in early development of zebrafish (Danio rerio). Biochim Biophys Acta 2012;1817:1002-1011.
7. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. Nature 1981;290:457-465.
8. Giles RE, Blanc H, Cann HM, Wallace DC. Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci U S A 1980;77:6715-6719.
9. Brown WM, George MJr, Wilson AC. Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A 1979;76:1967-
10. Ingman M, Gyllensten U. mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. Nucleic Acids Res 2006;34:D749-D751.

11. Macey JR, Papenfuss TJ, Kuehl JV, Fourcade HM, Boore JL. Phylogenetic relationships among amphibian reptiles based on complete mitochondrial genomic sequences. Mol Phylogenet Evol 2004;33:22-31.

12. Meusnier I, Singer GA, Landry JF, Hickey DA, Hebert PD, Hajibabaei M. A universal DNA mini-barcode for biodiversity analysis. BMC Genomics 2008;9:214.

13. Saxena N, Taneja N, Shome P, Mani S. Mitochondrial donation: a boon or curse for the treatment of incurable mitochondrial diseases. J Hum Reprod Sci 2018;11:3-9. 29681709

14. Handy SM, Deeds JR, Ivanova NV, Hebert PD, Hanner RH, Ormos A, et al. A single-laboratory validated method for the generation of DNA barcodes for the identification of fish for regulatory compliance. J AOAC Int 2011;94:201-210.

15. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 2012;28:1647-1649.

16. Laslett D, Canback B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 2008;24:172-175.

17. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 2016;33:1870-1874.