SPECIES IDENTIFICATION AND MITOCHONDRIAL GENOME SEQUENCING OF FEATHERBACK CHITAL FISH (*Chitala chitala*)

MAHMUDA AKTER

DEPARTMENT OF BIOTECHNOLOGY

SHER-E-BANGLA AGRICULTURAL UNIVERSITY

SHER-E-BANGLA NAGAR, DHAKA-1207, BANGLADESH

JUNE, 2022

SPECIES IDENTIFICATION AND MITOCHONDRIAL GENOME SEQUENCING OF FEATHERBACK CHITAL FISH (*Chitala chitala*)

BY

MAHMUDA AKTER

REGISTRATION NO. 20-11123

A Thesis
Submitted to the Faculty of Agriculture,
Sher-e-Bangla Agricultural University, Dhaka,
in partial fulfillment of the requirements
for the degree of

MASTERS OF SCIENCE
IN
BIOTECHNOLOGY

SEMESTER: JANUARY- JUNE, 2022

Approved by

Dr. Mohammad Nazrul Islam **Associate Professor** **Department of Biotechnology** **Sher-e Bangla Agricultural University** **Sher-e-Bangla Nagar, Dhaka-1207** **Supervisor**	**Dr. Shirin Sultana** **Senior Scientific Officer & Head** **Fisheries Biotechnology Division** **National Institute of Biotechnology (NIB),** **Ganakbari, Savar, Dhaka-1349** **Co-Supervisor**

Dr. Md Rafiqul Islam
Associate Professor
Department of Biotechnology, SAU
Chairman
Examination Committee

DEPARTMENT OF BIOTECHNOLOGY
Sher-e-Bangla Agricultural University
Sher-e-Bangla Nagar, Dhaka-1207

CERTIFICATE

This is to certify that the thesis entitles,**" SPECIES IDENTIFICATION AND MITOCHONDRIAL GENOME SEQUENCING OF FEATHERBACK CHITAL FISH (*Chitala chitala*)"** submitted to the Faculty of Agriculture, Sher-e-Bangla Agricultural University, Dhaka in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE (MS) IN BIOTECHNOLOGY,** embodies the result of a piece of bonafide research work carried out by **MAHMUDA AKTER,** Registration No. **20-11123** under my supervision and my guidance. No part of the thesis has been submitted for any other degree or diploma.

I further certify that such help or source of information, as has been availed of during the course of this investigation has duly been acknowledged.

Date: June, 2022

Place: Dhaka, Bangladesh

(Dr. Mohammad Nazrul Islam)

Associate Professor

Department of Biotechnology

Sher-e Bangla Agricultural University

Supervisor

DEDICATED TO MY

BELOVED PARENTS

AND

MY BROTHERS

LIST OF ABBREVIATIONS AND ACRONYMS

ABBREVIATION	FULL NAME
µl	Micro Liter
12S rRNA	12S Ribosomal Ribonucleic Acid
16S rRNA	16S Ribosomal Ribonucleic Acid
A	Adenine
ATP6	Adenosine Triphosphate 6
BLAST	Basic Local Alignment Search Tool
Bp	Base pair
C	Cytosine
CEB	Cytosol Extraction Buffer
CI	Chloroform:Isoamyl Alcohol
COI	Cytochrome C Oxidase Subunit I
Cytb	Cytochrome b
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide Triphosphate
DoF	Department of Fisheries
EDTA	Ethylenediaminetetraacetic Acid
EM	Enzyme Mix
EtBr	Ethidium Bromide
FAO	Food and Agriculture Organization
FASTA	Fast Adaptive Shrinkage Thresholding Algorithm
FBD	Fisheries Biotechnology Division
FRSS	Fisheries Resources Survey System

ABBREVIATION	FULL NAME
G	Guanine
GB	GeneBank
GDP	Gross Domestic Product
gDNA	Genomic DNA
GLD	Gel Loading Dye
HCl	Hydrochloric Acid
IUCN	International Union for Conservation of Nature
kb	Kilobase pair
MCL	Maximum Composite Likelihood
MEGA	Molecular Evolutionary Genetic Analysis
min	Minute
ML	Maximum Likelihood
MLB	Mitochondrial Lysis Buffer mM
MT	Metric ton
mtDNA	Mitochondrial Deoxyribonucleic Acid
NCBI	National Centre for Biotechnology Information
NIB	National Institute of Biotechnology
PBS	Phosphate Buffered Saline
PCI	Phenol:Chloroform:Isoamyl Alcohol
PCR	Polymerase Chain Reaction
RFLP	Restriction Fragment Length Polymorphism
rpm	Rotation per minute
T	Thymine

ACKNOWLEDGEMENTS

At first the author expresses her gratefulness to the Almighty Allah, the creator and sustainer of the universe for giving opportunity and ability to complete her research work and to prepare this manuscript for the degree of Master of Science (M.S.) in Biotechnology.

The author expresses her heartfelt gratitude, indebtedness, profound appreciation to her honorable Supervisor ***Dr. Mohammad Nazrul Islam****, Associate Professor, Department of Biotechnology, Sher-e-Bangla Agricultural University, (SAU), Dhaka for his untiring* inspiration,*scholasticguidance,valuable suggestions, encouragement and constant supervision in carrying out this research work as well as preparing the manuscript of this thesis.*

The author expresses her heartiest gratitude, sincere appreciation, ever indebtedness and deepest sense of respect to her co-supervisor, ***Dr. Shirin Sultana****, Senior Scientific Officer, Fisheries Biotechnology Division (FBD), National Institute of Biotechnology (NIB) for her scholarly guidance, innovative suggestions, kind cooperation and untiring perseverance throughout this research work.*

The author wants to express her heartiest thanks to the honorable Chairman, ***Dr. Md. Rafiqul Islam****, Associate Professor, Department of Biotechnology, Sher-e-Bangla Agricultural University, for his constant inspiration and utmost co-operation to conduct the research work.*

She is also grateful and expresses her deepest sense of respect to all of her honorable teachers at the Department of Biotechnology, SAU and all the scientific officers at the Fisheries Biotechnology Division, NIB for their valuable advice and help during the course of the study period.

The author feels heartfelt gratitude to the honorable Director General, Dr. Md. Salimullah, NIB, for his utmost support and permission to carry out this research at the FBD laboratory.

The author is very much grateful to specially Md. Nasir Khan and Moonmoon Ahamed Mim , Post Graduate Fellow, FBD, NIB for their cordial help, constant company and co-operation to the author in research experiment and thesis preparation.

The author is also thankful to Md. Moinul Hossain, Scientific Assistant and Md. Habibur Rahman, Laboratory Attendant, FBD, NIB for their help and friendly

behavior. The author says special thanks to Sania Sarwar, seniors, juniors, and well-wishers for their encouragement during the study period.

The author is also grateful to the Ministry of National Science and Technology for selecting her as an NST fellow and making significant cooperation to undertake this research work.

Finally, the author owes her deepest gratefulness to her beloved parents and her brothers. They sacrificed a lot for carrying out this study and the constant source of inspiration. Their sacrifice, blessings and inspiration have made all these possible.

Dated: June, 2022 ***The Author***

SAU, Dhaka

LIST OF CONTENTS

CONTENTS

CONTENTS (CONTINUED)

CONTENTS (CONTINUED)

CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF FIGURES

SPECIES IDENTIFICATION AND MITOCHONDRIAL GENOME SEQUENCING OF FEATHERBACK CHITAL FISH (*Chitala chitala*)

ABSTRACT

The purpose of this study was to characterize the mitochondrial genome sequencing of *Chitala chitala*. The complete mitogenome of *C. chitala* was 16248 bp long and contains 37 mitochondrial genes, including 13 typical protein coding genes, 22 tRNA genes, two ribosomal RNAs (12SrRNA and 16SrRNA), and two non-coding areas (control region, D-loop, and origin of light strand, OL). The heavy (H) strand encoded 28 genes, while the light (L) strand encoded the remaining 9 genes. A total of 31 bp of overlapped area was discovered across the *C. chitala* mitogenome in 13 distinct places. The mitogenome contained six intergenic spacers totaling 24 bp in length. The longest spacer was an 8-bp nucleotide sequence located between the $tRNA^{Leu}$ and ND1 genes. The 13 protein-coding genes (PCGs) were 11,423 bp in length and accounted for 70.30% of the mitogenome. The base composition was 25.16% T, 29.84% C, 31.4% A, and 13.5% G. *C. chitala*'s circular genome had a short subunit of rRNA (12S rRNA) and a large subunit of rRNA (16S rRNA), both of which were 956 bp and 1702 bp in length, respectively. The 12S rRNA gene had an overall base composition of A = 32.74%, T = 21.75%, C = 25%, and G = 20.50%, while the 16S rRNA gene had an overall base composition of A = 36.13%, T = 19.03%, C = 24.61%, and G = 20.21%. tRNA genes ranged in length from 67 to 76 bp, for a total length of 1,570 bp (9.6% of the total mitogenome). Fourteen tRNA genes were transcribed on the H-strand, while the remaining eight tRNA genes were transcribed on the L-strand. *C. chitala*'s largest non-coding region (control region) consists of 572 nucleotides, accounting for 3.5% of the total mitogenome. The morphologically detected fish species shared 99% of their DNA with *C. ornate* (Accession No. AP008923.1). Since there is no verified complete mitogenome of *Chitala chitala*. Currently, it has been considering as a provisional reference sequence (NC_070068.1).

CHAPTER- I

INTRODUCTION

CHAPTER I

INTRODUCTION

Bangladesh is one of the top fish producers in the world, and it has a wide variety of fisheries resources that can be largely divided into inland and marine fisheries. About 47.60 lakh acres are covered by inland fisheries, which are divided into two subsectors: inland capture and inland culture (FRSS, 2019). Rivers, estuaries, beels, Kaptai Lake, and flood plains make up the 39.27 lakh ha of inland capture, whereas ponds, ditches, baors, shrimp/prawn farms, cage- and pen-cultured fish, and seasonal cultured water bodies make up the 8.33 lakh ha of inland culture. A total of 1,18,813 km^2 and 200 nautical miles of EEZ are covered by marine capture fisheries from the baseline (DoF, 2022). The fish industry is one of the most productive and dynamic industries that has contributed significantly to the economy over the past few decades. Since gaining its independence in 1971, Bangladesh has made amazing advancements in the fisheries sector.

The fish industry is crucial for employment, sustenance, and earning foreign currency. It contributes more than one-fourth (26.50%) of the total agricultural GDP and 3.57% of the national GDP (DoF,2022). The average growth rate of this industry over the past ten years has been close to 5.43%. According to the FAO (2018), Bangladesh is the third-largest producer of inland open water fish, fifth in the world for aquaculture production, and eleventh for marine fish. Fish alone provides 60% of all animal protein, and each individual consumes 63g of fish each day (DoF,2022). Inspite of Covid-19 pandemic situation, the performance of this sector seems quite amazing.The total fish production reached at 46.21 lac Metric Ton (MT) in FY 2020-21 which exceeds the targeted fish production of 45.52 lac MT (DoF, 2022).

Fisheries and aquaculture are Bangladesh's second-largest export production and the main contributors to export earnings since the country produces and exports a variety of fish (Ahmed *et al*., 2019). The European Union (EU), the United States, and Japan are Bangladesh's top export destinations for fish and fishery products (Shamsuzzaman *et al*., 2020).

In our country, numerous fish species, including *Chitala chitala*, are alarmingly threatened by habitat fragmentation, siltation, overfishing, and other factors.

Chitala chitala is a freshwater primitive fish with evolutionary significance and widely dispersed throughout the Indian subcontinent, including Bangladesh, India, Nepal, Myanmar, and Pakistan. It is also known as giant featherback and clown knife fish (Chonder, 1999; Mitra *et al*., 2018).Although *C. chitala* typically grows to a length of approximately 75 cm (30 in), it can grow as long as 122 cm (48 in). In general, it has a silvery hue. Its back is covered in a row of golden or silvery bands, unlike all of its relatives. Besides, there are a number of relatively small, occasionally blurry dark dots. High nutritional value is included in this fish. Patients with measles are given instructions to consume fish flesh-based soup. It is a potential candidate species for aquaculture (Mandal *et.al*., 2012).

Among 64 threatened freshwater fishes , *Chitala chitala* is categorized as an endangered (EN) species in Bangladesh based on conservation status of freshwater fishes of Bangladesh (IUCN Bangladesh, 2015).

In order to support the biological activities of a cell, mitochondria can directly convert organic matter into energy (Avise *et al*., 1987; Wataru *et al*., 2013; Strohm *et al*., 2015; Parhi *et al*., 2019). Mitochondrial DNA, also known as mtDNA or mDNA, is the DNA found in mitochondria. It has a closed circular double-stranded structure (Prosdocimi *et al*., 2012). It is the second genetic information system in eukaryotic cells (Kim *et al*., 2008 , Cooke *et al*., 2012, Zhao *et al*., 2015 and Ruan *et al*., 2020). In eukaryotic cells, it serves as the primary site of ATP synthesis and oxidative phosphorylation (Wilson *et al*., 1985). Mitochondrial DNA is a relatively independent replication unit that exhibits maternal inheritance, a small size, a simple composition, quick progression, limited recombination, and heterogeneity in evolutionary rate at different loci (Harrison, 1989 and Javonillo *et al*., 2010). The mitochondrial genome of fish is a circular molecule that is about 15–18 kb in size and typically codes for 37 genes, 13 protein-coding genes, 2 rRNAs, and 22 tRNAs (Gray *et al*., 1989, Kim *et al*., 2009 and Alam *et al*., 2014). However, gene intervals and lengths differ between species.13 mitochondrial gene-encoded proteins direct cells to make the protein subunits of the enzyme complexes of the oxidative phosphorylation system that enable mitochondria to respire, serving as the cell's powerhouses. In phylogenetic and evolutionary studies of fishes, mitochondrial DNA sequences of numerous fish species have been identified, making them popular molecular guides (Brown *et al*., 1979; Wang *et al*., 2016; Wu *et al*., 2020).

Though small in size, the mitochondrial genome is responsible for ensuring that the powerhouses of cells function accurately. It acts asmodulators for regulating cellular metabolism, including the tricarboxylic acid (TCA) cycle, oxidative phosphorylation (OXPHOS), fatty acid metabolism, amino acid metabolism and nucleotide metabolism. Adenosine triphosphate (ATP) is the main energy source for maintaining cellular function (Clayton *et al*., 1992). For vertebrates, the mitochondrial genome has been extensively used as a useful tool. It is used as a molecular marker to study the evolutionary history and phylogenetic analyses of fish (Brown *et al*., 2008).

The maternal inheritance, high mutation rate, and high copy number per cell are only a few of the unique characteristics of mitochondrial genes. It is essential for controlling oxidative stress, apoptosis, and cellular metabolism (Burger *et al*., 2003).

Protein-coding genes are translated into messenger RNA (mRNA) molecules, which are then translated into polypeptide chains. Importance of the PCGs are given below:

NADH (Nicotinamide Adenine Dinucleotide) Dehydrogenase 1 protein is produced according to instructions from the ND1 gene. This protein is a part of complex I, a large enzyme complex. Complex I is one of numerous enzyme complexes required for oxidative phosphorylation. Within mitochondria, these complexes are contained in a compact folded structure. According to Shaklee *et al.* (1990), mitochondrial enzyme complexes compel the production of ATP during oxidative phosphorylation.

The ND2 gene promotes the activity of NADH dehydrogenase and participates in mitochondrial electron transport. According to Satoh *et al*., (2016), it is a component of mitochondrial respiratory chain complex I.

COX1 (Cytochrome C Oxidase Subunit 1) is a protein that helps to catalyze the reduction of water into oxygen in Eukaryotes. It is also encodes as an important enzyme that involves in the oxidation phosphorylation pathway and production of energy (Morita Ikuo, 2002).

COX-2 has been recognized and characterized as an essential moderator in a variety of physiologic and pathologic settings of fish such as ovulation, immunity and adipogenesis (Morita Ikuo, 2002).

The ATP6 gene encodes a protein that is necessary for normal mitochondrial function. A subunit of mitochondrial ATP synthase, which is present in both the inner mitochondrial membrane and the thylakoid membrane, is encoded by the ATP8 gene. According to Boominathan *et al*. (2016), mitochondrial ATP synthase uses an electrochemical gradient of protons across the inner membrane to catalyze ATP production during oxidative phosphorylation.

Mitochondrial membrane respiratory chain is made up of the ND3, ND4L, and ND4 genes, which are essential for catalyzing the transfer of electrons from NADH to ubiquinone, the electron acceptor, along the respiratory chain (Cardol *et al*., 2006).

CYTB gene provides directions to make a protein called cytochrome b. This protein plays a vital role in structures called mitochondria which convert the energy from food into a form that cells can use. Cytochrome b is one of 11 components of a group of proteins that is called complex III. This complex III performs one step of a process known as oxidative phosphorylation in which oxygen and simple sugars are used to create cell's main energy source adenosine triphosphate (ATP). In mitochondria, during oxidative phosphorylation, the protein complexes including complex III, force the production of ATP through a step-by-step transfer of negatively charged particles called electrons. It is involved to transfer these particles through complex III. According to Farias *et al.* (2001), the mitochondrial cytochrome b gene is also frequently utilized as a molecular marker for evolutionary relationships at different levels within the fish family.

Protein-coding genes make up only around 1% of DNA; the majority (99%) of the DNA is noncoding. Protein synthesis instructions are not found in noncoding DNA. Non-coding DNA was once considered to be "junk," with no known function. It is now evident that at least some of it, specifically the regulation of gene activity, is essential for cells to operate. For instance, regulatory elements in non-coding DNA are sequences that control when and where genes are turned on and off. These elements offer locations where specialized proteins, known as transcription factors, can bind and activate or repress transcription, the process of converting genetic information into proteins.

D-loop control area and OL region are non-coding sections of the mitogenome. A displacement loop, also known as a D-loop, is a DNA structure in which a third strand of

DNA holds apart the two double-stranded DNA molecules for a stretch of time. According to Elgar *et al*. (2008), D-loop is essential for meiotic recombinational healing of such defects.

Transfer RNA is a small RNA molecule that plays a key function in protein synthesis. The major function of tRNA is to transfer amino acids to form the right sequence of the polypeptides.It also acts as a link (or adaptor) between the messenger RNA (mRNA) molecule and the growing chain of amino acids that make up a protein (Suzuki *et al*., 2005).

Ribosomal RNA (rRNA) plays a key role in protein synthesis by forming a bond with transfer RNA and messenger RNA to ensure that the mRNA's codon sequence is accurately translated into the amino acid sequence of proteins. The 16s rRNA serves as a scaffold to define the positions of the ribosomal proteins and has a structural function. The anti-Shine-Dalgarno sequence is located at the 3′ -end and binds to the mRNA upstream of the AUG start codon. The proteins S1 and S21, which are known to be important in the beginning of protein synthesis, bind to the 3′ -end of 16S RNA.In phylogenetic analyses of vertebrates, the 12S ribosomal RNA (rRNA) gene sequence has frequently been employed. Since rRNA is primarily involved in the production of proteins (Dahlberg, 1989).

Gene order and content are in the fish mitochondrial genome exhibit fast divergence. A variety of unusual mutational derivatives of a few common mtDNA haplotypes can be found in a significant number of fish species. When compared to marine species, freshwater species have seen greater genetic divergence among dominant haplotypes, particularly those living in unglaciated regions. Hatchery stock analysis has frequently revealed their mtDNA variability, but most of the time, stocks are set for haplotypes that are also prevalent in wild populations (Satoh *et al*., 2016).

Environmental and cellular oxidative stress can be caused by a variety of factors. According to Burton and Jauniaux (2011), oxidative stress happens when the creation of reactive oxygen species exceeds the body's natural antioxidant defenses.Environmental toxins, xenobiotics, temperature, and a variety of other factors can cause oxidative stress in mitochondria. Reactive oxygen species (ROS), which are created in the mitochondria as part of regular cellular metabolism to produce cellular energy, can harm the structure and function of the mitochondria if they build up. The mitochondrial genome is more than five times more sensitive to oxidative stress than the nucleus, making it particularly vulnerable to these challenges.There are many information available regarding how mitochondrial activity is

affected by exposure to environmental toxins. Environmental contaminant and hazardous metal cadmium (Cd) can lead to mitochondrial dysfunction, especially by lowering ATP synthesis and mitochondrial efficiency (Kurochkin*et al*., 2011). The common environmental contaminant dibenzofuran (DBF) can change mitochondrial permeability and even result in mitochondrial malfunction (Duarte *et al*., 2013).

There have been numerous studies on *Chitala chitala* that have examined its mitochondrial genome, genetic variation, and biological parameters (Singh *et al*., 2019; Banik *et al*., 2014; Sarker *et al*., 2008; Mandal *et al*., 2009; Castro *et al*., 2018; and Mandal *et al*., 2012), but no noteworthy research has been done on the species mitochondrial genome of *Chitala chitala* in Bangladesh. In this study, we used next-generation sequencing (NGS) using the Illumina MiSeq platform to describe the whole mitochondrial genome of *C. chitala*. The complete mitochondrial genome sequence will be helpful to develop particular molecular markers, such as single nucleotide polymorphisms for population genetic structure research. *Chitala chitala* management and conservation in Bangladesh might benefit greatly from study of the mitochondrial genome sequencing.

Hence, the research work was carried out to fulfill the following objectives:

- ✓ To sequence complete mitochondrial genome of *C. chitala*.
- ✓ To characterize the mitochondrial genome of *C. chitala*.
- ✓ To study mitogenome based on phylogenetic history of *C. chitala*.

CHAPTER- II

REVIEW OF LITERATURE

CHAPTER II

REVIEW OF LITERATURE

Several kinds of research work on mitochondrial genome sequencing of *Chitala chitala* were conducted by many researchers for different purposes in different countries. But no significant study was done on mitochondrial genome sequencing of *C. chitala* in the Bangladesh context. A brief and concise literature review that's relevant to the previous study is provided in the following section.

Rawal *et al*. (2020) researched on distinction of two featherback species (Osteoglossiformes: Notopteridae) in India based on scale structure. The fish species were identified based on their morphological structures and molecular markers was used for differentiating between different genera and species. This investigation showed the utility of scale in distinctive two featherback species *Notopterus notopterus* (Pallas) and *Chitala chitala* (Hamilton). Both species morphology and ultra structure of the scales were studied. The scales of *N. notopterus* were granulated with drop like structures where the focus is downy in *C. chitala*. Depends on length/width ratio, the scales were relatively longer in *N. notopterus* than *C. chitala*. The lateral line was comparatively short in *C. chitala* whereas it was longer in *N. notopterus* with extended anterior .

Islam *et al.* (2020) worked on the sequencing and annotation of the entire mitochondrial genome of the endangered labeonine fish *Cirrhinus reba* .*Cirrhinus reba* was gathered from the Khulna district of Bangladesh in order to sequence and characterize the entire mitochondrial genome. The circular mitochondrial genome measured 16,597 base pairs and contained 37 mitochondrial genes (13 protein-coding, 2 ribosomal, and 22 transfer RNA genes), along with two non-coding regions, an origin of light strand replication (OL) and a displacement loop (D-loop), which share a structural similarity with other Teleostei fish. The phylogenetic tree showed that labeonine fishes are closely related to one another. The *Cirrhinus reba* mitogenome as a whole demonstrated 99.96% distinctness from another *C. reba* haplotype.

The Indian featherback fish, *Chitala chitala* (Hamilton-Buchanan, 1822), is an endangered featherback species with a wide distribution in the Indian subcontinent, according to research by Singh *et al*. (2019) on its complete mitochondrial genome and phylogenetic status. The mitogenome of *C. chitala* was sequenced (16375 bp) and mapped to identify 13 protein-

coding genes (PCG), 22 transfer R genes, and a total of Indicated by the ratio of synonymous to non-synonymous substitutions (Ka/Ks), purifying selection was responsible for the evolution of 10 genes. Phylogenetic trees were built using the concatenation of 12 PCGs with the other seven orders and osteoglossiformes to determine the taxonomic relatedness of the organisms.

Alam *et al*. (2019) worked on *Amblypharyngodon mola* (Hamilton, 1842)'s mitochondrial genome and its evolutionarily related subfamily Danioninae. Using the MiSeq platform, the whole mitochondrial genome of *Amblypharyngodon mola* was identified. The circular mitogenome was 16,545 base pairs in length and contains a regulatory section (D-Loop), conventional 13 protein-coding genes, 22 tRNAs, and 2 rRNAs. The Heavy strand had 28 genes, and the Light strand contained the remaining 9 genes. ND2, ND3, ND4, ATP6, COX2, COX3, and Cytb are the seven genes with the incomplete stop codon (T-/ TA-). *A. mola's* mitogenome shared 82% of its sequence with *Rasbora vaterifloris* (GenBank No. NC015531), according to analysis.

Baeza (2018) carried out a study on the entire mitochondrial genome of Caribbean spiny lobster *Panulirus argus*. The *Panulirus argus* contained AT-rich mitochondrial genome which was 15739 bp in length including 22 transfer RNA genes, 2 ribosomal RNA genes, and 13 protein-coding genes (PCGs). On the H-strand, the PCGs were mostly discovered. All mitochondrial PCGs with computed KA/KS ratios and values below 1 were determined to be evolving under purifying selection. The Achelata and other infraorders within the Decapoda were predicted to be monophyletic via a maximum likelihood phylogenetic study.

Li *et al*. (2017) conducted study on the whole mitochondrial genome sequencing and phylogenetic implications of *Sinocyclocheilusn jii* (Cypriniformes: Cyprinidae). The circular mitochondrial genome was 16,577 bp long and had one regulatory region, two rRNAs, 22 tRNAs, and 13 protein-coding genes. The mitochondrial genomes of 11 species were used to create a phylogenetic tree, which revealed that nine Sinocyclocheilus species clustered together to form one monophyletic clade, with *Sinocyclocheilus jii* being the most primitive species.

Through pyrosequencing, Perini *et al*. (2016) investigated the entire mitochondrial genome of the southern purple-spotted gudgeon *Mogurnda adspersa* (Perciformes: Eleotridae). Using an eighth of a 454 pyrosequencing plate, the circular mitochondrial genome of *Mogurnda adspersa* was discovered for the first time. The complete mitogenome was put together using

the bioinformatics program MIRA. The structure of the *M. adspersa* genome, which had a length of 16,523 bp and had 13 protein-coding genes, 22 transfer RNA genes, 2 ribosomal RNA genes, and 1 non-coding regulatory area, was extremely similar to that of most vertebrates. For investigations on *Mogurnda adspersa's* systematics and conservation efforts, the entire mitogenome sequence may be useful.

Liu *et al.* (2016) established a study on the whole mitochondrial genome of the striped raphael catfish, *Platydoras armatulus* (Siluriformes: Doradidae), through next-generation sequencing research. *P. armatulu's* whole mitochondrial genome was 16,470 nucleotides in length. 13 protein-coding genes, 22 tRNA genes, 2 rRNA genes, and 1 regulatory area made up the mitogenome. The overall base composition was 30.9% for A, 25.2% for T, 15.9% for G, and 28.0% for C. The establishment of mitogenome-enriched catfish (*P. armatulus*) molecular resources crucial for phylogenetic analysis and species identification.

Carvalho *et al.* (2016) established a study on the entire mitochondrial genome of the endangered catfish *Lophiosilurus alexandri* (Silurifomes: Pseudopimelodidae). Phylogenomic study shows monophyly of Pimelodoidea. An endangered catfish from Brazil's Sao Francisco River Basin is called Lophiosilurus alexandri.Known as Pacama locally, this area offers economic potential for aquaculture production. The 16,445 bp-long sequenced mitochondrial genome exhibits the normal gene organization of mitochondria. From the complete mitogenomes of 20 Siluriformes and two outgroups, a phylogenomic study was developed. The results confirmed the monophyly of the superfamily Pimelodoidea by establishing the monophyly of nine catfish families and clustering *L. alexandri* as a sister group to the family Pimelodidae.

Zhao *et al.* (2015) established a study on complete mitochondrial DNA sequence of the threatened fish (*Bahaba taipingensis*): Mitogenome characterisation and phylogenetic implications .Long PCR and primer walking techniques were used to determine the Chinese bahaba's entire mitochondrial genome (mitogenome) sequence. The circular mitogenome of other bony fishes is 16500 bp long and consists of 37 mitochondrial genes, including a regulatory region, 13 protein-coding genes, 2 ribosomal RNA genes, and 22 transfer RNA genes. The control area contained the extended termination associated sequence (ETAS), central conserved sequence block (CSB-D, SCB-E and CSB-F), and conserved sequence block (CSB-1, CSB-2 and CSB-3) domains. *Bahaba taipingensis* was found to be more

closely linked to Pseudosciaeniae than Argyrosominae and Sciaeninae by phylogenetic analysis.

The full mitogenome of the sheat fish *Pterocryptis cochinchinensis* (Siluriformes:Siliridae) was determined by Xu *et al*. (2015), along with its phylogenetic implications. Polymerase chain reaction and the direct sequencing technique were used to establish the whole mitogenome sequence. There were 22 tRNA genes, 2 rRNA genes, 13 protein-coding genes, 1 regulatory area (D-loop), 16,501 bp of total mitogenome sequence, and 1 gene that was identical to a typical vertebrate gene. It was the first time the entire pterocryptis mitogenome sequence has been reported. The sheatfish was shown to have formed the most basic branch, having sister relationships with the clade encompassing all other investigated genus Silurus fishes, according to phylogenetic analysis based on cytb gene and mitogenome sequences.

Rangel-Medrano *et al.* (2015) established a work on *Pseudoplatystoma magdaleniatum* (Siluriformes, Pimelodidae), a Neotropical catfish, has a fully sequenced mitochondrial genome. The MiSeq Illumina platform was used to sequence the whole mitochondrial genome of *P. magdaleniatum*. The circular mitogenome's length was 16,568 bp, and it contained 22 transfer RNA genes, 2 ribosomal RNA genes, 13 protein-coding genes, and 44.19% GC. It also showed that the mitogenome of Pimelodus pictus was flawless and identical in length.

Mohindra *et al.* (2015) worked on complete mitogenome sequences of two endangered species, *Clarias batrachus* (magur) and *Pangasius pangasius* (family Claridae and Pangasiidae, respectively) to revealed the complete mitochondrial genome sequences .Both circular mitogenomes had 1 non-coding (control) region, 22 transfer RNAs, 2 ribosomal RNAs, 13 protein-coding genes, and were 16,511 and 16,476 bp in size. The gene order was the same as that of other animals that have been seen. In phylogenetic investigations, the genome resource of the full mitogenome sequencing of the Indian catfish species would be helpful.

In Channa marulius (Hamilton, 1822), a large snake head, Singh *et al*. (2015) analyzed the mitochondrial genome's entire sequence and characterisation. 16,569 base pairs made up the whole mitogenome of *Channa marulius*, according to its established nucleotide sequence. On the mitogenome, various genes were discovered to be arranged similarly to other teleosts. L-strand's base was made up of T (19.1%), C (31.5%), A (34.8%), and G (14.6%). The control region was 915 nt in length and free of any repeated regions.

Clarias fuscus (Teleostei, Siluriformes: Clariidae) has a full mitochondrial genome, which Zhou *et al*. (2015) established as a result. They also discussed how extensively dispersed it is in China, South Asia, and Africa. *C. fuscus's* whole mitochondrial genome sequence was discovered by PCR.The mitochondrial genome sequence of *C. fuscus* was 16,518 base pairs in length, with 13 protein-coding genes, 22 tRNA genes, 2 rRNA genes, and a non-coding regulatory area. Its gene composition and organization were similar to those of other vertebrates. Except for eight tRNA genes and the ND6 genes, the majority of the genes were encoded on heavy-strand. The bias of G and C had been discovered in many regions (genes), like other vertebrates.

Li *et al.* (2014) conducted a study on *Liobagrus merginatus's* entire mitochondrial genome (Teleostei, siluruformes: Amblycipitidae). It took 16,497 base pairs to decode the entire mitochondrial genome of *Liobagrus marginatus*, which contains 22 tRNA genes, 13 protein-coding genes, 2 rRNA genes, and a non-coding regulatory area. With the exception of eight tRNA genes and the ND6 genes, which were similar to those in other fish, the majority of the genes were encoded on the heavy-strand. The bias of G and C was discovered in statistical results of several genes/regions, just like other vertebrates.

Banik *et al*. (2014) worked on Effects of climate change on the occurrence of *Chitala chitala* (Hamilton-Buchanan, 1822) in Tripura. *C. chitala* (Hamilton-Buchanan, 1822) is a near threatened fish species of freshwater ecosystem. Occurrence and abundance of this species is greatly reducing in North-Eastern India so in view of conserving the species it is undoubtedly important to recognize its environmental quality. During last two decades the climatic characteristics of Tripura showed that a little by little increase in temperature,a slow and gradual rise in humidity and rainfall was decreased quantitatively. A statistical correlation was noticed between air temperature and rainfall ($r = 0.99$, $P < 0.001$) and same correlation was also found between humidity and rainfall ($r = 0.99$, $P < 0.001$). On the other hand, a direct relationship was marked between air temperature and humidity.The physico-chemical characteristics were identified in which water temperature were 12.11-32.19C, water velocity 1.66 - 4.02 m/sec, pH 6.61-7.31, DO_2= 4.44 -6.89 ppm, HCO_3= 112.14 -152.76 ppm etc.

Zhang *et al*. (2013) discovered a novel gene rearrangement with three noncoding sequence insertions and a tRNA gene order that was different from that seen in other vertebrates. The mitochondrial circular genome (18,523 bp) contained the same set of 37 mitochondrial genes,

including 2 ribosomal RNA (rRNA), 22 transfer RNA (tRNA), 13 protein-coding genes, and a control region.

Prosdocimi *et al*. (2012) established a study on the entire mitochondrial genomes of two recently diverged species of the fish genus Nannoperca (Perciformes, Percichthyidae) and found the sequences for both *Nannoperca obscura* and *Nannoperca australis*. For Percichthyidae (Perciformes), these sequences constitute the first whole mitochondrial genomes. The design of mitochondrial genomes was completed using the Sanger sequencing method. *N. obscura* and *N. australis* both have full mitogenomes that were 16,496 and 16,494 base pairs in length, respectively. 13 protein-coding genes, two ribosomal RNA genes, 22 transfer RNA genes, and a regulatory area make up both genomes.

Cheng *et al*. (2012) worked on the complete mitochondrial genome sequence of the bighead croaker *Collichthys niveatus* (Perciformes, Sciaenidae) . They reported that the mitogenome was 16,450 base pairs (bp) long, contained 13 protein-coding genes, 2 ribosomal RNA genes, 22 transfer RNA genes, and a non-coding control region. The gene organization, base composition, and tRNA structures of the *C. niveatus* mitogenome were similar to those of other bony fishes.

A study was investigated by Mandal *et al.* (2012) on the mitochondrial DNA variation in wild populations of the critically endangered Indian Feather-Back Fish, *Chitala chitala* .The genetic diversity in the D-loop and mitochondrial cytochrome-b (cyt b) regions confirmed the presence of the critically endangered *Chitala chitala*, a prehistoric feather-back fish. Eight riverine groups in India provided samples, which were then examined for the cyt b area (307 bp) and the D-loop region (636-716 bp). The mitochondrial region sequencing revealed minimal nucleotide range and a large haplotype variation. The distribution of mismatches, haplotypes networks indicating two distinct mitochondrial lineages, and patterns of genetic diversity all strongly support a historical influence on the genetic makeup of *C. chitala*.

Song *et al.* (2012) worked on full mitochondrial genome sequence of the dragonet *Callionymus curvicornis* (Perciformes: Callionymoidei: Callionymidae) 16,406 base pairs made up the *C. curvicornis* mitogenome, which also contained 1 regulatory area, 2 rRNA genes, 22 tRNA genes, and 13 protein-coding genes. Vertebrate mitochondrial gene structure was notable. It was the first time that the entire mitochondrial genome of a fish suborder, Callionymoidei, had been reported.

He *et al*. (2011) investigated the full mitochondrial DNA sequences for the genome characterisation and phylogenetic applications of the Nile tilapia (*Oreochromis niloticus*) and Blue tilapia (*Oreochromis aureus*). *O. niloticus* and *O. aureus* have their full mitochondrial genomes sequenced, and phylogenetic studies using mitochondrial protein-coding genes were done to find out their evolutionary relationship. *O. niloticus* circular mitogenome measured 16,625 bp while *O. aureus* circular mitogenome measured 16,628 bp. It contained 13 protein-coding genes, two rRNA genes, 22 tRNA genes, and a potential regulatory area. Three alternative computer approaches (maximum parsimony, maximum likelihood, and Bayesian method) were used to establish the phylogenetic link. *O. niloticus* and *O. mossambicus* had a tight phylogenetic link, whereas *O. aureus* had a distant relationship from the aforementioned two species.

Using allozymes, RAPD, and microsatellites, Mandal *et al.* (2009) generated a conclusion regarding the evaluation of genetic variation in the clown knifefish, *Chitala chitala*. The endangered Indian featherback *Chitala chitala's* natural population was examined to determine intraspecific variation using twenty-seven enzyme systems, six RAPD primers, and two microsatellite loci. *C. chitala* samples totaling about 262 were gathered from the Narmada, Ganga Mahanadi, and Satluj river basins. With an FST value for RAPD and a combined FST for microsatellite loci of 0.0344 (95% confidence 0.0340-0.0350), the analysis established population subdivisions.

The molecular identification and phylogenetic connections of seven Indian Sciaenids were studied by Lakra *et al*. (2009) using mitochondrial 16S rRNA and cytochrome oxidase subunit I genes. The phylogenetic relationships among the commercially significant Indian sciaenids (*Otolithoides biauritus, Protonibea diacanthus, Johnius dussumieri, J. elongatus,* and *Otolithes cuvieri*) were established using the partial sequences of the 16S rRNA and cytochrome oxidase subunit I (COI) mitochondrial genes. According to this research, the seven species could be divided into three distinct groups that had the same phylogenetic resolution and were genetically distinct from one another.

Jondeung *et al*. (2007) established a result on the whole mitochondrial DNA sequence of the Mekong giant catfish (*Pangasianodon gigas*) and the evolutionary relationships among Siluriformes. The *Pangasianodon gigas* is a critically endangered species and the biggest freshwater fish in the world without scales. In order to understand the relative phylogenetic position of *P. gigas* and to reconstruct the phylogenetic relationships among 15 of the 33

families of Siluriformes, phylogenetic analyses based on mitochondrial protein and rRNA (12S and 16S rRNA genes) data sets were conducted on the complete nucleotide sequence of the mitochondrial genome of the Mekong giant catfish.

In a study on the genetic divergence of two featherback fish species, *Chitala chitala* and *Notopterus notopterus*, Lal *et al*. (2006) found that markers that classify *C. chitala* and *Notopterus notopterus* may be found in the Allozyme and RAPD profiles. From 23 allozyme systems, 35 allozyme loci were found, and 16 of these loci established species differences. At 244-2902 bp, 15 RAPD markers were found with 77 size pieces. Between *C. chitala* and *N. notopterus*, the theta estimations of 0.9798 (allozymes) and 0.9471 (RAPD) revealed a significant genetic variance. The genetic heterogeneity clearly demonstrated the separation of the two species, *C. chitala* and *N. notopterus*.

CHAPTER- III

MATERIALS AND METHODS

CHAPTER III

MATERIALS AND METHODS

The experimental research work was conducted step-by-step which was presented by the following study design.

3.1 Study Design

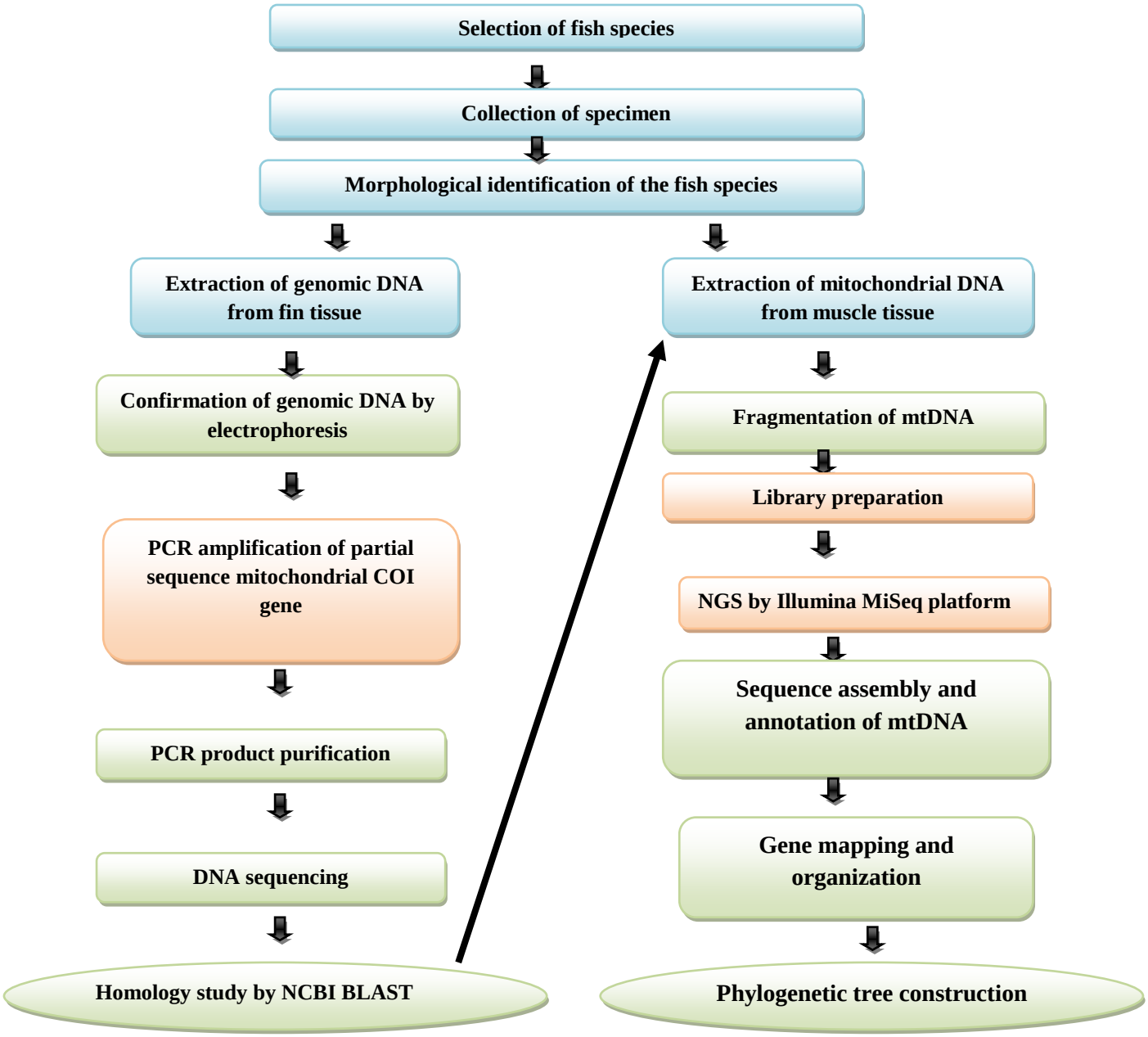


Figure 1. Layout of the study

3.2 Selection of Fish Species

Selection of fish species is an important task for research work. The specimens of *Chitala chitala* were collected from Kangsha river which is located in Netrokona district. The river is originated from the Garo Hills of India and flows as kangsha and later connected with the Shomeswari River at Jaria-Jhanjail.

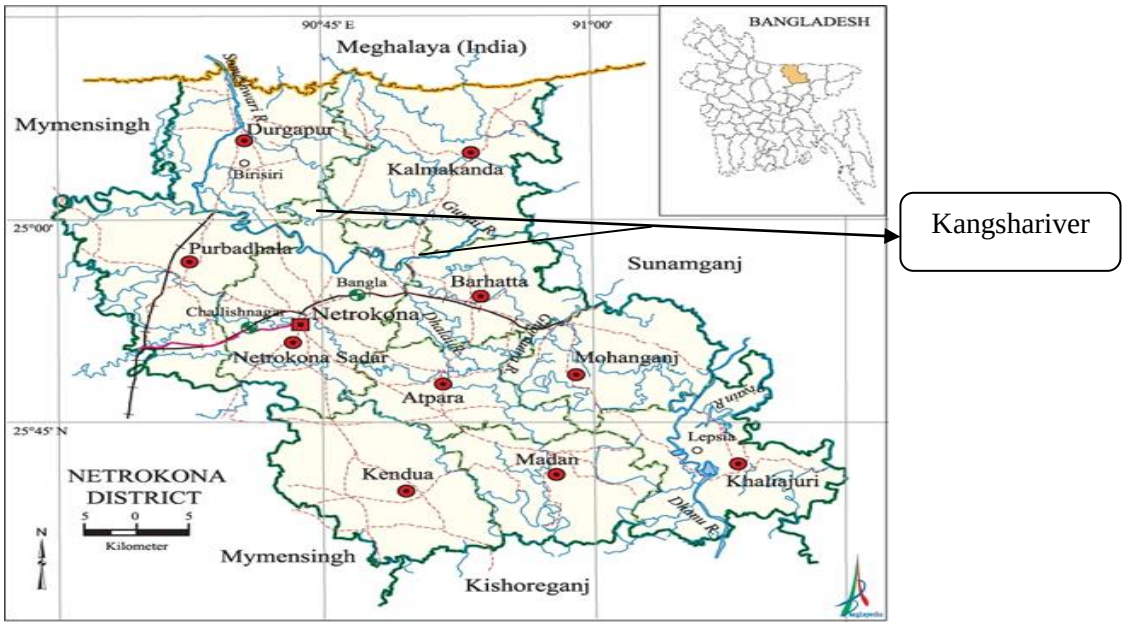


Figure 2. Location of Kangsha River in the Netrokona district (Source: Banglapedia)

3.3 Study Period and Location

The study was conducted during the period from February 2021 to June 2022.The study including morphological identification, genomic DNA extraction, quantification, PCR amplification, electrophoresis, purification, etc. was conducted at the laboratory of Fisheries Biotechnology Division, National Institute of Biotechnology (NIB), Savar, Dhaka-1349. Extraction of mitochondrial DNA, library preparation and next generation sequencing (NGS) were performed at abroad (South Korea). Sequence assembly and annotation were performed at the Department of Biotechnology, Sher-e-Bangla Agricultural University, Dhaka, Bangladesh.

3.4 Morphometric Taxonomy and General Information of Studied Fish Species

The morphological taxonomic study of the collected specimen was performed to ensure the correct initial identification of the fishes. Various parameters such as body shape, color, size, etc. were considered for morphological identification.

Body is elongated; head and body stoutly compressed laterally. The dorsal profile is rather concave. Short dorsal fin with very tiny scales. Long and confluent with the caudal fin is the anal fin. Pectoral fins are reduced. Dorsal part is coppery green colored and silvery at sides and below. 15 silvery bars present on each side of dorsal ridge. 5-9 small black spots near the end of the caudal fin. Lateral line is complete.

3.5 Materials Required for Tissue Sample Collection

As this study was laboratory based so well-equipped laboratory must be needed. To collect tissue sample following materials were required: Gloves, tissue, metallic tray, permanent marker, scissors, scalpels, forceps, 15- and 25-ml falcon tubes, 1.5 ml microcentrifuge tube, tube holders, distilled water, 70% ethanol, 95% ethanol, vortex mixer and weight balance etc.

3.6 Collection of Tissue Sample for DNA Extraction

After collection of samples, those were processed following aseptic techniques:

a) At first, all instruments were sterilized in an autoclave machine (AUTO CLAVE VS-1221, Vision Scientific Co., LTD) at 121°C for 20 minute under 15 psi and separate instruments were used for each sample.

b) After sterilizing hands with 70% ethanol and wearing gloves, fish was held and cut near about 1-2 g flesh tissue underneath the dorsal fin and pectoral fin with the help of a sterilized scalpel.

c) Then the sample was measured as 100-200 mg for mitochondrial DNA extraction and 400-500 mg fin tissues was collected by fin clipping for genomic DNA extraction using a balance machine (GIBERTINI, SER.NO. 153147, Italy), picked up with the help of sterilized forceps, and inserted into disinfected 1.5 ml labeled tube.

3.7 Isolation of Genomic DNA

A standard proteinase K incorporated with the phenol-chloroform-isoamyl alcohol method was used with slight alteration for genomic DNA extraction (Chowdhury *et al.*, 2016; Ahmed *et al*., 2019).

3.7.1 Reagent Required for Genomic DNA extraction

• **Lysis buffer :** A lysis buffer is a buffer solution utilize for the purpose of breaking open cells for use in molecular biology experiments that investigate the macromolecules of the cells.

Table 1. Composition of lysis buffer

Chemicals	Strength	Volume (1L)	Final concentration
Tris-HCL pH 8	1M	200 ml	0.2 M
EDTA pH 8.0	0.5 M	50	0.025 M
SDS	10%	50	0.5%
NaCl	5M	50	0.25 M
Ddh2o	-	650	-
	Total	=1000 ml	

• **Proteinase K:** In molecular biology, proteinase K is mostly used to break down protein and purify nucleic acid preparations of abnormalities.

• **Phenol/Chloroform/Isoamyl alcohol (PCI):** Phenol/Chloroform/Isoamyl alcohol (PCI) (25:24:1) is used in the refining of nucleic acids. This reagent consists of exceedingly pure chloroform, isoamyl alcohol and ultra Pure^{TM} Phenol saturated with Tris-HCL.

• **Chloroform/Isoamyl alcohol (CI) (24:1; v/v):** To eliminate proteins from preparations of nucleic acids, equilibrated phenol and chloroform are combined with isoamyl alcohol in a 24:1 ratio.

• **Ammonium acetate**: Ammonium acetate, also identified as spirit of Mindererus in aqueous solution which is frequently used with acetic acid to create a buffer solution.

3.7.2 Digestion of Tissue Sample

The collected fin tissue sample (100-200 mg) was cut into small pieces with a sterilized scissor and taken to aeppendorf tube containing 250 µl lysis buffer. The sample were grinded using pellet pestle grinder and added more 250 µl lysis buffer into that tube.Then 20 µl proteinase K (MP Biomedicals) were added. The tube was gently mixed by a vortex machine (Stuart, Bibby, CAT SA8, UK). After that, the tube was transferred to a floating stand and placed in the water bath (Figure 3) at 55°C for overnight until transparent the sample.

3.7.3 Isolation of DNA Pellets

After transparent the sample 520 µl PCI (the equal volume of Lysis buffer and proteinase k) was added into the eppendorf tube and carefully shook the tubes for a few minutes for proper mixing. Then the eppendorf tube were taken into centrifuge machine and centrifuged at10,000 rpm at 4°C temperature for 10 min. All the tubes were removed from the centrifuge machine and two phases were observed where the upper aqueous white phase contained DNA and the lower organic yellow phase contained lipids, proteins, and other impurities . Then 400 µl of sample from upper layer was carefully transferred to a fresh new tube without PCI contamination. After that, added 400 µl (equal amount of supernatant taken from the previous step) of Chloroform: Isoamyl alcohol(C.I=24:1) was added to each tube and mixed with them by manual inversion. Again, the tube was centrifuged at 10,000 rpm for 10 minutes at 4°C. Two layers were noticed (upper white DNA layer and lower white debris layer). Then 100 µl of sample from surface layer was taken with the help of a micropipette and placed into a new labeled eppendorf tube. Then, 900 µl of 100% frozen ethanol was added into each new tube containing supernatant and shake the tube. The tube was left at -20°C for 30 min. After that, the tubes were centrifuged at 14,000 rpm for 10 minutes. Following centrifugation, the DNA pellet was observed at the bottom of the tube. Then the tubes were made empty by removing all solution.

3.7.4 Purification of Extracted DNA

For the purification of DNA pellet, 1000 µl of 70% frozen ethanol was added to the effendorf tube. Then the tubes were centrifuged at 14000 rpm for 10 minutes and the surface liquid was poured from each tube. The eppendorf tubes were dried by air using PCR cabinet to remove the alcohol content and then the dried pellet was revoked in 50 µl of TE buffer (10mM Tris HCL and 1mM EDTA; pH 7.6) tapped for dissolving the pellet and centrifuged in a minicentrifuge (Dynamica, Velocity 6µ, Japan) at 4000 rpm for 2-3 minutes at room

temperature. Before adding the TE buffer was warmed up by keeping it in a water bath for a few minutes. Then the dissolved genomic DNA was preserved at -20°C for later use.

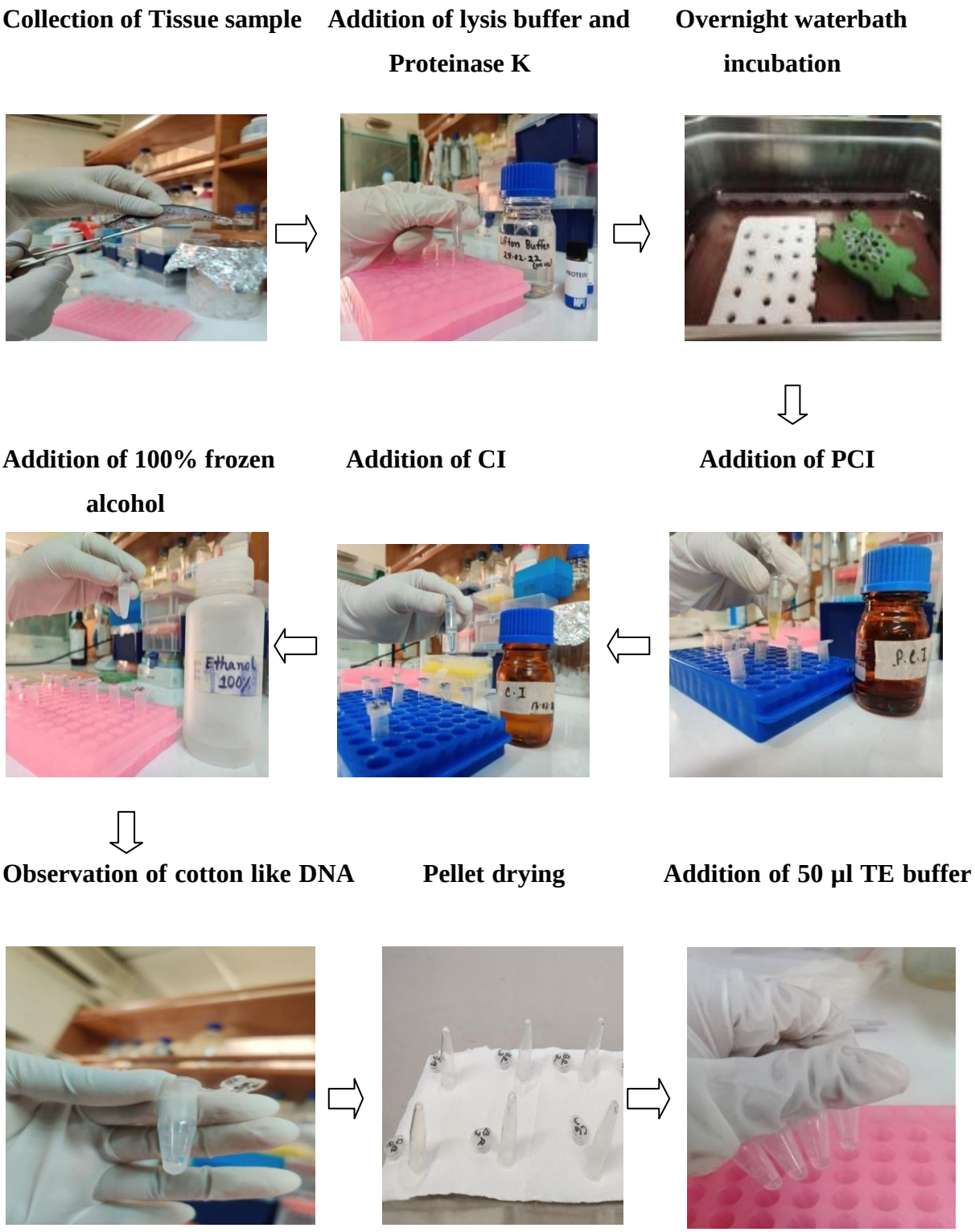


Figure 3. Pictorial views of genomic DNA extraction

3.7.5 Quantification of DNA

A nanodrop spectrophotometer (Nano Drop 2000 UV-Vis spectrophotometer, Thermo Fisher Scientific Inc., USA) was used for DNA quantification. It was done at 260/280 nm to determine the purity of DNA of fish samples. At first, the lid of the Spectrophotometer was raised and ultrapure water and lint-free laboratory tissue was used to wipe the upper and lower pedestal. Then, 2 µl of TE buffer was placed on the lower pedestal of the machine and closed the lid. After finishing the blank measurement, the lid was raised and again wiped with tissue and 2 µl of each DNA sample was placed. The concentration (ng/µl) and absorbance ratio (260nm/280 nm) of the extracted DNA were displayed on the monitor to show the DNA's concentration and purity.

a b

Figure 4. Quantification of DNA through Nanodrop Spectrophotometer (a) DNA sample loading, (b) Graphical result presented DNA concentration and purity.

3.7.6 Confirmation of genomic DNA through gel electrophoresis

By loading the isolated DNA onto a 1.5% agarose gel stained with ethidium bromide, the purity of the DNA was examined.

3.7.6.1 Materials Required for Gel Electrophoresis

- An electrophoresis chamber, power supply, electrophoresis buffer TAE 10X(Tris HCL,Glacial Acetic Acid ,EDTA),gel casting tray, and combs.
- Parafilm, Foil paper, Gloves, Measuring cylinder, Conical flask
- Microwave, UV transilluminator, Gel documentation system
- Agarose powder, Ethidium bromide
- DNA markers, Gel Loading Dye

3.7.6.2 Procedure of 1.5% Gel Preparation

0.3 gm agarose powder was measured

⇩

A conical flask was washed and measured 20 ml TBE buffer and poured into the flask

⇩

Added the agarose powder into the flask and shake to mix

⇩

The flask was placed into a microwave oven for 1.20 sec to dissolve the agarose powder

⇩

After dissolving, the gel was allowed to cool down

⇩

Then 2 µl of ethidium bromide (Sigma-Aldrich, Germany) was added to the flask

⇩

Mixed properly and poured into the tray

⇩

The agarose gel was allowed to be set at room temperature for 30 min

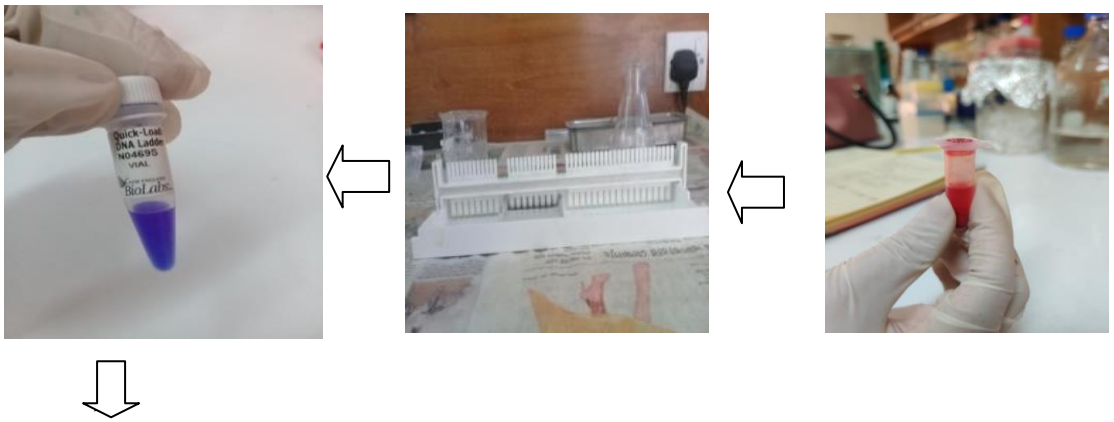
⇩

The comb was removed carefully and placed the gel in the electrophoresis tank.

Figure 5. Flow chart of 1.5% agarose gel preparation

Measure agarose powder **Measure TAE** **Flask in microwave oven**

DNA marker **Pour the gel** **Add Ethidium bromide**



Electrophoresis tank **Observation under UV light**

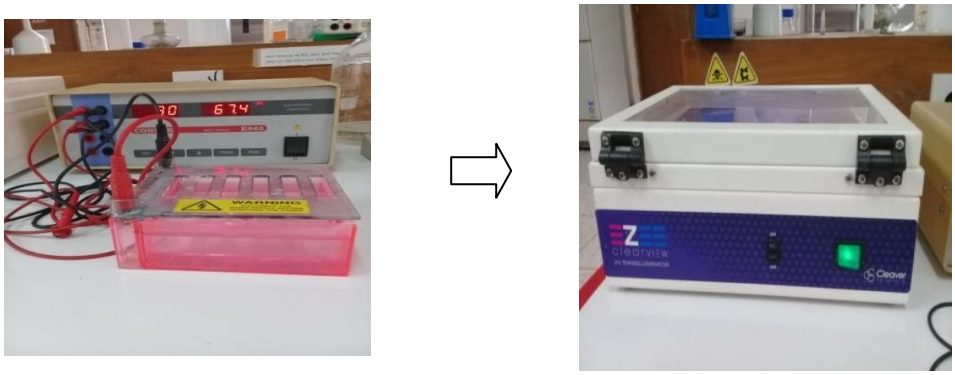


Figure 6. Pictorial views of 1.5% agarose gel preparation

3.7.7 DNA Sample Preparation and Electrophoresis

For preparing the DNA sample, gel loading dye (BioLabs 6X Blue, Gel Loading Dye) was mixed with the DNA. 5 µl DNA sample was mixed with 2 µl Gel Loading Dye on parafilm. Electrophoresis buffer (1X TBE) was added to the surface of the gel and carefully loaded the prepared DNA sample into the gel. DNA marker (QuickLoad® Purple 1 kb Plus DNA Ladder) was also loaded along with experimental DNA samples. Then, the lid was attached to the electrophoresis tank and turned-on power to run the gel for about 30 minutes. After running the gel, it was removed from the tray and exposed to UV light in UV transilluminator (3UV-UVP, LMS-20E, S/N-012006-001,USA) to confirm the DNA.

3.7.8 Polymerase Chain Reaction (PCR)

The PCR technique is one of the most well-known technique which allows detection and identification of gene sequences using visual techniques based on size and charge. It is based on enzymatic replication of DNA. For the species confirmation, the mitochondrial partial COI sequence was amplified using the BCL-BCH primers (Handy *et. al.,* 2011). The Cytochrome oxidase region of mitochondrial DNA is highly conserved which makes the design of universal primers and hence amplification possible. The steps of amplification of the COI gene were described in the following sector.

3.7.8.1 Materials Required for Polymerase Chain Reaction (PCR)

- DNA template
- Forward and reverse Primers
- Nuclease free water and Master mix
- Ice-cold PCR holder, PCR tube
- MiniSpin, PCR Cabinet, Thermal cycler

3.7.8.2 Primer Selection

In this study, Fish-BCL-F and Fish-BCL-R primers were used for fish species identification. These primers amplified at 655bp fragment of mitochondrial COI gene.

Table 2. Selected Primers for Mitochondrial COI Gene Amplification

BCL primers	**Primer sequence (5′-3′)**	**Amplicon size**
Fish-BCL-F	TCAACCAATCACAAAGATATCGGCAC	655 bp (Handy *et al*. 2011)
Fish-BCL-R	ACTTCCGGGTGACCGAAGAATCA	

3.7.8.3 Preparation of Polymerase Chain Reaction (PCR) Mixture

Following procedure was applied to prepare the PCR reaction mixture.

- At first, GoTaq® G2 Green Master Mix (DNA Polymerase, dNTPs, MgCl2, and reaction buffers), primer stock solutions, and nuclease-free water (Promega Madison, WI USA) were thawed from frozen stocks.
- Then, all components were placed on the icecold holder.
- The selected primers were diluted from stock and an exact amount of master mix, primers (forward and reverse), and nuclease-free water were mixed on a tube.
- Then, 24 µl of the mixture was taken in each labeled PCR tube and finally template DNA was added to each PCR tube. PCR tubes were tapped and taken into a Quick spin mini centrifuge (MiniSpin, Eppendorf AG, 22331, Germany) for 10-15 sec to proper mixing of all components.
- Whole procedure was performed inside the PCR safety cabinet (ESCO Airstream - PCR Cabinet) to avoid any type of contamination.

Table 3. Components of PCR Reaction mixture

Components	Volume(µl)	Total volume (µl) for six reaction
Master mix (2X)	12.5	12.5×6.5=81.25
Forward primer (100pmol/ µl)	1	6.5×1=6.5
Reverse primer (100pmol/ µl)	1	6.5×1=6.5
DNA template (50 ng/ µl)	2	6.5×2=13
Nuclease free water	8.5	6.5×8.5=55.25
Total reaction volume	25	162.5

3.7.9 Thermal profile

After proper mixing of all components, the tubes were placed in a thermal cycler (Applied Biosystems™ thermal cyclers: The ProFlex™ 3x32-Well PCR System. For initial denaturation, the reaction mixture was preheated for 4 minutes at 95°C. Then, 35 cycles comprising 45 sec denaturation at 94℃, 1 min annealing at 55℃ and 1min elongation at 72℃. Final step was done at 72℃for 10 min to complete extension of all of the amplified fragments.

Table 4 Thermal profile for the amplification of CO1 gene

Number of cycles	Step name	Temperature(℃)	Time
1	Pre heat	95	4 min
35	Denaturing	94	45sec
	Annealing	55	1 min
	Elongation	72	1min
1	Final extension	72	10min

Figure 7. PCR reaction programme at thermal cycler

3.7.10 Gel electrophoresis and documentation of amplified PCR products

Upon completion of PCR, it is important to know that the quantity of PCR product is enough for DNA sequencing (Ivanova *et al.*, 2006). Only the selected region will produce one single band during the electrophoresis if the PCR has amplified successfully (Brown, 2002). Once the sequences were amplified and they were analyzed using agarose gel and electrophoresis. The procedure of gel electrophoresis and documentation was described above. For PCR products, 1.5% agarose gel was prepared and Quick-Load® Purple 1kb bp DNA Ladder was used.As the product of PCR contained a green master mix that made the required DNA bands easier to observe on the UV transilluminator, gel loading dye was employed. Finally, a gel documentation system grabbed the image.

3.7.11 Purification of Amplified PCR Product

For sequencing, the next step of the procedure is hence to purify the amplified PCR product, i.e., clean the DNA from nucleotides that are not part of the sequence and residual primers. In this process, Monarch® PCR and DNA Cleanup Kit were used to purify the PCR product.

3.7.11.1 Materials Required for PCR Product Purification

• Kit components

✓ DNA cleanup binding buffer

✓ DNA wash buffer and elution buffer

✓ Spin Column and Collection tube

• Isopropanol, 95% ethanol

3.7.11.2 Steps of PCR Product Purification

14 ml isopropanol was added to Monarch DNA Cleanup Binding Buffer

20 ml ethanol was added to Monarch DNA Wash Buffer

Centrifugation process were completed at 13,000 rpm in a microcentrifuge

To bind the DNA, a starting sample volume of 50 µl is recommended

Nuclease-free water can be used for smaller samples to adjust the volume

Two sets of 25 µl volume PCR products were used for each sample and taken on 1.5 ml tube

100 µl DNA Cleanup Binding Buffer was added to the 50 µl sample.

300 µl ethanol was added and mixed well by pipetting

The column was inserted into the collection tube

Sample mixer was transferred onto the column and closed the cap

After 1 minute of spin, the flowthrough was discarded

Steps of PCR Product Purification (Continued)

⇩

To wash the DNA, the column was reinserted into the collection tube and 500 µl DNA Wash buffeer was added

⇩

After the 1-minute spin, the flow-through was discarded again

⇩

Washing step was repeated to remove enzymes

⇩

The column was transferred to a clean 1.5 ml microfuge tube and a re-spin was done for 1 minute to ensure the traces of salt and ethanol were not carried over to the next step

⇩

20 µl DNA Elution Buffer was added and after waiting for 1 minute at room temperature, centrifuged them for 1 minute to elute the DNA

⇩

A nanodrop spectrophotometer was used to determine the DNA concentration and purity of purified PCR product

⇩

The purified PCR product was stored at -20°C until sequencing.

Figure 8. Flowchart of PCR product purification

3.7.12 Sequencing of PCR Product

For sequencing, successfully generated PCR products were purified and taken away together with the forward and reverse primers used in the PCR to the DNA sequencing lab at NIB, Savar, Dhaka. The sequences were generated by using BigDye® Terminator v1.1 and 3.1 Cycle Sequencing Kits.

3.7.12.1 Materials Required for DNA Sequencing

- Kit components
  - ✓ Ready Reaction Premix (RRP)
  - ✓ Sequencing Buffer
- Purified PCR products
- Forward and reverse primes
- DNase/RNase-Free Distilled Water
- Hi-Di Formamide, EDTA, Sodium acetate
- 70% ethanol, absolute ethanol
- Sequencing machine

3.7.12.2 Procedure of DNA Sequencing

The quantity of PCR product is optimized to maximize the number of primer binding sites for the Big Dye reaction and is dependent upon the length and purity of the PCR product. The template volume was calculated from the value of Table 5 and the components were added as indicated in Table 6.

Table 5 . Template Volume Calculation for Cycle Sequencing

Template (PCR product)	Quantity
100-200 bp	1-3 ng
200-500 bp	3-10 ng
500-1000 bp	5-20 ng
1000-2000 bp	10-40 ng
>2000 bp	20-50 ng

Table 6. Components of Reaction Mixture for Cycle Sequencing

Reagents	Volume
BigDye Ready Reaction Premix (RRP)	4 µl
BigDye terminator buffer	2 µl
Template	Calculated the volume according to Table 3.5
Primer (Forward/reverse)	3.20 pmole
Ultrapure water	20 µl
Final volume	20 µl

3.7.12.3 Cycle Sequencing and Purification of Cycle sequencing Product

After reagent preparation and mixing, tubes were rotated to eliminate bubbles by using a rotator. Finally, the tubes were placed in a thermal cycler and performed the cycle sequencing. The process was taken up to two hours.To purify the cycle sequencing product, 1 µl 3M Sodium acetate, 1 µl EDTA, and 25 µl 100% ethanol were added into each PCR tube, mixed well and spun down. Then, incubated at -20°C temperature for 15 minutes and the tubes were centrifuged at 13,000 rpm for 15 minutes. The flow-through was discarded. Lastly, 35 µl of 70% ethanol was added, centrifuged at 13,000 rpm for 15 minutes and discarded the flow-through.

3.7.12.4 Capillary Electrophoresis

The samples were incubated at 37°C for 5 minutes inside the PCR machine. Then, 10µl highly de-ionized formamide/hidiformamide (Hi-Di) was added to each tube containing individual samples, vortex for 10-20 seconds and spun down. The samples were transferred to a thermocycler for denaturation of DNA (95°C for 5 min). Immediately after that, the samples were put into ice for making sure that the DNA stayed denatured. Then, the samples were loaded into 96 well plates of ABI 3500 Genetic Analyzer which uses 4 capillary

systems. For each sample, it took up to 4 hours to get a sequence. The raw data from the run result in Genetic Analyzer were analyzed using Sequencing Analysis Software (Version 5.2).

Figure 9. DNA Sequencer

3.7.12.5 Sequence Data Analysis

Sequences data were analyzed by using various bioinformatics software. The method for analyzing DNA data would be to describe the evolutionary relationship between species by performing phylogenetic analyses.

3.7.13 Confirmation of species by NCBI BLAST and barcoding

The DNA fragment was successfully sequenced and sequence identity was reviewed by searching GenBank using algorithms built into the Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi). BLAST helps estimate the similarity between the untrimmed barcode sequence records retrieved from sequencing and sequence records already existing in the comprehensive GenBank database. Chromas v. 2.6.6 was used for trimming and creating the fasta format of all the raw sequences and searching them on NCBI through BLAST to check the identity. The consensus size of the COI gene was ~655 bp (Handy *et al*., 2011).

3.8 Extraction of Mitochondrial DNA

Using a commercially available mitochondrial DNA isolation kit, mitochondrial DNA was isolated (Abcam, Cambridge, UK). The Mitochondrial DNA Isolation Kit from Abcam is a useful tool for isolating mtDNA in high yield and purity from a variety of cells and tissues without genomic DNA contamination. The pure mtDNA can be used for a variety of

purposes, including amplifications, cloning, Southern blotting, enzyme alterations, and PCR analysis.

3.8.1 Materials Required for Mitochondrial DNA Isolation

Kit components

Items	**Quantity**
5X Cytosol Extraction Buffer (CEB)	20 mL
Mitochondrial Lysis Buffer (MLB)	1.8 mL
Lyophilized Enzyme Mix (LEM)	1 vial
TE Buffer	1.5 mL

After opening the kit, all the buffers were stored at 4°C and during the experiment, they were kept on ice. 1X Cytosol Extraction Buffer was made by mixing 1 ml of the 5X Cytosol Extraction Buffer with 4 ml ddH_2O.Then Enzyme mix was suspended by adding 275 µl of TE buffer to lyophilized enzyme mix. They were aliquot, mixed well and re-freeze immediately at -80°C,as the enzyme mix would be stable for up to 3 months.

3.8.2 Additional Required Materials

- Absolute ethanol
- Tissue grinder
- Orbital shaker
- Pipettes and pipette tips
- Centrifuge and microcentrifuge

3.8.3 Procedure of Mitochondrial DNA Extraction

For the initial step, ice was collected from the ice maker and all the buffers working solutions and Phosphate Buffered Saline (Gibco, 1X PBS, pH 7.4) were placed in an ice box. Then, 1 ml PBS was added to the previously collected flesh tissue sample and homogenized using a mini hand homogenizer. The homogenate was transferred to a 15 ml falcon tube and 6 ml of ice-cold PBS was added as a washing solution. A centrifuge machine was used to centrifuge the tube containing sample at 600 g for 5 minutes at 4°C. The supernatant was carefully removed. Then the pellet was resuspended in 3 ml of 1X CEB in order to prevent the homogenized tissue from being too sticky to remove the insoluble materials during the low spin step. The tube was incubated on ice for 10 minutes.The tissue samples with CEB were again homogenized and the activity was carried out on ice.

In order to prevent damage to the mitochondrial membrane and the potential release of mitochondrial components, excessive homogenization was avoided. Homogenates were transferred to a 1.5 ml microcentrifuge tube and centrifuged at 1200 g for 10 minutes at 4°C. After this stage, the pellet was removed because it still had intact cells and nuclei. The supernatant was then put in a fresh 1.5 ml tube and centrifuged at 10,000 g for 30 minutes at 4 °C. The supernatant was removed after centrifugation and discarded. Again, the pellet was re-suspended in 1 ml of 1X CEB and centrifuged at 10,000 g for 30 minutes at 4 °C. Moreover, supernatants were removed, and the observed pellet was isolated mitochondria. 30 µl of the Mitochondrial Lysis Buffer (MLB) was added to lyse the mitochondria. Then, 10 µl enzyme mix was added and thoroughly blended. The pellet mixer was incubated in a 50°C water bath overnight. All proteins and DNAses were degraded by the Enzyme mix.

Then, 100 µl of absolute ethanol was added to each sample and the tube was maintained at -20°C for 10 minutes. Then, samples were centrifuged for five minutes at room temperature at high speed in a microcentrifuge. The supernatants were poured out from the tube.DNA from mitochondria was found in the pellet. Then, 1 ml 70% ethanol was added to each tube and centrifuged at 14,000 rpm for 10 min at 4°C. The supernatant was removed. Repeating the washing process with 70% ethanol. A pipette tip was used to remove the trace amount of ethanol from the tube. Since it could be challenging to dissolve the pellets if they were entirely dried, the pellet was dried for 5 minutes on a safety cabinet (Cleanair, CAH 1800, India). Finally, the mitochondrial DNA was resuspended in 20 µl TE buffer and stored extracted mitochondrial DNA at -20°C for future use.

3.8.4 Library Construction and NGS

The Covaris M220 Focused-Ultrasonicator (Covaris Inc., Woburn, MA, USA) initially fragmented the isolated mitochondrial DNA as 300–350 bp. The following steps were used to complete library preparation using the TruSeq RNA library preparation kit V2 from Illumina, San Diego, California, USA:

- End repair procedure
- 3'-end adenylation
- ligationof adaptor, and
- PCRassistedEnrichment of DNA fragments.

3.8.4.1 Performation of End Repair reaction

Using End Repair Mix, this process turned the overhangs left behind from fragmentation into blunt ends. 3' to 5' exonuclease activity eliminated the 3' overhangs, and polymerase activity filled in the 5' extensions.

3.8.4.2 Adenylation at 3' Ends

A single 'A' nucleotide was inserted to the 3' ends of the blunt fragments during the adapter ligation reaction to prevent ligation from one another. A similar single 'T' nucleotide on the adapter's 3' end provided a complementary overhang for ligating the adapter to the fragment. By using this method, chimera (concatenated template) generation was kept to a limit.

3.8.4.3 Adapter ligation

This step ligates indexing adapters to the ends of the ds cDNA to get them ready for hybridization onto a flow cell.

3.8.4.4 Enrichment of DNA Fragments

PCR was used to amplify the amount of DNA in the library as well as enrich those DNA fragments with adaptor molecules on both ends. The PCR was carried out using a PCR Primer Cocktail that annealed to the adaptor ends. To avoid affecting the representation of the library, the number of PCR cycles was kept to a minimum.

3.8.5 Purification and quantification of DNA templates

To obtain higher quality DNA, the enriched DNA library templates were purified using a DNA purification kit (RBC Bioscience, Jerusalem, Israel) and quantified using a qubit fluorometer. The 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) was used to validate the size, quality, and purity of the generated DNA libraries.

3.8.6 Illumina Sequencing

The Illumina MiSeq platform 2×300 bp pair ends (Illumina) were used for the NGS process.Illumina sequencing technique is a commonly used next-generation sequencing (NGS) technology throughout the world. Massively parallel sequencing is supported by Illumina sequencing devices and reagents, which use a proprietary approach that detects single bases as they are integrated into expanding DNA strands.For speedy and precise large-scale sequencing, this approach employs clonal array creation and patented reversible terminator technology. The novel and adaptable sequencing system allows for a wide range of applications in genomics, transcriptomics, and epigenomics.

3.8.7 Sequence assembly and annotation of the mitochondrial genome

MiSeq raw readings were assembled using the Geneious Prime 2020.0.3 software to create a comprehensive mitogenome of *Chitala chitala*. The ORF finder program (https://www.ncbi.nlm.nih.gov/orffinder/) was used to organize the gene order, sequences, and sizes of each of the 13 protein-coding genes and two ribosomal RNA genes based on the *Chitala chitala* reference mitogenome (GenBank no. ON764424). ARWEN software was used to identify all transfer RNA genes and their anticodon locations.

3.8.8 Gene mapping

Gene mapping is usually the first step of the detection of the gene. Genes can be viewed as one particular type of genetic markers in the creation of genome maps. Map of gene is used to identify the locus of a gene and the distance between genes. Gene mapping can also explain the distances between different sites within a gene. The OGDRAW software (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html) was used to create the *C. chitala* circular gene map.

3.8.9 Construction of predicted tRNA structure

The ARWEN software, as implemented in the MITOS web server, was used to identify tRNA genes, and the secondary structure of each tRNA was determined using the tRNAscan-SE v.2.0 web server (http://trna.ucsc.edu/tRNAscan-SE/).22. The RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) was used to visualize tRNA secondary structures.

3.9 Phylogenetic tree construction

A phylogenetic tree was constructed for the confirmation of the sequenced data using the nucleotide sequences of the studied species with seven reference sequences from the GenBank database (Accession no. of NCBI GeneBank database reference sequences was given in APPENDIX I). The alignment was performed by MEGA 11(Kumar *et al*., 2018). Using the Maximum Likelihood method and Tamura-Nei model (Tamura and Nei,1993) the evolutionary record was inferred. The evolutionary history of the taxa analyzed is represented by a bootstrap consensus tree calculated from 1000 replicates. Branches that are identical to partitions that are replicated in less than 50% of bootstrap replicates are collapsed. Next to the branches was the percentage of trees in which the connected taxa clustered together in the bootstrap test 1000 repetitions. The initial tree(s) for the heuristic search were generated automatically by applying the Neighbor-Join and BioNJ algorithms on a matrix of pairwise distances computed using the Tamura-Nei model and then choosing the topology with the highest log likelihood value. This analysis involved 8 nucleotide sequences. Codon positions integrated were 1st+2nd+3rd+Noncoding. There was a total of 17003 positions in the ending dataset. Evolutionary analyses were conducted in MEGA11(Kumar *et al*., 2018).

CHAPTER-IV

RESULTS AND DISCUSSION

CHAPTER IV

RESULTS AND DISCUSSION

Genomic DNA was isolated using the proteinase K/phenol-chloroform-isoamyl alcohol method to validate the species. Step-by-step procedures were used to ensure the extracted DNA's purity, integrity, PCR amplification, and gel electrophoresis etc. The PCR amplified mitochondrial COI gene was used to confirm the species.

4.1 Fish species identification based on COI gene sequence

In figure 10, it was found that all amplified PCR products of COI gene amplification showed bands at ~655 bp. Then the purified PCR products were sequenced and confirmation of *Chitala chitala* was performed based on PCR confirmation mt COI gene (Appendix III). The partial COI gene sequence homology was searched using NCBI BLAST (Appendix III). The query sequence was found to be 100% identical to the *Chitala chitala* COI gene sequence (GenBank Accession No. MK572123.1). Previously, partial COI gene was amplified sequenced and matched by NCBI BLAST for different fish species, *Cirrhinus reba* (Islam *et al*., 2020), *Amblypharyngodon mola* (Alam *et al*., 2019) and *Pangasius pangasius* (Mohindra *et al*., 2015).

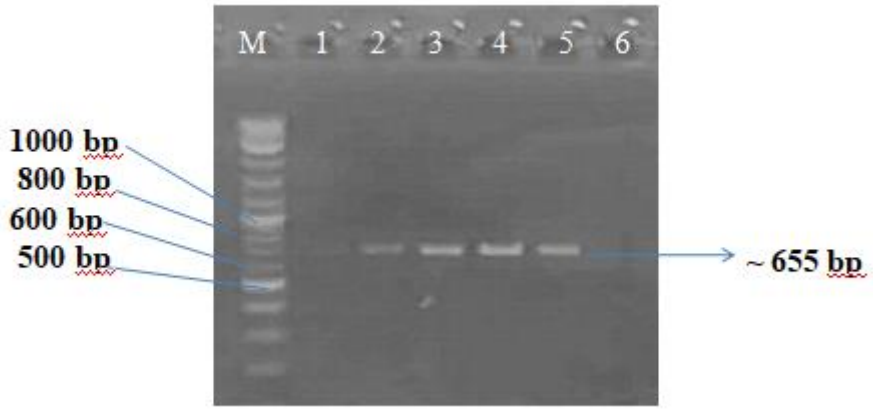


Figure 10. Gel electropherogram showing amplification of COI gene from genomic DNA of *Chitala chitala* (Lane M was 1kb plus DNA marker ; lanes 1-6, PCR products from six specimens)

4.2 Whole mitochondrial genome sequencing of Chital fish

The mitochondrial DNA was extracted from muscle tissue samples. The entire method, including mitochondrial DNA extraction, library preparation, and next generation sequencing (NGS), was carried out in South Korea. The feature of the complete mitogenome, gene organization, base composition, description of protein coding genes,non coding genes, tRNA structure etc. are described below.

4.2.1 Mitochondrial genome structure

C. chitala's circular complete mitochondrial genome (Accesion no. ON764424) was 16248 bp long and included 13 typical protein coding genes, 22 tRNA genes, two ribosomal RNAs (12SrRNA and 16SrRNA), and two non-coding regions (control region, D-loop, and origin of light strand, OL) (Table 7, 8, 9). The heavy (H) strand encoded 28 genes, while the light (L) strand encoded the rest of the genes (Figure 11).

A total of 31 bp of overlapping region have been identified across the *C. chitala* mitogenome in 13 distinct places. The mitogenome contained six intergenic spacers of 24 bp in length. The longest spacer was an 8-bp nucleotide sequence located between the tRNALeu and ND1 genes. In table 7, positive (+) values show intergenic space/gap between genes, negative (-) numbers indicate overlapping between genes, and zero (0) numbers indicate either overlap or no space exists between genes in the *C. chitala* mitogenome. *C. chitala*'s mitogenome exhibited striking similarities to that of other vertebrates (Mohindra *et al*., 2015; Zhou *et al.*, 2015; Singh *et al*., 2019; Islam *et al*., 2020).

4.2.2 Protein Coding Genes

The 13 protein-coding genes (PCGs) made up 70.30% of the mitogenome and had a total length of 11,423 bp. The shortest PCG was ATP8 (168 bp), while the largest was ND5 (1,838 bp). Twelve PCGs started with the identical translation initiation codon, ATG (Methionine), with the exception of the COX1 gene, which had an unexpected and alternative start codon, GTG (Valine). The remaining seven PCGs (ND2, COX1, COX2, ND3, ND4, ND6 and CYTB) finished with the standard stop codon (T--), as in other vertebrates, but the open reading frame of ND5 used TA- as a stop codon. Incomplete termination codons (TAA) occurred at the ends of five PCGs: ND1, ATP8, ATP6, COX3, and ND4L (Table 7). It is expected that posttranscriptional polyadenylation, or poly-A tail, will be used to finish these incomplete stop codons.

Table 7. **Organizational features of 13 Protein Coding Genes**

Gene	Nucleotide position		Size(bp)	Codon		Intergenic nucleotide (bp)	Strand	A+T(%)
	From	To		Start	Stop			
ND1	2883	3851	969	ATG	TAA	3	H	53.4
ND2	4064	5110	1047	ATG	T--	-2	H	55.6
COX1	5493	7041	1549	GTG	T--	-2	H	57.2
COX2	7191	7881	691	ATG	T--	0	H	57.6
ATP8	7957	8124	168	ATG	TAA	-10	H	61.3
ATP6	8115	8798	684	ATG	TAA	-1	H	58.6
COX3	8798	9583	786	ATG	TAA	-1	H	53.6
ND3	9655	10003	349	ATG	T--	0	H	54.7
ND4L	10075	10371	297	ATG	TAA	-7	H	53.8
ND4	10365	11745	1381	ATG	T--	0	H	57.7
ND5	11956	13793	1838	ATG	TA-	0	H	58.2
ND6	13794	14316	523	ATG	T--	0	L	57.2
CYTB	14390	15530	1141	ATG	T--	0	H	57.0

4.2.3. Transfer RNA genes

The 22 tRNA genes encoded in the mitochondrial genome of *Chitala chitala* ranged in length from 67-76 bp, estimating a total length of 1,570 bp (9.6% of the total mitogenome) (Table 8) and all tRNA exhibited a standard 'cloverleaf' secondary structure as predicted by both ARWEN and tRNAs can-SE v.2.0 (Figure 13). Fourteen tRNA genes were transcribed on the H-strand, however the remaining eight tRNA genes were transcribed on the L-strand (Fig. 11).

Table 8. Organizational features of 22 tRNA genes

Gene	Nucleotide position		Size (bp)	Anti-codon	Intergenic nucleotide (bp)	Strand	A+T (%)
	From	To					
$tRNA^{Phe}$	1	69	69	GAA	0	H	54.6
$tRNA^{Val}$	1026	1096	71	TAC	0	H	45
$tRNA^{Leu}$	2799	2874	76	TAA	8	H	50.0
$tRNa^{Ile}$	3855	3925	71	GAT	-1	H	56.3
$tRNA^{Gln}$	3925	3995	71	TTG	-2	L	56.3
$tRNA^{Met}$	3994	4064	71	CAT	-1	H	53.5
$tRNA^{Trp}$	5109	5177	69	TCA	-1	H	59.4
$tRNA^{Ala}$	5177	5248	72	TGC	0	L	62.5
$tRNA^{Asn}$	5249	5321	73	GTT	0	L	60.2
$tRNA^{Cys}$	5354	5421	68	GCA	-1	L	48.5
$tRNA^{Tyr}$	5421	5491	71	GTA	1	L	61.9
$tRNA^{Ser}$	7040	7114	75	TGA	4	L	50.6
$tRNA^{Asp}$	7119	7190	72	GTC	0	H	59.7
$tRNA^{Lys}$	7882	7956	75	TTT	0	H	52.0
$tRNA^{Gly}$	9583	9654	72	TCC	0	H	63.8
$tRNA^{Arg}$	10004	10074	71	TCG	0	H	59.1
$tRNA^{His}$	11746	11815	70	GTG	0	H	74.2
$tRNA^{Ser}$	11816	11882	67	GCT	-1	H	49.2
$tRNA^{Leu}$	11882	11956	75	TAG	-1	H	52.0
$tRNA^{Glu}$	14317	14384	68	TTC	5	L	57.3
$tRNA^{Thr}$	15531	15603	73	TGT	3	H	54.7
$tRNA^{Pro}$	15607	15676	70	TGG	0	L	62.8

4.2.4 Ribosomal RNA genes and non coding regions

As in other bony fishes (Song *et al.*, 2012; Xu *et al*., 2016; Li *et al.*, 2017; Alam *et al*., 2019; Perini *et al*., 2016; Prosdocimi *et al*., 2012), the mitogenome of *C. chitala* contained small subunits of rRNA (12S rRNA) and large subunits of rRNA (16S rRNA). The two ribosomal RNA genes collectively comprised 16.35% (2,658 bp) of the circular mitogenome and were both found on the H-strand. As in other vertebrate genomes, 12S rRNA gene was located between the $tRNA^{Phe}$ and $tRNA^{Val}$ genes and 16S r RNA gene was located between $tRNA^{Val}$ and $tRNA^{Leu}$ genes, respectively.

C. chitala, like other vertebrates, had no introns and two non-coding regions, an OL and a control region or displacement loop (D-loop). *C. chitala's* primary non-coding area (control region) consisted of 572 nucleotides, accounting for 3.5% of the total mitogenome, and the control region was dominated by A+T content (68.7%). The AT rich regulatory region

comprises promoters and an origin of replication of mtDNA, both of which are required for transcription and replication of mtDNA.The control region is particularly flexible to size variations. The 32-nucleotide OL region was positioned between $tRNA^{Asn}$ and $tRNA^{Cys}$ and was orientated on the L-strand in a set of five tRNA genes.By comparing with the recognition sites in some reported fishes (Rawal *et al*., 2020; Li *et al*., 2017; Castro *et al*., 2018; Jondeung *et al*., 2007 and Perini *et al.,* 2016) found the almost similar result.

Table 9. Organizational features of two rRNA and two non coding region

Gene	Nucleotide position		Size(bp)	Intergenic nucleotide (bp)	Strand	A+T(%)
	From	To				
12S rRNA	70	1025	956	0	H	54.5
16S rRNA	1097	2798	1702	0	H	55.1
OL	5322	5353	32	0	-	25.0
Control Region (D-loop)	15677	16248	572	0	-	68.7

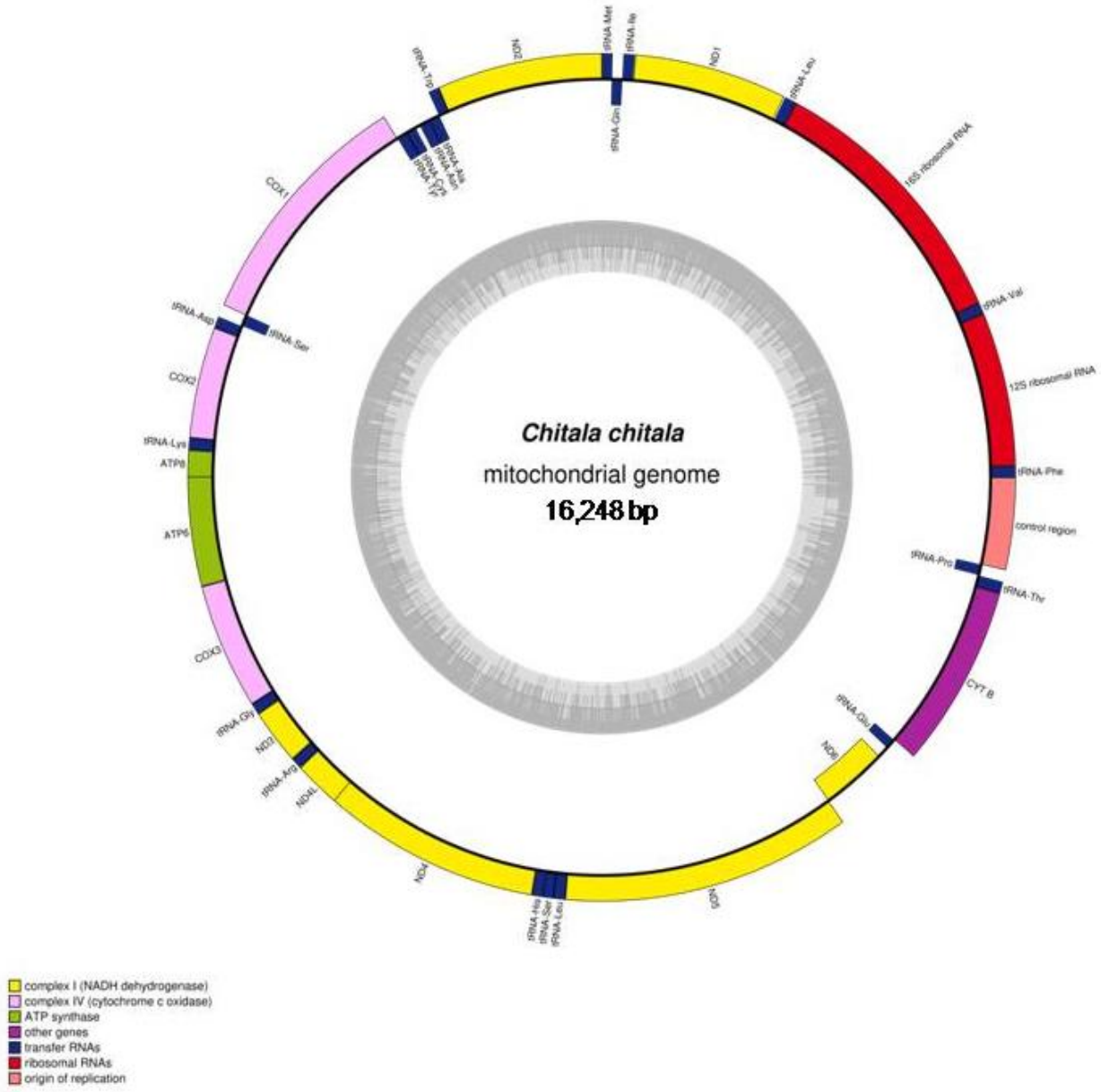


Figure 11. Gene organization of mitochondrial genome of the *Chitala chitala*

4.2.5 Overall base composition of *Chitala chitala*

The mitogenome of *C. chitala* had an overall base composed of 56.74% for A + T contents (A = 32.00% and T = 24.74%) and 43.26% for G + C contents (G = 15.31% and C = 27.95%) (Figure 12), respectively indicating an obvious anti-guanine bias. Similar works was established by Singh *et al*. (2019) on *Chitala chitala* in India.

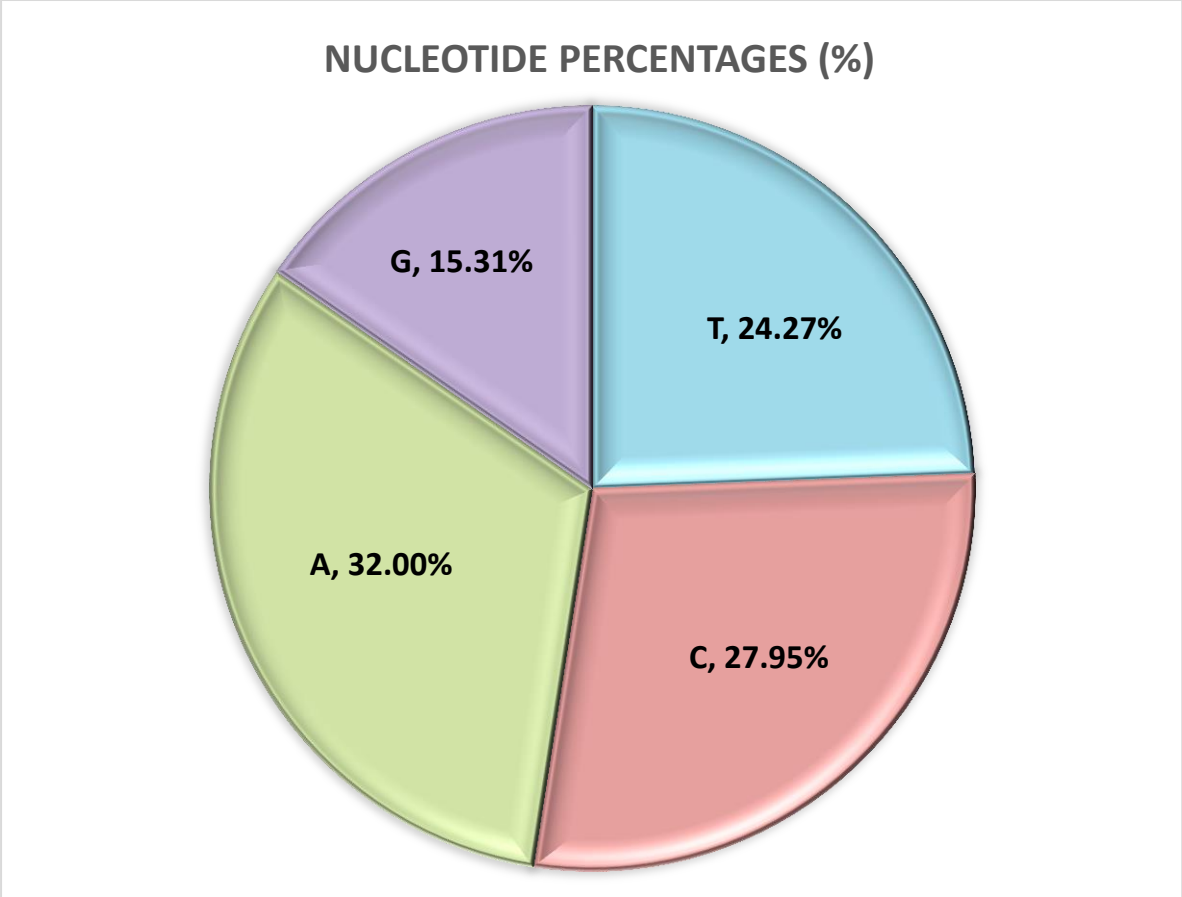


Figure 12. Base composition of *Chitala chitala*

4.2.6 Base composition of Protein-coding genes

In table 10, T had a base composition of 25.16%, C of 29.84%, A of 31.50%, and G of 13.50%. With the exception of the ND6 gene, which is on the light (L) strand, the concatenated sequence of all protein-coding genes on the heavy (H) strand was used to compute the nucleotide composition and codon use frequencies. The sequences of 13 PCGs were analyzed, and the results revealed a 57% total A+T content. In all the positions of codons (1^{st},2^{nd} and 3^{rd}) estimated the A plus T contents as 49.9%, 58.3%, and 61.6% respectively (Table 10). In this table, we have observed that T (37.04%) and A (31.50%) contents predominate over C and G contents in the second and third positions of the codons. Guanine content (4.76%) was visible at the third position in every codon. Guanine content (4.76%) was seen at the third position each codon, which is identical to *Cirrhinus reba* (Islam *et al*., 2020).

Table 10. Base composition of PCGs of *Chitala chitala* mitochondrial genome

PCGs	T%	C%	A%	G%	A+T%	Total(bp)
ND1	23.7	32.0	29.7	14.5	53.4	969
ND2	20.9	33.2	34.7	11.08	55.6	1047
COX1	28.9	25.8	28.2	16.8	57.2	1549
COX2	26.0	26.6	31.5	15.7	57.6	691
ATP8	25.6	30.9	35.7	7.74	61.31	168
ATP6	28.0	29.5	30.5	11.8	58.6	684
COX3	25.7	29.9	27.9	16.4	53.6	785
ND3	27.2	30.9	27.5	14.3	54.7	349
ND4L	28.2	30.6	25.5	15.4	53.8	297
ND4	26.7	28.6	30.9	13.6	57.7	1381
ND5	25.1	30.3	33.1	11.3	58.2	1838
ND6	13.7	29.8	43.4	12.8	57.2	523
CYTB	26.8	29.2	30.2	13.6	57.0	1141
Position in the codon						
1^{st}	21.52	27.62	28.43	22.43	49.95	3811
2^{nd}	37.04	28.31	21.35	13.30	58.39	3806
3^{rd}	16.93	33.60	44.71	4.76	61.64	3805
Across all PCGs	25.16	29.84	31.50	13.50	56.66	11422

4.2.7 Base composition of tRNAs, ribosomal RNAs and non-coding regions

In table 11, The highest content of A and T was observed in $tRNA^{His}$(74.3%), whereas that of the lowest content was observed in $tRNA^{Val}$ (45%). In ribosomal RNAs, the A+T content (54.84%) was higher than the G+C content (45.16%), whereas the overall base composition of the 12S rRNA gene was A = 32.74%, T = 21.75%, C = 25%, and G = 20.50%, and the 16S rRNA gene was A = 36.13%, T = 19.03%, C = 24.61%, and G = 20.21%. The overall base composition of non-coding region OL was A = 31.6%, T = 32.4%, C = 20%, and G = 17% and the non-coding region was A = 34.4%, T = 34.2%, C = 17.1%, and G = 14.1%.

Table 11. Base composition of tRNAs, ribosomal RNAs and non-coding regions

Genes	T%	C%	A%	G%	A+T%
$tRNA^{Phe}$	17.1	26.5	37.5	18.7	54.7
$tRNA^{Val}$	14.08	25	30.1	23.9	45
$tRNA^{Leu}$	25	30.2	25	19.7	50
$tRNa^{Ile}$	28.1	18.3	28.2	25.3	56.3
$tRNA^{Gln}$	22.5	28.1	33.9	15.5	56.3
$tRNA^{Met}$	23.9	28.1	29.5	18.3	53.5
$tRNA^{Trp}$	23.1	21.7	26.2	18.8	59.4
$tRNA^{Ala}$	27.7	23.6	34.7	13.9	62.5
$tRNA^{Asn}$	24.6	24.6	35.6	15	60.3
$tRNA^{Cys}$	26.4	25	22	26.4	48.5
$tRNA^{Tyr}$	26.7	25.3	35.2	12.7	62
$tRNA^{Ser}$	25.3	29.4	25.3	20	50.7
$tRNA^{Asp}$	26.3	20.8	33.3	19.4	59.8
$tRNA^{Lys}$	20	26.7	32	21.3	52
$tRNA^{Gly}$	29.1	20.8	34.7	15.2	63.9
$tRNA^{Arg}$	26.7	23.9	32.4	17	59.1
$tRNA^{His}$	37.1	10	37.1	15.7	74.3
$tRNA^{Ser}$	20.8	28.3	28.3	22.3	49.2
$tRNA^{Leu}$	21.3	26.7	30.7	21.3	52
$tRNA^{Glu}$	25	26.4	32.3	16.2	57.3
$tRNA^{Thr}$	24.6	24.6	30.1	20.5	54.8
$tRNA^{Pro}$	28.5	24.2	34.2	12.9	62.9
12S rRNA	21.7	25	32.7	20.5	54.4
16S rRNA	19	24.6	36.1	20.2	55.1
OL	32.4	20	31.6	17	58.5
Control region	34.2	17.1	34.4	14.1	68.7

4.2.8. Predicted Secondary structure of all tRNAs

All tRNA exhibited a standard 'cloverleaf' secondary structure as predicted by both ARWEN and tRNAs can-SE v.2.0 (Figure 13).Similar type of cloverleaf secondary structure of tRNA genes for whole mitogenome was established by Carvalho *et al*., 2016 and Satoh *et al*., 2016.

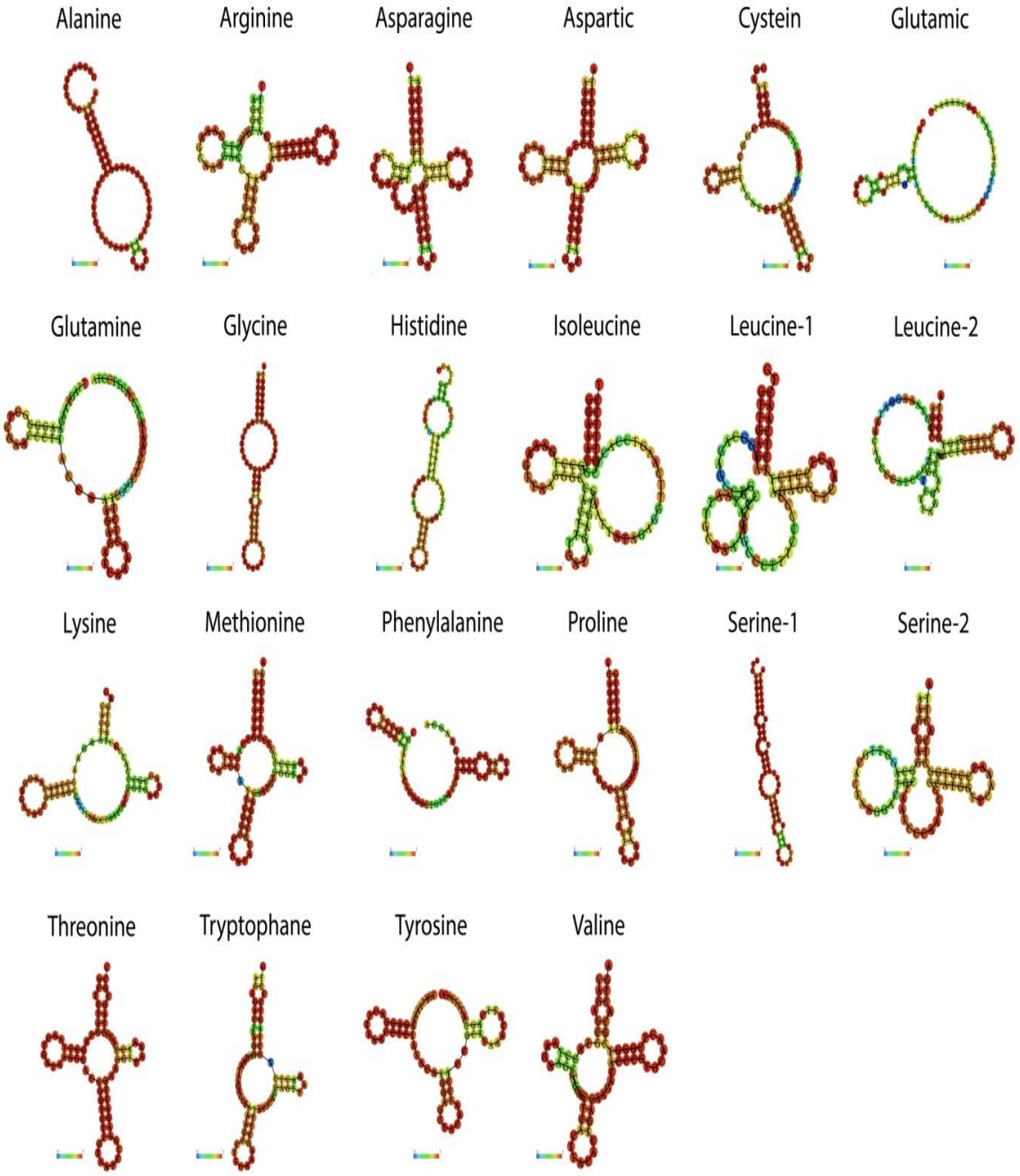


Figure 13. Predicted Secondary structure of all tRNAs in the mitochondrial genome of *Chitala chitala* predicted by ARWEN and tRNAscan-SE v.2.0.

4.3. Sequence homology and Phylogenetic tree construction

The sequenced data was exported as FASTA sequence format for comparing with other reference sequences available on the NCBI GenBank database using the Basic Local Alignment Search Tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The morphologically identified fish species showed about 97% or higher identity with the nearest related species except for *Chitala lopis*.

Phylogenetic tree construction and evolutionary analysis of the *Chitala chitala* with other notopteridae fishes was carried out by using the MEGA 11 software. Kyle and Wilson (2007) used phylogenetic analysis and a BLAST search tool to compare sequences between species providing accurate statistical metrics for species identification. A phylogenetic tree was constructed for the validation of the sequenced data using the nucleotide sequences of the studied species with reference sequences from NCBI GenBank databases. The Tamura-Nei model and the Maximum Likelihood approach were used to infer the evolutionary history. The percentage of trees with the associated taxa clustered together appeared adjacent to the branches. In the phylogenetic tree, the studied species *Chitala chitala* (GenBank no. ON764424) showed 100% similarity with *Chitala blanci*(GenBank No.AP008921) and 99% similarity with *Chitala ornata*(GenBank no.AP008923). The Nearest-Neighbor-Interchange method was used to assess tree inference. The Tamura-Nei distance approach (Tamura and Nei, 1993) was used to determine the heuristic and inequality of nucleotide frequencies. Ghouri et al. (2020) amplified the COI gene through Polymerase Chain Reaction (PCR). The studied species clustered with the same species from the NCBI GeneBank database that were previously submitted (Appendix IV). After validation,the complete mitochondrial genome was submitted to NCBI (National Center for Biotechnology Information).Recently, the authority of NCBI completed verification of whole mitochondrial genome of *Chitala chitala* and released under the GenBank Accession no. ON764424 (Appendix V), since there is no verified complete mitogenome of *Chitala chitala*. Currently, it has been considering as a provisional reference sequence (NC_070068.1).

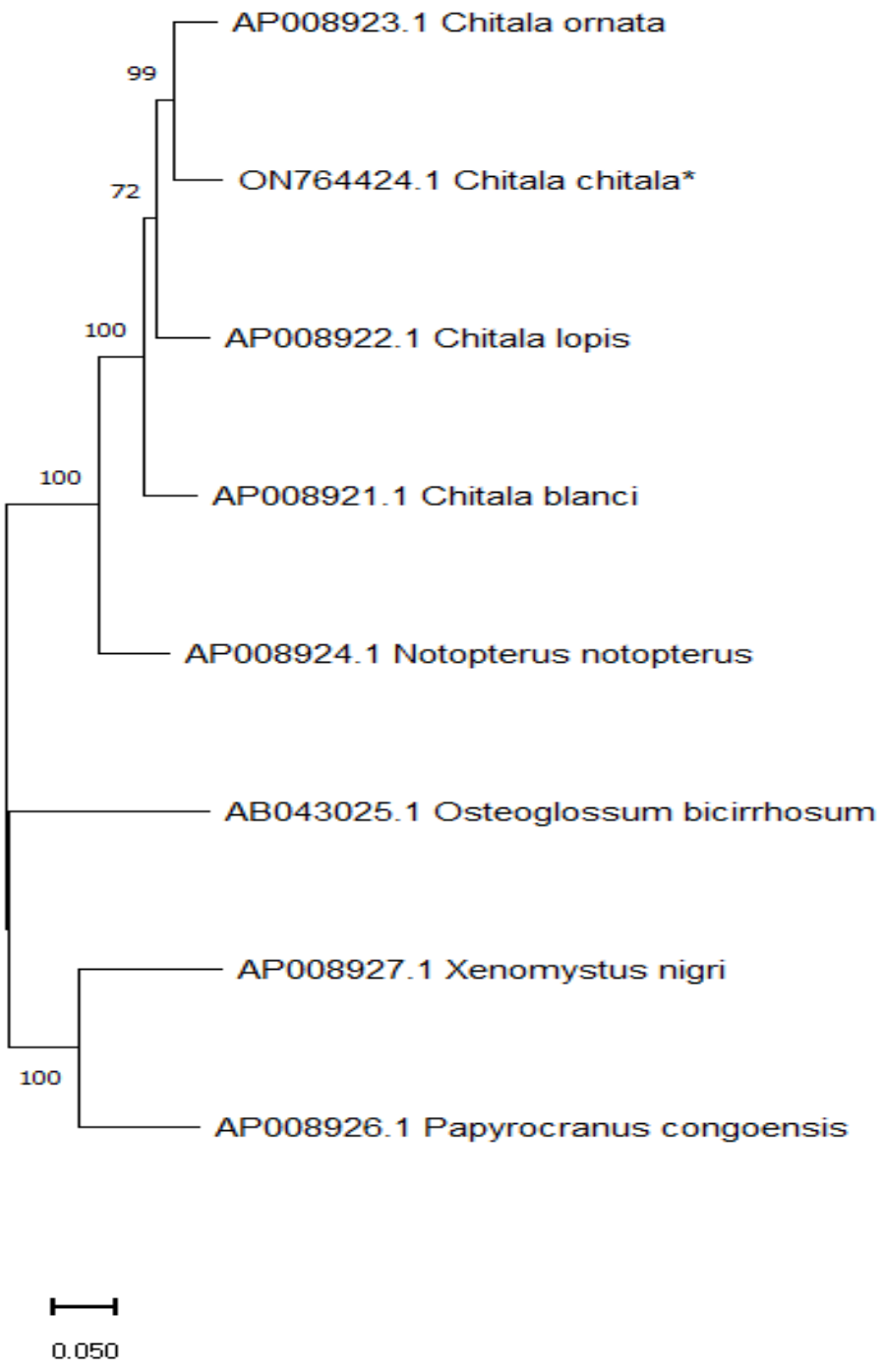


Figure 14. Phylogenetic tree of eight species of the subfamily Notopteridae based on mitochondrial genome sequence following Maximum Likelihood method and bootstrap replications 1000. The star mark (*) indicates present study.

CHAPTER- V

SUMMARY AND CONCLUSION

CHAPTER V

Summary and Conclusion

The mitochondrial genome of *C. chitala* was 16248 bp long and contains 37 mitochondrial genes, including 13 typical protein coding genes, 22 tRNA genes, two ribosomal RNAs (12SrRNA and 16SrRNA), and two non-coding areas (control region, D-loop, and origin of light strand, OL). The 13 protein-coding genes (PCGs) were 11,423 bp in length and accounted for 70.30% of the mitogenome. The basal composition was 25.16% T, 29.84% C, 31.4% A, and 13.5% G. The shortest PCG was ATP8 (168 bp), while the largest was ND5 (1,838 bp). Except for the COX1 gene, which has a unique and alternative start codon, GTG (Valine), the other twelve PCGs began with a unique translation beginning codon, ATG (Methionine). The open reading frame of ND5 ended with TA-, whereas the remaining seven PCGs (ND2, COX1, COX2, ND3, ND4, ND6, and CYTB) ended with the conventional stop codon (T--), as in other vertebrates. Five PCGs, including ND1, ATP8, ATP6, COX3, and ND4L, have an incomplete termination codon.

C. chitala's circular genome had a short subunit of rRNA (12S rRNA) and a large subunit of rRNA (16S rRNA), both of which were 956 bp and 1702 bp in length, respectively. The ribosomal RNA genes were both found on the H-strand and made up 16.35% (2,658 bp) of the entire circular mitogenome. As in other vertebrate genomes, 12S rRNA gene was located between the $tRNA^{Phe}$ and $tRNA^{Val}$ genes and 16S r RNA gene was located between $tRNA^{Val}$ and $tRNA^{Leu}$ genes, respectively. In ribosomal RNAs, the A+T content (54.84%) was higher than the G+C content (45.16%), while the overall base composition of the 12S rRNA gene was A = 32.74%, T = 21.75%, C = 25%, and G = 20.50%, and the 16S rRNA gene was A = 36.13%, T = 19.03%, C = 24.61%, and G = 20.21%.

This study also addressed 22 tRNA genes and their secondary structure. They ranged in length from 67 to 76 bp, for a total length of 1,570 bp (~9.6% of the entire mitogenome). Fourteen tRNA genes were transcribed on the H-strand, while the remaining eight tRNA genes were transcribed on the L-strand. The highest A and T content was found in $tRNA^{His}$ (74.2%), while the lowest level was found in $tRNA^{Val}$ (45%).

C. chitala's primary non-coding area (control region) consisted of 572 nucleotides, which represented 3.5% of the total mitogenome, and the control region was dominated by A+T content (68.7%). The AT rich regulatory region comprises promoters and an origin of replication of mtDNA, both of which are required for transcription and replication of

mtDNA.The control zone is particularly adaptable to size changes.The 32-nucleotide OL region was located between $tRNA^{Asn}$ and $tRNA^{Cys}$ and was orientated on the L-strand in a set of five tRNA genes.

Using the Basic Local Alignment Search Tool, the sequenced data was exported as a FASTA sequence for comparison with reference sequences in the GenBank database. Except for *C . lopis*, the morphologically detected fish species had 99 to 100% identity with the closest comparable GenBank species.

Because of its rarity and delicacy, *C. chitala* is regarded as one of the most promising, important, and expensive fish for food, sport, and aquarium reasons. Mitochondrial DNA (mtDNA) is a helpful raw material as a molecular marker for research of phylogenetics, phylogeography, genetic barcoding, species biodiversity, genetic disorders, and mitochondrial illnesses. Complete genomic analysis of fish has been widely employed in the study of fish evolution.

Decoding the mitochondrial genome in detail might provide vital information for conservation and management of *C. chitala* in Bangladesh.

REFERENCES

REFERENCES

Ahmed, M. S., Chowdhury, M. M. K. & Nahar, L. (2019). Molecular characterization of small indigenous fish species (SIS) of Bangladesh through DNA barcodes. Gene, 684, 53-57.

Alam, M. J., Andriyono, S.,Eunus, A. T. M., Rahman, M. M..& Kim, H. W. (2019). Mitochondrial genome of Mola carplet, *Amblypharyngodon mola* (Hamilton, 1822) and its evolutional relationship in subfamily Danioninae. Mitochondrial DNA Part B, 4(1), 650-651.

Alam, M.T., Petit, R. A., Read, T. D. & Dove, A. D. (2014). The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species. Gene, 539(1), 44-49.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E. & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics, 18(1), 489-522.

Baeza, J. A. (2018). The complete mitochondrial genome of the Caribbean spiny lobster *Panulirus argus*. Scientific Reports, 8(1), 1-10.

Banik, S. & Roy, R. (2014). Effects of climate change on the occurrence of *Chitalachitala* (Hamilton-Buchanan, 1822) in Tripura. Int J Fish Aquat Stud, 2(2), 249-255.

Baxter-Roshek, J. L., Petrov, A. N.& Dinman, J. D. (2007). Optimization of ribosome structure and function by rRNA base modification , 2(1), 174.

Boominathan, A., Vanhoozer, S., Basisty, N., Powers, K., Crampton, A. L., Wang, X. & O'Connor, M. S. (2016). Stable nuclear expression of ATP8 and ATP6 genes rescues a mtDNA Complex V null mutant. Nucleic Acids Research, 44(19), 9342-9357.

Brown, K. H. (2008). Fish mitochondrial genomics: sequence, inheritance and functional variation. Journal of Fish Biology, 72(2), 355-374.

Brown, W. M., George, Jr, M.& Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences, 76(4), 1967-1971.

Burger, G., Gray, M. W.& Lang, B. F. (2003). Mitochondrial genomes: anything goes. Trends in Genetics, 19(12), 709-716.

Burton, G.J., Jauniaux, E. (2011).Oxidative stress. Best Pract Res Clin Obstet Gynaecol,25:287–99.

Cardol, P., Lapaille, M., Minet, P., Franck, F., Matagne, R. F.& Remacle, C. (2006). ND3 and ND4L subunits of mitochondrial complex I, both nucleus encoded in

Chlamydomonas reinhardtii, are required for activity and assembly of the enzyme. Eukaryotic Cell, 5(9), 1460-1467.

Carvalho, D. C., Perini, V. D. R., Bastos, A. S., Costa, I. R. D., Luz, R. K., Furtado, C. & Prosdocimi, F. (2016). The complete mitochondrial genome of the threatened Neotropical catfish *Lophiosilurus alexandri* (Silurifomes: Pseudopimelodidae) and phylogenomic analysis indicate monophyly of Pimelodoidea. Genetics and Molecular Biology, 39, 674-677.

Castro Paz, F. P., Batista, J. D. S. & Porto, J. I. R. (2014). DNA barcodes of rosy tetras and allied species (Characiformes: Characidae: Hyphessobrycon) from the Brazilian Amazon basin. PLoS One, 9(5), e98603.

Cheng, J., Ma, G. Q., Song, N.& Gao, T. X. (2012). Complete mitochondrial genome sequence of bighead croaker *Collichthys niveatus* (Perciformes, Sciaenidae): a mitogenomic perspective on the phylogenetic relationships of Pseudosciaeniae. Gene, 491(2), 210-223.

Chondar, S. L. (1999). Biology of finfish and shellfish. SCSC Publishers (India), West Bengal.

Chowdhury, M. M., Rahman, A. S., Nahar, L., Rahman, M., Al Reza, H. and Ahmed, M. S. (2016). The efficiency of different DNA extraction methods for fish tissues: a comparative analysis. IOSR Journal of Pharmacy and Biological Sciences, 11(3), 11-15.

Clayton, D. A. (1992). Structure and function of the mitochondrial genome. Journal of Inherited Metabolic Disease, 15(4), 439-447.

Cooke, G. M., Chao, N. L.& Beheregaray, L. B. (2012). Divergent natural selection with gene flow along major environmental gradients in Amazonia: insights from genome scans, population genetics and phylogeography of the characin fish *Triportheus albus*. Molecular Ecology, 21(10), 2410-2427.

Dahlberg, A. E. (1989). The functional role of ribosomal RNA in protein synthesis. Cell, 57(4), 525-529.

DoF,(2022).Yearbook of Fisheries Statistics of Bangladesh, (2020-21). Fisheries Resources Survey System (FRSS), Department of Fisheries; Ministry of Fisheries and Livestock, 2022. Volume 38;138 p.

Duarte, F. V., Gomes, A. P., Teodoro, J. S., Varela, A. T., Moreno, A. J. M., Rolo, A. P.& Palmeira, C. M. (2013). Dibenzofuran-induced mitochondrial dysfunction: Interaction with ANT carrier. Toxicology in Vitro, 27(8), 2160-2168.

Elgar, G. &Vavouri, T. (2008). Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. Trends in Genetics, 24(7), 344-352.

FAO. Fisheries and Aquaculture Department(2018). http://www.fao.org.

FRSS. Yearbook of Fisheries Statistics of Bangladesh (2018-19), (Fisheries Resource Survey System, Department of Fisheries, Dhaka, Bangladesh). 2019; 135.

Farias, I. P., Ortí, G., Sampaio, I., Schneider, H. & Meyer, A. (2001). The cytochrome b gene as a phylogenetic marker: the limits of resolution for analyzing relationships among cichlid fishes. Journal of Molecular Evolution, 53, 89-103.

Galib, S.M., Samad, M.A., Mohsin, A.B.M., Flowra, F.A. and Alam, M.T.(2009). Present Status of Fishes in the Chalan Beel- the Largest Beel (Wetland) of Bangladesh, International Journal of Animal & Fisheries Science. 2(3): 214-218.

Ghouri, M. Z., Ismail, M., Javed, M. A., Khan, S. H., Munawar, N., Umar, A. B.& Ahmad, A. (2020). Identification of edible fish species of Pakistan through DNA barcoding. Frontiers in Marine Science, 7, 554183.

Gray, M. W. (1989). Origin and evolution of mitochondrial DNA. Annual Review of Cell Biology, 5(1), 25-50.

Hamilton, F. (1822). An account of the fishes found in the river Ganges and its branches, Archibald Constable and Company, Edinburg. 405 pp.

Handy, S. M., Deeds, J. R., Ivanova, N.V., Hebert, P. D., Hanner, R. H. and Ormos, A. (2011). A single-laboratory validated method for the generation of DNA barcodes for the identification of fish for regulatory compliance. Journal of Association of Official Agricultural Chemists , 94, 201-210.

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends in Ecology & Evolution, 4(1), 6-11.

He, A., Luo, Y., Yang, H., Liu, L., Li, S. & Wang, C. (2011). Complete mitochondrial DNA sequences of the Nile tilapia (*Oreochromis niloticus*) and Blue tilapia (*Oreochromis aureus*): genome characterization and phylogeny applications. Molecular Biology Reports, 38, 2015-2021.

Islam, M. N., Sultana, S. &Alam, M. J. (2020). Sequencing and annotation of the complete mitochondrial genome of a threatened labeonine fish, *Cirrhinus reba*. Genomics & Informatics, 18(3).

IUCN (2015). Red List of Bangladesh volume5: Freshwater Fishes.IUCN, International Union for Conservation of Nature,Bangladesh Country Office, Dhaka, Bangladesh. 2015; 16:360.

Ivanova, N. V., Dewaard, J. R. and Hebert, P. D. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes, 6(4), 998-1002.

Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T. P. & Nishida, M. (2013). MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Molecular biology and Evolution, 30(11), 2531-2540.

Javonillo, R., Malabarba, L. R., Weitzman, S. H.& Burns, J. R. (2010). Relationships among major lineages of characid fishes (Teleostei: Ostariophysi: Characiformes), based on molecular sequence data. Molecular Phylogenetics and Evolution, 54(2), 498-511.

Jondeung, A., Sangthong, P., &Zardoya, R.(2007). The complete mitochondrial DNA sequence of the Mekong giant catfish (*Pangasianodon gigas*), and the phylogenetic relationships among Siluriformes. Gene, 387(1-2), 49-57.

Kim, K. Y., Lee, S. Y., Bang, I. C.& Nam, Y. K. (2008). Complete mitogenome sequence of an Endangered freshwater fish, *Iksookimia choii* (Teleostei; Cypriniformes; Cobitidae) full length research paper. DNA Sequence, 19(5), 438-445.

Kim, M. I., Baek, J. Y., Kim, M. J., Jeong, H. C., Kim, K. G., Bae, C. H. & Kim, I. (2009). Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. Molecules and Cells, 28, 347-363.

Kumar, S.,Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution, 35(6), 1547.

Kurochkin, I. O., Etzkorn, M., Buchwalter, D., Leamy, L. & Sokolova, I. M. (2011). Top-down control analysis of the cadmium effects on molluscan mitochondria and the mechanisms of cadmium-induced mitochondrial dysfunction. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 300(1), R21-R31.

Kyle, C. J. & Wilson, C. C. (2007). Mitochondrial DNA identification of game and harvested freshwater fish species. *Forensic Science International*, *166*(1), 68-76.

Lakra, W. S., Goswami, M. & Gopalakrishnan, A. (2009). Molecular identification and phylogenetic relationships of seven Indian Sciaenids (Pisces: Perciformes, Sciaenidae) based on 16S rRNA and cytochrome c oxidase subunit I mitochondrial genes. Molecular Biology Reports, 36, 831-839.

Lal, K. K., Mandal, A., Singh, R. K., Punia, P., Kapoor, D., Chauhan, U. K. & Mohindra, V. (2006). Genetic divergence in two featherback fishes, *Chitala chitala* and *Notopterus notopterus*. Journal of Applied Ichthyology, 22(5), 369-373.

Li, C., Xiang, X., Ning, T., Chen, S. & Xiao, H. (2017). The complete mitochondrial genome sequence of *Sinocyclocheilus jii* (Cypriniformes: Cyprinidae) and phylogenetic implications. Mitochondrial DNA Part B, 2(2), 638-639.

Li, Q., Du, J., Liu, Y., Zhou, J., Ke, H., Liu, C. & Liu, G. (2014). The complete mitochondrial genome of *Liobagrus marginatus* (Teleostei, Siluriformes: Amblycipitidae).

Liu, S., Yao, J., Zhang, J. & Liu, Z. (2016). Next generation sequencing yields the complete mitochondrial genome of the striped raphael catfish, *Platydoras armatulus* (Siluriformes: Doradidae). Mitochondrial DNA Part A, 27(3), 1963-1964.

Mandal, A., Lal, K. K., Mohindra, V., Singh, R. K., Punia, P., Chauhan, U. K. & Lakra, W. S. (2009). Evaluation of genetic variation in the clown knifefish, *Chitala chitala*, using allozymes, RAPD, and microsatellites. Biochemical Genetics, 47, 216-234.

Mandal, A., Mohindra, V., Singh, R. K., Punia, P., Singh, A. K.& Lal, K. K. (2012). Mitochondrial DNA variation in natural populations of endangered Indian Feather-Back Fish, *Chitalachitala*. Molecular biology reports, 39, 1765-1775.

Mitra, A., Mukhopadhyay, P. K.&Homechaudhuri, S. (2018). An overview of biology and culture potentials of humped featherback *Chitala chitala* (Hamilton, 1822)–a new candidate for aquaculture diversification. Reviews in Fisheries Science & Aquaculture, 26(3), 371-380.

Mohindra, V., Singh, R. K., Kumar, R., Sah, R. S. & Lal, K. K. (2015). Complete mitochondrial genome sequences of two endangered Indian catfish species, *Clarias batrachus* and *Pangasius pangasius*. Mitochondrial DNA, 26(5), 678-679.

Morita, I. (2002). Distinct functions of COX-1 and COX-2. Prostaglandins & other lipid mediators, 68, 165-175.

Ng, P. C.&Kirkness, E. F. (2010). Whole genome sequencing. Genetic variation: Methods and Protocols, 215-226.

Noller, H. F. (1984). Structure of ribosomal RNA. *Annual Review of Biochemistry*, 53(1), 119-162.

Parhi, J., Tripathy, P. S., Priyadarshi, H., Mandal, S. C.& Pandey, P. K. (2019). Diagnosis of mitogenome for robust phylogeny: A case of Cypriniformes fish group. Gene, 713, 143967.

Perini, V. D. R., de Carvalho, D. C., Beheregaray, L. B. & Prosdocimi, F. (2016). The complete mitochondrial genome of the southern purple-spotted gudgeon *Mogurnda adspersa* (Perciformes: Eleotridae) through pyrosequencing. Mitochondrial DNA Part A, 27(1), 380-382.

Prosdocimi, F., de Carvalho, D. C., de Almeida, R. N. & Beheregaray, L. B. (2012). The complete mitochondrial genome of two recently derived species of the fish genus Nannoperca (Perciformes, Percichthyidae). Molecular Biology Reports, 39, 2767-2772.

Rahman,A.K.A. (1989). Freshwater Fishes of Bangladesh, 1st edition, Zoological Society of Bangladesh, Department of Zoology, University of Dhaka, Dhaka-1000, pp. 237-238.

Rangel-Medrano, J. D., Alzate, J. F. & Márquez, E. J. (2016). Complete mitochondrial genome of the Neotropical catfish *Pseudoplatystoma magdaleniatum* (Siluriformes, Pimelodidae). Mitochondrial DNA Part A, 27(6), 4033-4034.

Rawal, Y. K., Dhillon, O. & Sidhu, I. S. (2020). Distinction of two featherback species (Osteoglossiformes: Notopteridae) in India based on scale structure. Records of the Zoological Survey of India, 120(3), 241â-246.

Ruan, H., Li, M., Li, Z., Huang, J., Chen, W., Sun, J. & Zou, K. (2020). Comparative analysis of complete mitochondrial genomes of three Gerres fishes (Perciformes: Gerreidae) and primary exploration of their evolution history. International Journal of Molecular Sciences, 21(5), 1874.

Satoh, T. P., Miya, M., Mabuchi, K. & Nishida, M. (2016). Structure and variation of the mitochondrial genome of fishes. BMC genomics, 17, 1-20.

Shaklee, J. B., Allendorf, F. W., Morizot, D. C.& Whitt, G. S. (1990). Gene nomenclature for protein-coding loci in fish. Transactions of the American Fisheries Society, 119(1), 2-15.

Shamsuzzaman, M. M., Mozumder, M. M. H., Mitu, S. J., Ahamad, A. F. & Bhyuian, M. S. (2020). The economic contribution of fish and fish trade in Bangladesh. Aquaculture and Fisheries, 5(4), 174-181.

Singh, R. K., Dutta, N., Pathak, A., Mohindra, V., Kaur, G., Kumar, R. & Lal, K. K. (2019). The complete mitochondrial genome of Indian featherback *Chitala chitala* (Osteoglossiformes) and phylogenetic status. Conservation Genetics Resources, 11, 473-480.

Singh, R. K., Mohindra, V., Pathak, A., Lal, K. K. & Jena, J. K. (2015). Complete sequence and characterization of mitochondrial genome in great snakehead, *Channa marulius* (Hamilton, 1822). Mitochondrial DNA, 26(3), 473-474.

Song, H. Y., Satoh, T. P. & Mabuchi, K. (2012). Complete mitochondrial genome sequence of the dragonet *Callionymuscurvicornis* (Perciformes: Callionymoidei: Callionymidae). Mitochondrial DNA, 23(4), 290-292.

Strohm, J. H., Gwiazdowski, R. A.& Hanner, R. (2015). Fast fish face fewer mitochondrial mutations: Patterns of dN/dS across fish mitogenomes. Gene, 572(1), 27-34.

Suzuki, T. (2005). Biosynthesis and function of tRNA wobble modifications. In Fine-tuning of RNA functions by modification and editing (pp. 23-69). Berlin, Heidelberg: Springer Berlin Heidelberg.

Talwar, P.K. and Jhingran, A.G,.(1991). Inland Fishes of India and Adjacent Countries, Vol. I, Oxford & IBH Publishing Co. Pvt. Ltd. New Delhi-Calcutta, pp. 63-64.

Tamura, K. & Nei, M. (1993).Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.Molecular Biology and Evolution, 10(3), 512-526.

Wang, L., Chen, Z., Gao, J., Zhao, Y., Sun, P., & Lu, K. (2016). The complete mitochondrial genome of Indonesian tiger fish *Datnioides microlepis* (Bleeker 1854). Mitochondrial DNA Part B, 1(1), 328-329.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M. & Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. Biological Journal of the Linnean Society, 26(4), 375-400.

Wu, Y., Wang, X., Liu, S., Luo, H. & Lin, Q. (2020). Population genetic structure and phylogenetic analysis of gray's pipefish, *Halicampus grayi* in the South China Sea. Genes and genomics, 42, 155-164.

Xu, J., Han, C.& Huang, J. R. (2016). The complete mitogenome of the sheatfish *Pterocryptis cochinchinensis*(Siluriformes: Siluridae) and phylogenetic implications. Mitochondrial DNA Part A, 27(4), 2785-2786.

Zhang, Z., Zhao, L., Song, N.& Gao, T. (2013). The complete mitochondrial genome of *Johnius grypotus* (Perciformes: Sciaenidae). Mitochondrial DNA, 24(5), 504-506.

Zhao, L., Gao, T.& Lu, W. (2015). Complete mitochondrial DNA sequence of the endangered fish (*Bahaba taipingensis*): Mitogenome characterization and phylogenetic implications. ZooKeys, (546), 181.

Zhou, C., Wang, X., Guan, L.& He, S. (2015). The complete mitochondrial genome of *Clarias fuscus* (Teleostei, Siluriformes: Clariidae). Mitochondrial DNA, 26(2), 270-271.

APPENDICES

APPENDIX I (Supplementary Table 1)

Concentration and Quality of the Extracted Genomic DNA

Number of specimen	Genomic DNA	
	Concentration (ng/µl)	Quality (260/280nm)
01	93	2.10
02	184.6	1.88
03	68.1	1.73
04	111	1.99
05	123.7	2.02
06	92.2	1.82

APPENDIX II (Supplementary figure 1)

Gel Electropherogram of genomic DNA of *Chitala chitala*.

Gel electrophoresis show the integrity of DNA. This procedure was done for the extracted genomic DNA. In Figure, the gel was loaded as lane M- Marker (1kb plus), lane 1-specimen 1, lane 2-specimen 2, lane 3-specimen 3, lane 4-specimen 4, lane 5- specimen 5, lane 6-specimen 6.

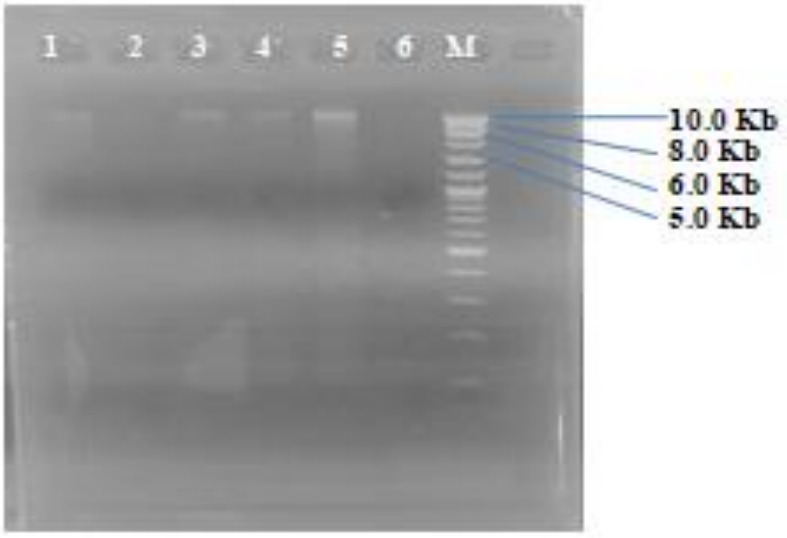


Figure : Gel Electropherogram of genomic DNA of *Chitala chitala*.

APPENDIX III (Partial COI sequence of *Chitala chitala*)

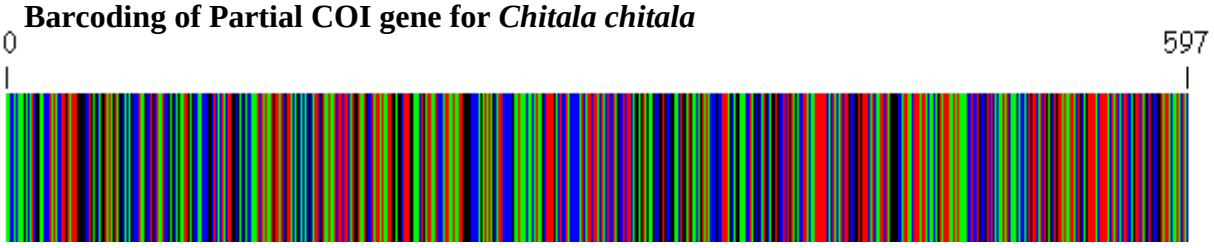
>Partial COI

AACCACAAAGACATCGGAACCCTATACCTTGTATTTGGGGCCTGAGCAGGTATA
GTAGGCACAGCCCTAAGCCTGCTAATCCGAGCAGAATTGAGCCAACCCGGCTCA
CTACTTGGCGACGACCAAATCTATAATGTTATCGTTACAGCACACGCATTCGTAA
TAATCTTCTTCATGGTAATGCCTATTATAATTGGAGGCTTTGGAAACTGATTAATC
CCATTAATAATTGGGGCCCCAGATATAGCATTCCCCCGAATAAACAACATAAGCT
TTTGACTCCTACCCCCATCATTCTTACTACTCCTAGCCTCTTCAGGAGTAGAAGCC
GGTGCCGGAACTGGATGAACAGTATACCCGCCTTTAGCAGGAAACCTAGCGCAT
GCAGGTGCCTCTGTAGACCTTACAATTTTTTCACTACATCTTGCCGGTGTTTCATC
AATTCTAGGGGCCATTAACTTTATTACAACAGTATTTAATATAAAACCTCCTGCC
GTCTCACAATATCAAACACCACTGTTCATCTGAGCTGTTATAATTACTGCAGTTTT
ACTTTTACTATCACTTCCAGTTCTAGCTGCCGGTATTACAATACTAC

Sequence Identity of Partial COI sequence by NCBI BLAST

	Description	Scientific Name	Common Name	Taxid	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☐	Chitala chitala from Bangladesh cytochrome oxidase subunit ...	Chitala ...	clown k...	112163	1103	1103	100%	0.0	100.00%	655	MK572123.1
☐	Chitala chitala voucher OM 109 cytochrome oxidase subunit 1...	Chitala ...	clown k...	112163	1103	1103	100%	0.0	100.00%	652	JQ667555.1
☐	Chitala chitala voucher ZMUD:001 cytochrome oxidase subun...	Chitala ...	clown k...	112163	1098	1098	100%	0.0	99.83%	689	MF140393.1
☐	Chitala chitala voucher UC-CH1 cytochrome oxidase subunit ...	Chitala ...	clown k...	112163	1098	1098	100%	0.0	99.83%	655	JX891536.1
☐	Chitala chitala voucher OM 108 cytochrome oxidase subunit 1...	Chitala ...	clown k...	112163	1098	1098	100%	0.0	99.83%	652	JQ667554.1
☐	Chitala chitala isolate E14 cytochrome c oxidase subunit I (C...	Chitala ...	clown k...	112163	1096	1096	99%	0.0	100.00%	639	MN259187.1
☐	Chitala chitala voucher NBFGR:CC8077A cytochrome c oxida...	Chitala ...	clown k...	112163	1092	1092	100%	0.0	99.66%	655	FJ459464.1
☐	Chitala chitala voucher NBFGR:CC8077B cytochrome c oxida...	Chitala ...	clown k...	112163	1092	1092	100%	0.0	99.66%	655	FJ459465.1
☐	Chitala chitala voucher NBFGR:CC8077D cytochrome c oxid...	Chitala ...	clown k...	112163	1086	1086	100%	0.0	99.50%	655	FJ459467.1
☐	Chitala chitala cytochrome oxidase subunit I (COI) gene, parti...	Chitala ...	clown k...	112163	1083	1083	98%	0.0	99.83%	654	KY909148.1
☐	Chitala chitala isolate bf134 cytochrome c oxidase subunit I (...	Chitala ...	clown k...	112163	1072	1072	97%	0.0	99.83%	600	MK359982.1

Barcoding of Partial COI gene for *Chitala chitala*



APPENDIX IV (AccessionNo. of NCBI GeneBank Reference Sequences)

Fish Species	Accession No.
Chitala ornata	AP008923.1
Chitala chitala	ON764424 (Present Study)
Chitala lopis	AP008922.1
Chitala blanci	AP008921.1
Notopterus notopterus	AP008924.1
Osteoglossum bicirrhosum	AB043025.1
Xanomystus nigri	AP008927.1
Papyrocranus congoensis	AP008926.1

APPENDIX V (*C. chitala* mitochondrion, complete genome)

Chitala chitala mitochondrion, complete genome

GenBank: ON764424.1

FASTA Graphics:

LOCUS ON764424 16248 bp DNA circular VRT 17-MAR-2023
DEFINITION *Chitala chitala* mitochondrion, complete genome.
ACCESSION ON764424
VERSION ON764424.1
KEYWORDS .
SOURCE mitochondrion *Chitala chitala* (Clown Knifefish)
ORGANISM *Chitala chitala*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteoglossocephala;
Osteoglossomorpha; Osteoglossiformes; Notopteridae; Chitala.
REFERENCE 1 (bases 1 to 16248)
AUTHORS Islam,M.N., Akter,M., Sultana,S. and Alam,M.J.
TITLE Assembly and annotation of the complete mitochondrial genome of endangered clown knifefish (*Chitala chitala*) by using next generation sequencing
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 16248)
AUTHORS Islam,M.N., Akter,M., Sultana,S. and Alam,M.J.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-2022) Department of Biotechnology, Sher-e-Bangla Agricultural University, Sher-e-Bangla Nagar, Dhaka 1207, Bangladesh
FEATURES Location/Qualifiers
source 1..16248
/organism="*Chitala chitala*"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:112163"
/country="Bangladesh"
/note="common: Clown Knifefish"
tRNA 1..69
/product="tRNA-Phe"
/anticodon=(pos:31..33,aa:Phe,seq:gaa)
rRNA 70..1025
/product="12S ribosomal RNA"
tRNA 1026..1096
/product="tRNA-Val"
/anticodon=(pos:1058..1060,aa:Val,seq:tac)
rRNA 1097..2798
/product="16S ribosomal RNA"

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

tRNA 2799..2874
/product="tRNA-Leu"
/anticodon=(pos:2834..2836,aa:Leu,seq:taa)
gene 2883..3851
/gene="ND1"
CDS 2883..3851
/gene="ND1"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 1"
/protein_id="WAS32720.1"

/translation="MDMIILIITPLTYIVPVLLAVAFLTLLERKVLGYMQLRKGPNIVGPWGLLQPIAD
GVKLFIKEPVRPYASAPLLFLATPTLALTLALTLWAPMPMPHSVTDLNLGILFILALSSLAVYS
ILGSGWASNSKYALIGALRAVAQTISYEVSLGLILLSAGGFTLHTFNVTQESIWLLAPSWPLAA
MWYISTLAETNRAPFDLTEGESELVSGFNVEYAGGPFALFFLAEYANILLMNTLSTILFLGAT
YNPLLPELTAINLMTKAAILSVLFLVRASYPRFRYDQLMHLVWKSFLPMTLALVLWHTSLPL
SMAGIPPQT"
tRNA 3855..3925
/product="tRNA-Ile"
/anticodon=(pos:3886..3888,aa:Ile,seq:gat)
tRNA complement(3925..3995)
/product="tRNA-Gln"
/anticodon=(pos:complement(3961..3963),aa:Gln,seq:ttg)
tRNA 3994..4064
/product="tRNA-Met"
/anticodon=(pos:4025..4027,aa:Met,seq:cat)
gene 4064..5110
/gene="ND2"
CDS 4064..5110
/gene="ND2"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 2"
/protein_id="WAS32721.1"

translation="MNPYVLTILISSLGLGTTITFASSHWLLAWMGPEINTCAILPLMAKQHHPRAIEAA
TKYFLTQATAAAMILFASTMEAWASGEWNIQQISNQTAMTLLTLAALKIGLAPLHFWMPEV
LQGLDLTTGLVLSTWQKLAPFALIYQISPNTNHTLLVLLGLMSTLIGGWGGLNQTQTRKIMA
YSSIAHLGWMITVLQFMPDLTVLNLTIYITMTSAIFLTLKNISATKINTMATTWSKTPALTATT
MLCLLSLGGLPPLTGFMPKWLILQELLLATLMAMSALLSLFFYLRLCYATTLTISPNTNSQPTP
WRLKTNGTTMPVTISTTLYLLLLLLTPALMALTT"

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

tRNA 5109..5177
/product="tRNA-Trp"
/anticodon=(pos:5139..5141,aa:Trp,seq:tca)
tRNA complement(5177..5248)
/product="tRNA-Ala"
/anticodon=(pos:complement(5215..5217),aa:Ala,seq:tgc)
tRNA complement(5249..5321)
/product="tRNA-Asn"
/anticodon=(pos:complement(5286..5288),aa:Asn,seq:gtt)
tRNA complement(5354..5421)
/product="tRNA-Cys"
/anticodon=(pos:complement(5392..5394),aa:Cys,seq:gca)
tRNA complement(5421..5491)
/product="tRNA-Tyr"
/anticodon=(pos:complement(5457..5459),aa:Tyr,seq:gta)
gene 5493..7041
/gene="COX1"
/gene_synonym="COI"
CDS 5493..7041
/gene="COX1"
/gene_synonym="COI"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:7041,aa:TERM)
/transl_table=2
/product="cytochrome c oxidase subunit I"
/protein_id="WAS32722.1"

/translation="MAITRWLFSTNHKDIGTLYLVFGAWAGMVGTAGSLLGDDQIYNVIVTAHAFV
MIFFMVMPIMIGGFGNWLIPLMIGAPDMAFPRMNNMSWLLPPSFLLLLASSGVEAGAGTGW
TVYPPLAGNLAHAGASVDLTIFSLHLAGVSSILAINFITTVFNMKPPAVSQYQTPLFIWAVMIT
AVLLLLSLPVLAAGITMLLTDRNLNTTFFDPAGGGDPILYQHLFWFFGHPEVYILILPGFGMIS
HIVAYYSGKKEPFGYMGMVWAMMAIGLLGFIVWAHHMFTVGMDVDTRAYFTSATMIIAIP
TGVKVFSWLATLYGGSIKEAPFLWALGFIFLFTVGGLTGIVLANSSLDIILHDTYYVVAHFHY
VLSMGAVFAIMGGFVHWFPLFSGYTLHGTWTKIHFMVMFIGVNLTFFPQHFLGLAGMPRRY
SDYPDAYTNTISSIGSLISLVAVIMFLFILWEAFAAKREVLSVEMTSTNAEWLHGCPPPYHTFE
EPAFVQAKLV"
tRNA complement(7040..7114)
/product="tRNA-Ser"
/anticodon=(pos:complement(7078..7080),aa:Ser,seq:tga)

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

tRNA 7119..7190
/product="tRNA-Asp"
/anticodon=(pos:7152..7154,aa:Asp,seq:gtc)
gene 7191..7881
/gene="COX2"
/gene_synonym="COII"
CDS 7191..7881
/gene="COX2"
/gene_synonym="COII"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:7881,aa:TERM)
/transl_table=2
/product="cytochrome c oxidase subunit II"
/protein_id="WAS32723.1"

/translation="MAHPAQVGLQDAASPVMEELIHFHDHTLMVIFLISNFVLYIIVAVVSTKLTNKY
AHDAQEIEIVWTVLPAVILILIALPSLRILYLMDEINNPHLTVKAIGHQWYWSYEYTDYKDLA
FDSYMIPTQDLTPGQFRLLEVDHRMVIPAESPIRMLITAEDVLHSWAVPSLGIKMDAVPGRLN
QATFIASRPGVYYGQCSEICGANHSFMPIAVEAVPLTHFEDWSTSMLEEA"
tRNA 7882..7956
/product="tRNA-Lys"
/anticodon=(pos:7916..7918,aa:Lys,seq:ttt)
gene 7957..8124
/gene="ATP8"
CDS 7957..8124
/gene="ATP8"
/codon_start=1
/transl_table=2
/product="ATP synthase F0 subunit 8"
/protein_id="WAS32724.1"
/translation="MPQLNPAPWLLMLLFSWLVFLTMIPTKITQHHFMGDPAPQITKK
YTPTPWTWPWH"
gene 8115..8798
/gene="ATP6"
CDS 8115..8798
/gene="ATP6"
/codon_start=1
/transl_table=2
/product="ATP synthase F0 subunit 6"
/protein_id="WAS32725.1"

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

/translation="MTLSFFDQFSITTYLGIPLVALALVLPWILIPTPQKRCLNNRLITLQAWFIRQFTH
QLFMPINKEGHKWALLLASLLIFLMTLNLLGILPYTFTPTTQLSMMGFAVPLWLAAVLIGVRN
QLTHTLAHFLPVGTPGPLIPILIVIETISLLIRPIALLTANLTAGHLLIQLISTAAFAMTSIMPTVSL
LTMALLLLLTILELAVAVIQAYVFVLL LSLYLQESV"

gene 8798..9582
/gene="COX3"
/gene_synonym="COIII"

CDS 8798..9582
/gene="COX3"
/gene_synonym="COIII"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:9581..9582,aa:TERM)
/transl_table=2
/product="cytochrome c oxidase subunit III"
/protein_id="WAS32726.1"
/translation="MARQAHAYHMVDPSPWPLTGATAALLLTSGLAIWFHYNSTILMASGLALMLLT
MYQWWRDIVREGTYLGHHTPPVQKGLRFGMILFITSEVFFFLGFFWAFFHSSLAPTPELGGC
WPPTGIAPLDPFEVPLLNTAVLLASGVTVTWAHHSLMEGARKEAVQSLALTILLGCYFTALQ
AMEYYEAPFTIADGVYGSTFFVATGFHGLHVIVGTTFLAVCLLRQIKYHFTSQHHFGFEAAA
WYWHFVDVVWLFLYVSIYWWGS’

tRNA 9583..9654
/product="tRNA-Gly"
/anticodon=(pos:9616..9618,aa:Gly,seq:tcc)

gene 9655..10003
/gene="ND3"

CDS 9655..10003
/gene="ND3"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:10003,aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 3"
/protein_id="WAS32727.1"

/translation="MNLITTTLIIAAALSCILATIAFWIPQMNPDTEKLSPYECGFDLGSARLPFSLRFFL
VAILFLLFDLEIALLLPLPWGDQLTTPTLTFIWASAILALLTLGLIYEWLQGGLEWAE"

tRNA 10004..10074
/product="tRNA-Arg"
/anticodon=(pos:10036..10038,aa:Arg,seq:tcg)

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

gene 10075..10371
/gene="ND4L"
CDS 10075..10371
/gene="ND4L"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 4L"
/protein_id="WAS32728.1"

/translation="MTPMHFTFSSAFILGLMGLAFHRTHLLSALLCLEGMMLSLFIAALWSLQLESIAY
SAAPMLLLAFSACEASAGLALLVATARTHGTDHLQNLNLLQC"
gene 10365..11745
/gene="ND4"
CDS 10365..11745
/gene="ND4"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:11745,aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 4"
/protein_id="WAS32729.1"

/translation="MLKILIPTIMLFPTTWLVPKQWLWTTTTAQSLIVAALSLTWFKWSSEAGWTSLN
LHLATDQLSTPLLVLTCWLLPLMIIASQNHISTEPINRQRTYISLLILQTFLIMAFGATEIIMFYI
MFEATLIPTLIIITRWGNQTERLNAGTYFLFYTLAGLLVALLITQKNTGTLSMMTMYYTQPLG
FTTWADNIWWLGCLMAFLVKMPLYGVHLWLPKAHVEAPIAGSMVLAAVLLKLGGYGMMR
IVMMLDPLTKQLAYPFIILALWGIIMTGSCLRQTDLKSLIAYSSVSHMGLVAGGILIQTPWGFT
GAIILMIAHGLVSSSLFCLANTNYERTHSRTLLLARGLQTILPLMATWWFIANLANLALPPLPN
LMGELTIITSMFNWSYPTIIITGLGTLITAGYSLYMFLMTQRGPTTTHTISLTPSHTREHLLMVL
HIIPVLLLIVKPELLWGWCA"
tRNA 11746..11815
/product="tRNA-His"
/anticodon=(pos:11777..11779,aa:His,seq:gtg)
tRNA 11816..11882
/product="tRNA-Ser"
/anticodon=(pos:11842..11844,aa:Ser,seq:gct)
tRNA 11882..11956
/product="tRNA-Leu"
/anticodon=(pos:11916..11918,aa:Leu,seq:tag)
gene 11956..13794
/gene="ND5"

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

CDS 11956..13794
/gene="ND5"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 5"
/protein_id="WAS32730.1"

/translation="MHTSLIFNSTLMLILTTLSFPIITSMWTEPLNKTWATTHVKTSIKMAFLTSLIPLFI
FLDQGLEAIITNWNWMNTLTFNFNISFKFDHYSIIFTPVALYVSILEFATWYMHSDPNMNRFF
KYLLLFLIAMITLVTANNMFQLFIGWEGVGIMSFLLIGWWYARADANTAALQAVIYNRVGDI
GLILTMVWLAVNLNSWEIKQIFTLSKDMDLTLPLMGLILAATGKSAQFGLHPWLPSAMEGPT
PVSALLHSTTMVVAGIFLLIRLHPLIENNQMALTTCLCLGALTTFFTATCALTQNDIKKIVAFS
TSSQLGLMMVTIGLNQPQLAFMHICTHAFFKAMLFLCSGSIIHSLYDEQDIRKMGGLNNLLPL
TSSCLIIGSLALTGTPFLAGFFSKDAIIEALNTSHLNAWALTLTLLATSFTAVYSFRVVFFALMG
HPRFLPLTPINENTKTVINPIKRLAWGSIIAGLIISSNQIPMKTQVMTMHPTLKLTALLISITGLIT
AMALANLTAMQHKLKPHTTTHNFSNMLGYYPMTIHRLIPKLNLILGQTMATQLVDQTWFEK
TGPKGISSIQLTPITTVSDTQQGIIKTYLTIFFLTTTVAVTALLLI"

gene complement(13795..14316)
/gene="ND6"
CDS complement(13795..14316)
/gene="ND6"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 6"
/protein_id="WAS32731.1"

/translation="MAILFSMLLIGLLLGLVAVASNPAPYFAALGLVLAAAVGCGILVGCGGSFLSFV
LFLIYLGGMLVVFAYSAALAAEPYPDSWGDWSVFGYVFFYVLGLLFLGLVVSGTWYLGSWF
FVDELKEFSIFRGDFSGVALMYSSGGVMLVIWGWVLLLTLFVVLELTRVLSREALRAI"

tRNA complement(14317..14384)
/product="tRNA-Glu"
/anticodon=(pos:complement(14353..14355),aa:Glu,seq:ttc)
gene 14390..15530
/gene="CYTB"
CDS 14390..15530
/gene="CYTB"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:15530,aa:TERM)
/transl_table=2
/product="cytochrome b"
/protein_id="WAS32732.1"

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

/translation="MASLRKTHPIAKIVNDALIDLPAPVNISAWWNFGSLLGIILTGLFLAMHYTSDIST
AFSSVTHICRDVNYGWLIRNIHANGASFFFICIYLHVARGYYGSYLYKETWNVGVILLLLVM
MTAFVGYVLPWGQMSFWGATVITNLLSAVPYIGDAVQWIWGGFSVDNATLTRFFAFHFLFP
FLIAGATIMHLLFLHETGSNNPMGPFHPYFSYKDLLGFIIMLLALTMLALFSPNLLGDPENFTP
ANPLVTPPHIKPEWYFAYAILRSIPNKLGGVLALLFSILVLVLVPILHTSKMRAMTFRPLSQLLF
WSLVADMAILTWIGGMPVEDPYIIIGQIASTIYFALFLILIPAAGYVENKILQMN"

tRNA 15531..15603
/product="tRNA-Thr"
/anticodon=(pos:15564..15566,aa:Thr,seq:tgt)

tRNA complement(15607..15676)
/product="tRNA-Pro"
/anticodon=(pos:complement(15643..15645),aa:Pro,seq:tgg)

D-loop 15677..16248
/note="control region"

ORIGIN

```
    1 gccagcgtag cttatataaa gcacaacact gaaactgtta agacgggccc taataagacc
   61 ccgcaagcat ggtcccgact tttatgtcag ctacaactaa aattacacat gcaagtgtcc
  121 gcgcccccgt gagaatgccc tcactgccta gcggtttaga ggagccggta tcaggcacac
  181 aaatgtagcc taaaacgcct tgcttagcca cacccccaag ggcacccagc agtgattaac
  241 attaaatata agcgaaagct tgatttagtt atagttaaga gggtcggtta aactcgtgcc
  301 agccgccgcg gttatacgag agaccccagt tgacactatc ggcgtaaagt gtgactacag
  361 aaaaatataa aactaaagcc aaaacctctc aaagccgtta tacgcatatc gagactcgta
  421 ggtccctaaa cgtaagtagc tttaacatat ctgaactcac gaaagctggg aaacaaactg
  481 ggattagata ccccactatg cccagtcata aactttaatg gtaacacact cataccactc
  541 gccagggaac tacgagcgca tcgcttaaaa cccaacggac ttggcggtgc ctcacactcc
  601 acctagagga gcctgttcta taactgaaac tacccgttaa acctcaccgt ttctagccat
  661 cagtctatat accgccgtcg caagctcacc ctatgaagga aatagtagca agcaagaagg
  721 gcttcgccca gaacgtcagg tcgaggtgta gcaaatgaaa cgggaagaaa tgggctacat
  781 tttctgaaca acagaacata cgaacataat tatgaaacac gttacttgaa ggaggattta
  841 gtagtaaaaa gaaaatagag tgttcttttg aaactggctc tgaggcgcgc acacaccgcc
  901 cgtcactctc cactaaacat tttatcactg acattcaaca aaataactaa tataaatatg
  961 tcaccaacac agcggagaca agtcgtaaca cggtaagtgt accggaaggt gcacttggat
 1021 taacccagag cgtagctaag aagccaagca cctcacttac actgagaaga cacccgtgca
 1081 actcgggtcg cactgagcca gagagctagc cgtcacacca cccgcaaata caagacaata
 1141 actcaaaata ttatataaac caaccattaa aacatttttc cgccctagta tatgagatag
 1201 aaaaggatga tagcgccata gaaaaagtac cgcaagggaa agctgaaaaa ggaagtgaaa
 1261 caaaccgcca aagcacaaaa aagcagagat taactctcgt acctttttgca tcatgattta
 1321 gccagaaccg ccaggcgaag agacccttaa gtctgagccc ccgaaactaa acgagctact
 1381 tcgagacagc atatttagag ccaacccgtc tctgtggcaa aagagtggga agatctccaa
 1441 gtagaggtga caaacctaac gagtttagtg atagctggtt gcttaagaaa tgaatactag
 1501 ttcagcctcg tgagacactt acctcaaacc ataaaacgta agaacctaag acgctcacaa
 1561 gagttagcca aaggggggtac agccccttta aaaaagaaca caaccttaac agttgggcaa
```

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

1621 ggatcataat ttataaagca cttgaccacg gtgggcctaa gagcagccat ccgaaaagaa
1681 agcgttgaag ctcaagcaaa caacaagctt attatcccga taaatatatc caacccccact
1741 atcaatatta agccccccta tgactacata ggagagataa tgctaaaatc agtaacaaga
1801 aagcacgact ttctcccggc acatgtgtta gtcagatcgg accacccact gacatataaa
1861 gaactcaata aaagagagta ttgtacccta taaaataacc aagaaaacta cacaaataac
1921 atcgttaacc caacacagga gtgccacaac agggatagac aaaatgaaga aagaaggaac
1981 tcggcaatcc acggccccgc ctgtttacca aaaacatggc ctctcgcaat aacacagatg
2041 agaggtccaa cctgcccagt gaccaagagt taaacggccg cggtatctta accgtgctaa
2101 ggtagcgtaa tcacttgtct tttaaatgaa gacctgtatg aaaggcatta cgagggccca
2161 actgtctcct acttcatgtc agtgaaattg atctgtccgt gcagaagcgg acatacccac
2221 ataagacgag aagaccctgt ggagcttaag atatcaaatt aaccgcgcct agaaactaac
2281 aagcccacgg gccccaaaca ccaaacaagc atagcggcca taattaaact tatcttcggt
2341 tggggcgacc atggaggata aaaaagcctc caagaagaaa caggggggtca gtcacactac
2401 ccctaagagc caagagccac acctctaggc aacagaaaac tctgactaat aatgacccag
2461 gcccagcctg attaacgaac caagttaccc cagggataac agcgcaatcc tttccaagag
2521 cccatatcgc cgaaagggtt tacgacctcg atgttggatc aggacatcct ggtggcgaaa
2581 attttaccaa gggttcgttt gttcaacgat taaagtccta cgtgatctga gttcagaccg
2641 gagtaatcca ggtcggtttc tatctatgaa ctttcccttc ctagtacgaa agggccggaa
2701 ggcgaggggc caataccaaa agcaagcccc actcctacat aatgaacaca actaaaacaa
2761 aaaaggagga tacaaaccaa gcccaagata agggctaagt tgagatggca gagcctggta
2821 attgcaaaag acctaagccc tttcccccag aggttcaact cctcttctca actcttacac
2881 aaatggacat aattattctc atcattaccc cactcactta cattgttccc gtactattag
2941 cggtagcctt cctaacacta ttagaacgaa aagtattagg ctatatacaa ctccgaaaag
3001 gacccaacat tgtaggacca tgaggccttc tacaaccaat cgcagacggt gtaaaactat
3061 tcattaaaga gccagtacga ccatacgcct ccgcccccct gctattccta gccacccta
3121 ccctagctct tacactagcc ctcacactgt gggcaccaat accaataccc cactcagtaa
3181 cagacctaaa cctagggatc ctatttatcc tcgcactatc gagcctggcc gtatactcaa
3241 tcctgggctc gggctgagcc tcaaactcaa aatatgccct catcggcgcc ctacgcgccg
3301 tagcccaaac aatctcgtat gaagtaagcc taggcctaat cttactttcc actattatac
3361 tagcaggagg ctttacactc cacactttca atgtaaccca agaatcaatc tgattactag
3421 ccccaagttg accactagca gcaatatgat acatttcaac tctagcagaa acgaaccgag
3481 ccccattcga cctaactgaa ggtgaatcag agctagtatc cggattcaac gtagaatacg
3541 caggaggacc attcgccctt ttcttcctag cagaatatgc aaacatcctc ctaataaaca
3601 ccctctccac aatcctattt ttaggggcca cctataaccc tctcctcccc gaactcacag
3661 caattaactt aataacaaaa gctgcaatcc tatcagtcct attcctatga gtccgcgcat
3721 cttacccacg attccgatac gaccaactca tgcacctagt gtgaaaaagc ttcctaccaa
3781 taacactagc tctagtactt tgacacacct ccctcccccct atccatagca ggcatccccc
3841 cacaaaccta aaccggaaac gtgcctgaaa gttaagggcc actttgatag agtgaattat
3901 agaggttcaa gtccactcgt ttccttagaa agaaggggct cgaaccctta ccaaagagat
3961 caaaactctt cgtgcttcca atacaccact tcctagtaaa gtcagctaaa caagctttcg
4021 ggcccatacc ccgagaatgt gggttaaacc ccctccttta ctaatgaacc catatgtact
4081 caccatccta atttcaagct taggactagg aacaactatt acatttgcca gctcacactg

4141 actactagct tgaataggcc ctgaaatcaa cacttgcgcc atccttccac ttatagccaa
4201 acaacatcac ccacgagcaa ttgaagccgc caccaaatat ttcctaacac aagccaccgc
4261 agcagcaata atcctattcg ctagcacaat agaagcctga gcatcaggag aatgaaacat
4321 tcaacagatt tcaaaccaaa cagccataac actcctcaca ctagctctcg ccctaaaaat
4381 tggactagct cccctacact tctgaatacc agaagtactt caaggattag accttacaac
4441 cggactcgtc ctttcaacat gacaaaaatt agccccattc gctctcatct accaaatctc
4501 accaaacaca aaccacaccc tactagtact gctcggccta atatccacac tgatcggagg
4561 atgaggtggc ctaaaccaga cccaaacacg aaaaattata gcctactcat caatcgccca
4621 cctaggctga atgatcacag tactacaatt tataccagac ctaacagtac tcaaccttac
4681 aatctacatc acaataacat cagccatctt cctcacccta aaaaacatct ccgccacaaa
4741 aattaacaca atagcaacaa cctggtcaaa aacaccagcc ctcaccgcaa caacaatact
4801 ctgcctccta tccctaggag gcctcccacc acttacagga tttataccta aatgacttat
4861 cctccaagaa ctcactaaac aaaacctacc actactcgcc acactcatgg caataagcgc
4921 tctcctaagc ctctttttct acctacgact atgctacgca acaacactaa caatctcacc
4981 aaacacaaac tcgcaaccca ccccctgacg actcaaaaca aacggcacaa ctatacccgt
5041 aaccatctct acaaccctct acctattact actactacta accccagcct tgatagcact
5101 aacaacctag agacttagga taataagacc aaaagccttc aaagctttaa gcaggagtta
5161 aactctccta gtccctgata agacttgcag gactttagcc cacatcttct aaatgcaact
5221 cagacacttt aattaagcta aagccttact agatgagaag gcttcgaacc tacaaactct
5281 tagttaacag ctaagcgcta aaaccaacta gctttcatct accttcctcc cgccgcggcg
5341 ggaaaaggcg ggaggaagcc ccggcaggca gttagcctac attttcaggt ttgcaatctg
5401 acgtgtttac actacagggc tgataaaaag aggactttaa cctctataca tggagctaca
5461 atccaccgct taaacattca gccattttac ccgtggcaat cacccgctga ctattttcta
5521 ccaaccacaa agacatcgga accctatacc ttgtatttgg ggcctgagca ggtatagtag
5581 gcacagccct aagcctgcta atccgagcag aattgagcca acccggctca ctacttggcg
5641 acgaccaaat ctataatgtt atcgttacag cacacgcatt cgtaataatc ttcttcatgg
5701 taatgcctat tataattgga ggctttggaa actgattaat cccattaata attggggccc
5761 cagatatagc attcccccga ataaacaaca taagcttttg actcctaccc ccatcattct
5821 tactactcct agcctcttca ggagtagaag ccggtgccgg aactggatga acagtatacc
5881 cgcctttagc aggaaaccta gcgcatgcag gtgcctctgt agaccttaca atttttcac
5941 tacatcttgc cggtgtttca tcaattctag gggccattaa ctttattaca acagtattta
6001 atataaaacc tcctgccgtc tcacaatatc aaacaccact gttcatctga gctgttataa
6061 ttactgcagt tttacttta ctatcacttc cagttctagc tgccggtatt acaatactac
6121 ttacagaccg caaccttaac acaacattct ttgacccggc aggcggagga gatccaattc
6181 tttaccaaca cctattctga ttctttggac acccagaagt atatatctta attctcccag
6241 gcttcggtat aatctcccat atcgtcgcct attattctgg taaaaaagaa ccatttggct
6301 atatgggcat ggtctgagct ataatagcca ttggactgct aggctttatt gtatgagctc
6361 accatatatt tacagtaggt atagatgtcg acactcgggc atactttaca tctgcaacaa
6421 taattatcgc aattccaacc ggcgtaaaag tatttagctg attagctaca ctatacggcg
6481 gatcaatcaa atgagaagca ccattcctat gagccctagg attcattttc ctatttacag
6541 taggaggcct aacaggaatt gtactagcca actcatccct agacattatc ctgcatgata
6601 cctactacgt tgttgcccac ttccactatg tcttatcaat aggcgctgta tttgcaatta
6661 taggaggatt cgtacactga ttccctctct tctcaggcta caccctccat ggaacatgaa
6721 ccaaaatcca ctttatagta atatttattg gagttaacct gactttcttt cctcaacact

6781 tccttggact agcaggaata cctcgccgat actccgacta cccagacgcc tacacgctat
6841 gaaatacaat ctcctcaatc ggatcattaa tttcactcgt agcggtaatc atattcctat
6901 ttattctttg agaagcattt gcagctaaac gagaagtcct atccgttgaa ataacctcca
6961 caaacgctga atgacttcac ggctgccctc caccgtacca tacatttgaa gaaccagcct
7021 tcgtccaggc aaaactggtg tcgagaaagg aggcctttga acccccatat gctagtttca
7081 agccagccgc ataaaccgtt ctgccacttt cttacatcaa ggtactagta aaaccgacat
7141 tacattgcct tgtcaaggca aaattgtggg ttaaagcccc acgtatctta atggcccacc
7201 cagcacaagt aggactccaa gacgcagcat ctccagtaat ggaagaactt atccacttcc
7261 acgaccacac attaatagta attttttaa ttagcaactt cgtactatat attattgtag
7321 cagtagtttc aactaaactg acaaataaat atgcacatga tgcccaagaa attgaaattg
7381 tatgaacagt attaccagca gttatcctta ttctaatcgc ccttccatca ctccgcatcc
7441 tctacctgat ggacgaaatc aataacccac acttaacagt caaagctatc ggccaccaat
7501 gatactgatc ttatgaatat accgattaca aagacctagc cttcgactcg tacataattc
7561 caacacaaga ccttaccccc ggccaattcc gccttttaga agttgatcac cgtatagtaa
7621 tccctgccga gtctccaatt cgtatactaa ttacagcaga agacgttctc cactcctgag
7681 cagtcccatc cttaggaatt aagatggatg cggtcccagg ccgtctaaac caagctacat
7741 ttattgcctc ccgcccagga gtttattacg gacaatgctc tgaaatttgc ggggcaaacc
7801 acagctttat accaatcgct gttgaagcag taccattaac acactttgaa gactgatcga
7861 catctatact agaagaagcc tcactaagaa gctaacaagg agatagcgtt agccttttaa
7921 gctaaagatc ggtgactacc gcccacccct agtgacatgc cacaacttaa ccccgcccct
7981 tgacttctaa tattattatt ttcatgatta gtattcctta caataatccc caccaaaatt
8041 acgcaacatc attttatagg agaccccgca ccacaaatta ctaaaaaata cacaccaacc
8101 ccctgaacct gaccatgaca ctaagcttct ttgaccaatt ctccattacc acctatctag
8161 gaatcccgct agttgcccta gcattagtcc tcccatgaat cctaattccc acaccacaaa
8221 aacgatgctt aaacaaccgc ctaattacac tacaagcatg atttatccgc caatttacac
8281 accaactgtt tatgccaatt aataaagaag gacacaaatg agcccttcta ctagcctcat
8341 tattaatctt cctaataaca ctaaatctcc taggcatcct accatacaca tttaccccaa
8401 caacacagct ctcaataaat ataggatttg cagttccact ttgactcgca gcagtattaa
8461 ttggcgtacg taaccaatta acacacacct tagcccactt cttgccagta ggcacccccg
8521 gcccacttat ccccattctg atcgtaatcg aaaccattag cctcttaatt cgcccaattg
8581 ctctcggtgt ccgactcaca gcaaatttaa cagcaggtca cttactaatt caacttatca
8641 gcactgccgc attcgcaata acatcaatca taccaaccgt ttctctacta acaatagccc
8701 tactattatt attaacaatt cttgagctcg cagttgcagt cattcaggca tacgtctttg
8761 ttctacttct aagcctgtac ttacaagaat ccgtataatg gcccgccaag cacacgcata
8821 tcacatagta gaccctagcc cctgaccttt aacaggagca accgctgcct tattactaac
8881 atccggccta gccatctgat tccactataa ttccacaatc cttatagcct caggccttgc
8941 ccttatactt ctaaccatat atcaatgatg acgagatatt gtacgagaag ggacatatct
9001 cggacaccac acacccccag tacagaaagg ccttcgattt ggaataattt tattcatcac
9061 atccgaagta ttcttctttc ttggcttttt ctgagccttc ttccactcta gcctagctcc
9121 aacccccgaa ctaggaggtt gctgaccacc cacaggaatt gcaccacttg acccatttga
9181 agtgccacta ctaaacacag ccgtcctact agcctccgga gtcacagtaa cctgagcaca
9241 ccacagcctc atagaaggtg cacgaaaaga agccgtccaa tcattagcac taacaatcct
9301 actaggctgc tacttcacag ccctacaagc aatagaatac tatgaagcac catttactat
9361 cgccgacggc gtctacggct ccactttctt cgtagcaaca ggatttcacg gcctacatgt

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

9421 tattgtcggc acaactttcc tagcagtttg cctactacga caaattaaat accactttac
9481 atctcaacac catttcggat ttgaagccgc tgcctgatac tgacacttcg tcgacgtagt
9541 atgactcttc ttatacgtct caatctactg atgaggatca taatctttct agtatcaatg
9601 ccagtacatg tgacttccaa ttacacagtc ttggttaaaa cccaaggaaa gataatgaat
9661 ttaatcacaa caacactaat tattgcagca gccctatctt gcatcttagc aaccatcgcc
9721 ttctgaatcc cacagataaa ccctgacaca gaaaaactat caccatacga atgcggcttc
9781 gacccattag gatctgcccg actaccattc tccctccgct tctttttagt agctattcta
9841 tttcttctat ttgacctaga aattgcactc ctcctccccc ttccatgagg agaccaactg
9901 accaccccta cactgacctt catctgagcc tcagcaattt tagcactact aaccctaggc
9961 ctcatttacg agtgattaca aggcggcctt gaatgggcag aataggtaat tagtccaaag
10021 caaagacctc tgatttcggc tcagaaaatt gtggttcaac tccacaatca ccttatgacc
10081 cctatacatt tcacctttag ctcagcattt atcctaggct taatgggctt agcattccat
10141 cgaactcact tattatctgc ccttctctgc ctagagggaa tgatattatc cctatttatt
10201 gcaacagcat tatgatccct acaactagaa tcaatcgcct attcagccgc tcccatgcta
10261 ctcctcgctt tctcagcctg cgaagcaagt gcaggcttgg ccttgctagt agctacagcc
10321 cgaacccacg ggactgacca cctccaaaac ctaaaccttc tacaatgtta aaaattctaa
10381 tccccaccat tatgctgttc ccaacaacat gactggttcc aaagcaatga ctctggacca
10441 ccaccacagc ccaaagccta attgtagctg cactaagcct gacatgattt aaatgaagct
10501 cagaggcagg atgaaccagc ctaaacctac acctagcaac agaccagtta tctacacctc
10561 tcttagtcct aacctgctga ctacttcccc taataattat tgccagccaa aaccacattt
10621 ccacggaacc tattaaccga caacgaacct acatctcctt attaatttca ctgcaaactt
10681 tcttaatcat agccttcggg gcaacagaaa tcattatatt ttatattata tttgaagcta
10741 cactaatccc aaccctaatt attatcacac gatgaggcaa ccagacggag cgattaaacg
10801 caggaaccta cttcctattt tatacactag ctggctcctt accacttctg gtcgccctct
10861 taattacaca aaaaaacaca ggcaccctat ccataatgac catatattat acacaacccc
10921 taggctttac aacatgggcg gacaatattt gatgacttgg ctgtctcata gcgttccttg
10981 ttaaaatacc cctatatgga gtacacctat gattacccaa agctcatgta gaagccccaa
11041 ttgcaggctc aatagttctt gccgcagtat tactaaaact aggaggatac ggcatgatac
11101 gaattgtaat aatactagac cccttaacaa aacaactagc ctacccattt atcatccttg
11161 ccctatgagg cattatcata accggatcaa tctgcttacg ccaaactgac ctaaaatcat
11221 tgatcgccta ttcatccgtt agccacatag gcctggtcgc aggaggtatt ttaatccaaa
11281 ctccatgggg tttcactgga gccattatct taataattgc acatggcctt gtatcctcct
11341 cactattctg cttagccaat acaaactatg aacgcactca cagccgcacc ctactactag
11401 cccgaggcct ccagacaatt ctcccactca tagctacctg atgatttatc gccaacctag
11461 ccaaccttgc tttaccacca ctacccaatc taataggaga actaacaatc atcacctcaa
11521 tattcaactg atcctatcca acaattatca ttacaggcct tggcacccta attacagctg
11581 gctattcact ttacatattc ctaataactc aacgcggtcc aaccacaaca cacacaattt
11641 ctctaacccc atcccacact cgagaacatc tcctcatggt cctccatatt atcccagttt
11701 tactattaat cgtaaaacca gagttattat gaggctgatg tgcctgtaaa tatagtttta
11761 ataaaaatgc tagattgtga ttctagagat ggaagttaaa atcttcttac ttaccaagcg
11821 aggcaggttg cactaaggac tgctaatccc aaaccccatg gttcaaatcc gtggctcact
11881 tagcccctaa aggataacag cccatccgtt ggtcttagga accaaaaact cttggtgcaa

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

11941 ctccaagtag tggctatgca cacctcatta atctttaact ccaccctaat actaatcctg
12001 accacactta gcttcccaat cattacatca atatgaactg aacccctaa caaaacctga
12061 gcaacaaccc atgtcaaaac atccattaaa atagccttcc taactagcct tatcccccta
12121 tttatttttt tagaccaagg tttagaggcc atcattacaa actgaaactg aataaatacc
12181 ctaacattca acttcaacat tagctttaag ttcgaccact actcaatcat ttttacccca
12241 gtagccctct atgtcacctg gtccattcta gaattcgcca cctgatatat acactccgac
12301 ccaaacataa atcgattctt caaatatctt ctcctcttcc taattgcaat aattacccta
12361 gttacagcaa ataacatatt tcaactcttt attggatggg aaggcgtggg aattatgtca
12421 ttcctcctaa ttggctgatg atatgctcga gccgatgcaa acactgcagc cctccaagcc
12481 gtcatctaca accgcgtcgg cgacatcgga ctaatcttaa caatagtatg actagccgta
12541 aacctcaact cctgagaaat caaacaaatc ttcacactat caaaagacat agaccttaca
12601 ctccctctaa taggcctaat ccttgcagca acaggaaaat cagcccaatt tggacttcac
12661 ccatgactac catctgccat agaaggccct acgccagtct ctgccctact ccactcaacg
12721 acaatagtag tcgcaggaat cttcctatta atccgcctac acccattaat tgaaaataac
12781 caaatagccc taacaacttg cctatgctta ggagccctca ccactttctt taccgccact
12841 tgcgctctaa cacaaaatga tatcaaaaaa atcgtagctt tttcaacatc aagtcaacta
12901 ggcctaataa tagtcaccat cggactaaac caaccacaac ttgcattcat acatatctgc
12961 acccatgctt tcttcaaagc catactattt ctatgctccg gctcaattat ccacagcctt
13021 tacgacgaac aagacatccg aaaaataggt ggtctcaaca acctgctccc cctaacatct
13081 tcctgcctaa ttatcggaag cctcgcttta actggcactc ccttttagc cggcttcttc
13141 tcaaaagacg caatcatcga agccctaaac acatcccact taaacgcctg agcccttaca
13201 ctaaccctgt tagccacatc attcacagct gtatacagct tccgagtcgt atttttgcc
13261 ctaataggac acccacgatt cctacccctc accccaatta atgaaaatac taaaacagta
13321 atcaatccta ttaaacgcct agcctgagga agtattattg ccggccttat tatttcatcc
13381 aatcaaattc ctataaaaac acaagtaata acaatacacc caactctcaa attaactgca
13441 ctactaatct ctatcacagg cctaatcaca gcaatagcac tcgcaaacct caccgccata
13501 caacacaaac taaaaccgca cacaacaaca cataacttct ctaatatact aggctattac
13561 ccaataacca tccaccgact aatcccaaaa cttaacctca tcctgggcca aacaatagcc
13621 actcaactag tagaccaaac atgatttgaa aaaacaggcc caaaaggaat ttcatccatc
13681 caactaacac ccatcacaac cgtcagcgat acacaacagg gcattatcaa aacctaccta
13741 accatcttct tcctaaccac cacagtagct gtaacagccc tgctcctaat ctaatctaat
13801 ggcccgcaat gcctcacgac tcaaaacacg agtcaactcc aaaacaacaa acaatgtcaa
13861 caaaagtacc caccctcaaa taactaacat aaccccacca gacgagtata tcaacgccac
13921 cccactgaaa tctccacgaa agatagaaaa ctcctttaat tcatcaacaa aaaatcaaga
13981 ccccagatat caagtgcccg acactactaa acctaaaaac aacaacccta aaacataaaa
14041 gaaaacataa ccaaaaacag accagtctcc ccaagaatca gggtacggct cagcagccaa
14101 tgctgccgaa taagcaaaga caacaagcat acccccctaaa taaattaaaa ataaaacaaa
14161 agacagaaac gacccaccac acccaaccaa aataccacaa ccaactgctg ctgccagtac
14221 caaccccaga gcagcgaaat atggtgcagg gttggaagcc acagccacca accctagcaa
14281 taacccaatt aaaagtatgg agaaaagaat tgccattaat tttgccagga ctttaaccag
14341 gacctgcgac ttgaaaaacc gccgttgcat tcaactacaa aatccaataa tggccagcct
14401 ccgaaaaacc cacccaatcg caaaaatcgt taacgacgca ttaattgacc tacccgcccc
14461 agttaatatc tcagcctgat gaaactttgg ctctctacta ggaatctgcc tcatcgtcca

APPENDIX V (*C. chitala* mitochondrion, complete genome) continued

14521 aattctcacc ggactatttc tagccataca ttatacatcg gatatttcaa ccgccttctc
14581 ctcagtaaca cacatttgcc gagacgtcaa ctatggttga ctaatccgaa atattcatgc
14641 aaacggtgcc tcattcttct ttatctgcat ttatctacac gtagcccgag gcctctacta
14701 cggctcatac ctctacaaag aaacatgaaa cgtgggagtt attctactac ttctagtaat
14761 aatgaccgcc tttgtgggat acgttctacc ttgaggacag atatcattct gaggagccac
14821 agtcattaca aacctcttat ccgccgtccc ctacatcgga gatgctttag tccaatgaat
14881 ctgaggagga ttttcagtag acaacgcaac actaacccga ttcttcgcat tccacttcct
14941 atttccattt ctaattgcag gcgcaactat catgcacctc ctttttttac acgaaacagg
15001 atccaataat ccaataggac taaattcaga cacagacaaa gtgccatttc acccatactt
15061 ttcatacaaa gacctactcg gattcattat tatactccta gccctcacaa tactagcgct
15121 attctcacca aacctactag gagacccaga aaacttcaca cctgcaaacc ccctagttac
15181 ccctccacac atcaaacccg aatgatactt cctattcgca tacgcaatct tacgatccat
15241 ccccaataaa ctaggaggtg tcctagccct attattctca atcctcgtac tagtactagt
15301 accaatccta catacatcca aaatacgagc cataacattc cgaccactat cacaacttct
15361 attctggtcc ttagttgcag acatagccat cctcacatga atcggaggta taccagtaga
15421 agacccatac attattattg gacaaatcgc atccacaatc tacttcgcac tattccttat
15481 cctcatccca gccgcaggtt acgtagaaaa caaaattcta caaataaact gctctagtag
15541 cttaactgat aaagcaccgg tcttgtaatc cgaagattga aggttaaaac ccatcctagc
15601 gcctcctcag agaaaggaga ttttaactcc cgcccttaac tcccaaagct aagattctaa
15661 ttaaaactat tctctgccac acacgaccct atgtatgtcc atttatgtta tatgctatat
15721 tgtaaactac atgttatgta cttaaccatg tctatgcgtg tacccagtaa tgtacatatg
15781 taatatatta tatataatgt attattaaca tgtatgtact ataccatata tgtatgcact
15841 ggtacatact atgtattata ttacatatat atatgtacta gtacatacta tgtattatat
15901 tacatatata tatgtactag tacatactat gtattatatt acatatatat atgtactagt
15961 acatactatg tattatatta catatatata tgtactagta catattatgc attatattac
16021 attaatgtgc ttaataaata ccatgagtgg agccagacat agatttcaaa tcagctacga
16081 acccatcaga caaacgaaat tgtaggctca aataacttgt cctccaatgc ttccttgcgt
16141 tatcaacact tctcgcccca gtacaaaaat gtagtaagag accaccaacg atttaggcag
16201 ggatacactc ttaatgatgg gtcagggaca aatatcgtgg gggtcgca