DNA Barcoding of Selected Aquarium Fishes of Bangladesh

A thesis submitted to the Department of Fisheries, University of Dhaka

in partial fulfillment of the requirements for the degree of

Masters of Science (MS) in Fisheries

Submitted By

Exam Roll: Curzon 809 (MS 2015-16)

Registration Number: 2011-612-782 (Session: 2011-12)

Department of Fisheries

University of Dhaka

Dhaka-1000, Bangladesh

12 February 2017

Declaration

I hereby declare that the dissertation entitled "**DNA Barcoding of Selected Aquarium Fishes of Bangladesh**" submitted to the Department of Fisheries, University of Dhaka for the degree of Master of Science (MS) is based on self investigation, carried out under the supervision of **Mohammad MamunChowdhury**, Associate Professor, Department of Fisheries, University of Dhaka, Dhaka-1000, Bangladesh.

I also declare that this or any part of this work has not been submitted for any other degree anywhere. All sources of knowledge used have been duly acknowledged.

Md. Fayez Mahmud

Exam Roll: Curzon 809 (MS 2015-16)

Registration Number: 2011-612-782

Department of Fisheries

University of Dhaka

Dhaka-1000, Bangladesh

Certificate

I certify that the research work embodied in this thesis entitled "**DNA Barcoding of Selected Aquarium Fishes of Bangladesh**" submitted by **Md. Fayez Mahmud**, Roll Number: Curzon 809, Session: 2015-2016, Registration Number: 2011-612-782, has been carried out under my supervision.

This is further to certify that it is an original work and suitable for the partial fulfillment of the degree of Master of Science (MS) in Fisheries from the Department of Fisheries, University of Dhaka.

I wish every success in his life.

Mohammad MamunChowdhury
Associate Professor
Department of Fisheries, Faculty of Biological Sciences
University of Dhaka, Dhaka-1000, Bangladesh

Acknowledgements

Without the mercy of Almighty Allah, this kind of noble work was not successful for me. So, I would like to give my heartiest gratitude to the Almighty Allah to save me from any kind of problems during the interval of my thesis work.

I convey my deepest appreciation to my reverent teacher and supervisor **Mohammad MamunChowdhury**, Associate Professor, Department of Fisheries, University of Dhaka, for his continual supervision, designing the study, guidance and inspiration throughout the work. I am also indebted to his for his ceaseless endurance for careful reading and reviewing the write-up. Without his constant encouragement and continuous help it would not have been possible to complete the study.

I express my heartiest gratitude to **Dr. Md. Sagir Ahmed**, Professor, Department of Zoology, University of Dhaka, for permitting me to work at his laboratory and providing me the proper support during my thesis work.

I am also grateful to Ministry of Science and Technology for funding for conducting my thesis work. Special thanks to my two elder sisters, Sabrina Rahman Dina and NafisaNawal Islam for untiring help in the laboratory.

I would like to express my appreciation to my friends, Utpal Chandra Ray, Md. MostaviEnanEshik, Md. Monwarul Islam and Md. MunjurHossainfor their friendly company during the preparation of this thesis.

And last of all, I am very grateful to the Department of Fisheries, University of Dhaka.

Author

12 February 2017

Abstract

The ornamental fish sector is a widespread and global component of international trade, fisheries, aquaculture and development. However, the scope of this sector is vast in our country, but we cannot go ahead because of unconsciousness, lack of knowledge, lack of care of government and no government institutions. This sector should be given priority with extra care because it may a lot of foreign exchange every year by exporting the native ornamental fishes that remain unused in our country.

The overall objective of the study was the molecular characterization of selected aquarium fishes to investigate the genetic relationship and to construct the phylogenetic tree among those selected fish species collected from different markets of Dhaka city.

In this study, we used a DNA barcoding approach to identify and characterize the selected species. The partial sequence (651bp) of COI (Cytochrome Oxidase Subunit I) gene was analyzed for species identification, and we compared them with reference sequences from different databases (GenBank and BOLD). The mitochondrial cytochrome oxidase I (COI) region of all samples were successfully amplified using PCR.

Multiple sequence alignment was done using examined DNA sequences and we got 66.42% similarity among selected species. We found three Orders (Cyprinodontiformes, Cypriniformes and Perciformes) among these five samples which indicate the vast genetic divergence of aquarium fishes. This study also shows the purity of species collected from the markets because the sequences of collected sample were matched with the sequences collected from the GenBank of respected species. Phylogenetic tree was constructed then to show the proper genetic divergence among our experimental species.

The findings indicate that DNA barcoding is a powerful technique, accurately identifying samples regardless of sample source. There are several species of aquarium fish in Dhaka city which could not be possible to take under the present study. So, more study is needed to take as many species as possible under investigation to make this sector as a highly profitable business.

Contents

List of Tables

List of Figures

List of Plates

List of Flowcharts

Chapter 1
Introduction

1.1 General Background

Ornamental fish trade is rapidly expanding and there is a growing recreational demand for aquarium fishes in the domestic and international market (Rushna et al., 2016). A total of 19 aquarium shops were found to be working with new promise in Barisal region. It was found that 29 varieties of ornamental fishes including both exotic and indigenous were traded in Barisal division of which the representatives from the fish order Cypriniformes are dominant (Rushna et al., 2016). Price of aquarium fishes varied due to size of each species ranged from BDT 30 to BDT 800 per pair (Rushna et al., 2016). Most of the ornamental fishes were imported by trader. Then they were supplied to district level suppliers. Aquarium accessories such as feed, medicine, artificial light also were traded in aquarium shops in this area.

According to the field survey, the aquarium business in Khulna region is reported to be started in 1988-89 and that time there was only one shop which was the pioneer of this business in this region. At present there are more than 10 shops providing the demand of aquarium and related accessories in this region. Having vast opportunities as well as the demand not only in the local area but also in the world aquarium market this sector till now didn't develop according to the desire of the consumer as well as the provider due to the lack of enough scientific researches. To develop and manage the sector it should be known their aquaculture techniques and genetic characteristics to produce more color variations to increase the market demand. (https://sites.google.com/site/fisheriesinbd /trade/ornamental-fish-and-aquarium).

The species level characteristics analysis of aquarium fish species was carried out through morphological characters and controversial due to phenotypic variation. In this study, molecular characteristics are considered as identification tool by DNA barcoding using COI (Cytochrome oxidase subunit I) gene which is more authentic and will help to create genetic color variations among species.

The use of a globally recognized short DNA sequence, DNA barcode, for identification of species has gained global support as an applicable tool for species identification,

particularly with respect to fishes as coordinated by the fish barcode of life. Fish biodiversity in tropical Africa demonstrate an amazing variety of shape, size, and color. However, many of these fishes are under immense pressure from overfishing and climate change. In addition, the lack of appropriate methods of identification has limited our ability for classification, thus limiting the information available for fishery management. About 65% of fishes captured worldwide have been identified to species level ranging from about 90% in temperate areas to less than 40% in tropical regions. However, there is the need to identify more fish species from Africa, where there is a dearth of information on indigenous fish species. The paucity of taxonomic data on local fish can be ascribed to the limitations imposed by traditional-basedmorphological identification, which can be confusing and unreliable due to problems of intraspecific, phenotypic and cryptic variation often overlapping among sister taxa in nature. Consequently, the limitations imposed by morphological identification, have made the use of molecular diagnostic tools as a prerequisite for effective species identification. DNA-based identification techniques have been developed and shown to be analytically important for characterization of organisms. DNA taxonomic techniques such as DNA barcoding have been useful for species identification and description. DNA barcoding has been used to identify species and is important in characterizing biological diversity. This technique involves the amplification and sequencing of short universal molecular tags from a highly conserved gene. The mitochondrial cytochrome oxidase I (COI) gene is commonly used for thispurpose. The method is rapid, accurate and useful in delineatingdifferences between species. Consequently, the mitochondrial genesCOI and 16S have been successfully employed in species identificationbased on DNA barcodesand a series of barcoding projectsinvolving various organisms from different geographic regions isavailable at the public barcode library.

DNA barcoding is a taxonomic method that uses a short genetic marker in an organism's DNA to identify it as belonging to a particular species. It differs from molecular phylogeny in that the main goal is not to determine patterns of relationship but to identify an unknown sample in terms of a preexisting classification. Although barcodes are sometimes used in an effort to identify unknown species or assess whether species should be combined or separated, the utility of DNA barcoding for these purposes is subject to debate. The most commonly used barcode region, for animals, at least, is a segment of approximately 600 base pairs of the mitochondrial gene cytochrome oxidase I (COI).

Hebert et al. (2003) focused this discussion by proposing that a DNA barcoding system for animal life could be based upon sequence diversity in cytochrome c oxidase subunit 1 (COI). They established that diversity in the amino acid sequences coded by the 5 section of this mitochondrial gene was sufficient to reliably place species into higher taxonomic categories (from phyla to orders). They also found that diversity in nucleotide sequences of the same gene region regularly permitted the discrimination of closely allied species of lepidopterans, a group with modest rates of molecular evolution and high species diversity. As such, these insects provided a challenging test for the ability of COI diversity to resolve species boundaries.DNA barcoding has the potential to identify species (Hebert et al.,2003). The 5region of cytochrome c oxidase I (COI) gene is recommend as the universal and standard barcoding marker for all animals (Hebert et al., 2004).An understanding of the taxonomy and systematics of fish species is a pre requisite for sustainable management of genetic resources. DNA barcoding enhances the prospects for species-level identificationsglobally using a standardized and authenticated DNAbased approach. The application of cytochrome c oxidase I (COI) gene for species identification in fish triggered the international initiative for barcoding all fishes. DNA barcoding provides an accurate and automated species identification system through the use of molecular tags based on short andstandardized mitochondrial genes .TheUS Food and Drug Administration (FDA) is pushing adoption ofDNA barcoding system to solve the species substitution problemin the United States.DNA barcoding appears to hold greatpotential for fish species authentication monitoring system by both regulatory bodies and industry, the utilization of which could enhance transparency and fair trade on the domestic fisheries market.

It's effectiveness has been validated for various animal group and most investigate species (>94%) possess distinct barcodes arrays, with low intraspecific variation and high divergences from closely allied taxa (Ward et al., 2005; Hajbabaei et al., 2006). Barcode sharing has been found between a few congeneric species, largely among taxa that are known to hybridize. Moat prior barcodes studies have generated hypothesis concerning overlooked (Cryptic) species (Hebert et al., 2004), many of which have subsequently been recognized as having morphological and ecological differences (Ward et al., 2005; Hajbabaei et al., 2006). Taxonomists have traditionally utilized morphological characters as taxonomic tools to identify fish species. However, during fish development, morphological characters are not always stable and these characters often cannot be assessed in sectioned specimens. Even when experienced taxonomists

have intact adult specimens to work with, fish identification may not be straightforward if morphological characters are too subtle or if the existing literature and taxonomic history are contradictory. Lack of taxonomic rigour has impeded sustainable use and conservation of worldwide fish resources (Ward et al., 2009), so a reliable and efficient means to authenticate fish species is urgently needed. Molecular identification, that is employing molecular markers to authenticate species, is nowadays widely applied. Initially, this approach employed protein markers but now mainly relies on mitochondrial DNA (mtDNA). Molecular identification based on mtDNA has several advantages over a morphological approach. First, DNA is more resistant to degradation than morphological characters. For example, DNA extraction is still workable from samples that have undergone food processing or digestion. Many mitochondrial genes, such as cytochrome b (cyt b), 16S ribosomal RNA (rRNA) and 12S rRNA, have been utilized as genetic markers for molecular identification.However, the accuracy of molecular identification relieson having a reliable and complete reference database(Hebert, 2007), so inconsistent geneticmarker usage could impede the application of molecularauthentication. Since Hebert et al. (2003) first employedthe cytochrome c oxidase subunit I (COI, which encodespart of the terminal enzyme of the mitochondrial respiratorychain), for species identification, it has been demonstratedthat this genetic fragment could serve as a 'DNAbarcode' for biological authentication in many differentkinds of animals, from invertebrates to vertebrates.

Although Hebert et al. (2003) argued that a COI-based identification system could be developed for all animals, scepticism has been expressed (Mallet andWillmot, 2003). Primary objections have focused on the concern that DNA sequence differences among closely allied species will often be too small to allow their discrimination. Although this issue has never been tested comprehensively, Johns andAvise (1998) demonstrated that closely related species of vertebrates regularly show more than 2% divergence at another mitochondrial gene, cytochrome b. The present study addresses this issue further by examining the extent of sequence diversity at COI among congeneric taxa in the major animal phyla. The most intensive analysis focuses on the arthropods because sequence information for these organisms is particularly detailed owing to their high taxonomic diversity. However, COI divergences are also examined among closely related species in all animal phyla where data are available. In total, sequence divergences are examined in more than 13 000 congeneric pairs including representatives from 11 phyla. These results

support, with the exception of a single phylum, the conclusion that species-level diagnoses can routinely be obtained through COI analysis.

DNA barcoding are already a powerful tool for the identification of aquarium fishes to the species level. We conclude that barcodes have broad application for authenticity testing and the phylogeographic patterning of genetic diversity. It can also inform aspects of traceability. The Fish Barcode of Life (FISH-BOL) campaign (Ward et al. 2009) advocates the use of both COI and BOLD by the ichthyological community for constructing a global reference sequence library designed to enable the rapid, accurate, and cost-effective identification of eggs, larval, juvenile, and fragmentary remains, even by non-specialists.

DNA barcoding aims to provide an effective method for species-level identifications using an array of species specific molecular tags derived from the 5' region of the mitochondrial cytochrome c oxidase I (COI) gene (Namree et al., 2015). The effectiveness of the method hinges on the degree of sequence divergence among species and species level identifications are relatively straight forward when the average genetic distances between sister species. Fishes constitute a high diverse group of vertebrate development. In this context, the identification of fish species is challenging and DNA barcoding provide new perspective in ecology and systematic of fishes.

DNA barcoding is designed to provide accurate and automated species identifications through the use of molecular species tags based on short, standardized gene region. While humanity is facing increasing evidence of the erosion of earth's biodiversity, this approach is providing it's effectiveness in characterizing the complexity of the biodiversity real, at a pace unequalled by other characters. The primary goal of DNA barcoding focus on the assembly of reference tools for species identification in nature. Current result suggest that, in a large array of organism, species are generally well delineated by a particular sequence or by a tight cluster of very similar sequence or that allow unambiguous identification.

With the aim of assigning specimens to known species based on molecular tags, a 648bpsegment of the 5' region of mitochondrial cytochrome c oxidase I (COI) gene forms the library of primary barcodes for the animal kingdom. Mitochondrial DNA (mtDNA) presents several advantages that it well suited for large scale molecular tagging. this genome is present in a large number of copies yielding substantial amounts of

genomic DNA from a variety of extraction methods. Moreover, the high mutation rate and small effective population size make it often an informative genome about evolutionary patterns and process. For a barcoding approach to species identification to succeed, however, within-species DNA sequences need to be more similar to one another than to sequences in different species. Several process such as pseudogenes ontogenesis, introgressive hybridization and retention of ancestral polymorphism pose potential difficulties in capturing species boundaries using mtDNA sequence. The detection of mixed genealogy between closely related species has been previously estimated to occur in nearly 20 percent of the case in the wild. Recent barcoding studies emphasized that this percent can vary widely among phyla, yet species assignment failures typically do not exceed 5 to 10 percent in a large array of organisms.

DNA barcoding exhibits a sufficient level of variation to discriminate among species. The key advantage of DNA barcoding is that once solid reference database has been established, the method does not require expert taxonomic knowledge in order to identify specific samples (Lockley et al., 2000). Further identification can be done with small tissue samples from virtually any part of the organism, does not require reproductive material, and the identification process is generally fast and reproducible. A limitation of the method is that no single universal DNA regions that can be used across all taxonomic groups have been identified. While many DNA barcoding studies in animals have achieved high rates of species discrimination using a single region, COI for plants it has proven necessary to use a combination of regions to obtain sufficient discrimination success. DNA barcoding technology is an emerging molecular-based authenticity technology that uses variation within a single genetic market region (i.e., the COI region of mitochondrial DNA) to identify plants and animals species through DNA sequencing (Hebert et al., 2003).

1.2 Problem Statement

Identification and characterization of aquarium fishes were usually based on morphological characters. However considerable ambiguity exists due to morphological similarity and they have led to some controversial hypotheses on species identification (Uchid, 1939; Jayaram, 1968; Lee, 1990; Kim et al., 2003). As aquarium fish are closely related and has potentiality to lead their whole life in a closed and small aquarium, the detection of their species is so important in the aquarium fish business to find the correct

species with required characters. On the other hands, color variation is also a vital factor among aquarium fishes which creates business demand.

So, only morphological characters are not sufficient to characterization of aquarium fishes. On the other hand molecular tools are advantageous for fish identification because large number of fish species from distinct live history stages (Eggs, fry and adults) can be examined.

DNA barcoding has the potentiality to be used for characterization of aquarium fishes as an effective and authentic tool. Indigenous and exogenous fishes can be separated easily by DNA barcoding which leads a better management of this sector and minimize the hazardous effects of exogenous fishes during import of the fishes.

1.3 Rationale

Fish species identification is traditionally based on external morphological features. Yet, in many aquarium fishes and especially their diverse developmental stages are difficult to identify by morphological characters. DNA based identification methods offer an analytically powerful addition or even an alternation. This work intends to provide an updated and extensive overview on the PCR methods for fish species identification. Besides approach of molecular toots can provide valuable information for species identification to complement the taxonomic data and validation of systematic positions and phylogeny.

1.4 Research needs

Very few literatures are available on aquarium fishes found in Bangladesh and most of them are based on morphological characters and business issue. A literature was published in 2016 entitled "A trend of ornamental fish business in Barisal division, Bangladesh" in the journal of "International Journal of Fisheries and Aquatic Studies" by RushnaAlam and Jahangir Alam. Another literature was published in 2016 entitled "A study on aquarium fish business in Jessore district, Bangladesh" in the journal of "Journal of Scientific Research" by Shams Muhammad Galib.

Literature on molecular characterization and phylogenetic assessment of aquarium fishes in Bangladesh was not found. This study was thus aimed to characterize some selected

aquarium fishes by DNA barcoding with the phylogenic relationship to distinguish the species of selected fishes which are available in different markets in Dhaka city.

1.5 Socio-economic importance

Traditionally, the people of Bangladesh are pleasure seeker in nature and it has a great potential for ornamental fish domestication not only through indigenous species but also to furnish with exotics verities. Globally ornamental fish keeping treats as an industry due to its high market demand and profitability as well as it is a hobby that has been rated as second most popular and most popular type of pet in industrialized country. Ornamental fish keeping is a multi-dollar industry comprising the harvest, sale and displaying live fish in aquarium and garden pond and lakes. Although some reported that some aquatic plants and some other aquatic animal such as snails and turtle also involved in this sector with associate aquarium fishes. It was projected that 150 million ornamental fish both marine and freshwater were sold each year worldwide and the trade and its associated aquarium accessories was worth over US$7 billion annually, in the late 1980s (Rushna et al., 2016). After that this trend was increased by approximately 14% per year. More recently, it was estimated that 1,471 marine ornamental fish and more than 4,000 species of ornamental freshwater fish were trade globally each year (Rushna et al., 2016). In spite of the availability of limited data are available about the status of ornamental fish business in Bangladesh, it can be said that it become a profitable business in the capital as well as in some other big cities.

1.6 Objectives

The overall objective of the study was the molecular characterization of selected aquarium fishes collected from Dhaka city. The specific objectives were,

1. To characterize 5 species of aquarium fishes through DNA barcodingusing mitochondrial COI (Cytochrome oxidase subunit I) gene.

2. To investigate the genetic relationship among those selected aquarium fishes collected from different markets of Dhaka city.

3. To construct the phylogenetic tree of selected fishes using DNA sequences.

Chapter 2
Materials and Methods

2.1 Experimental fish

In the present study, 5 different species of the fish were examined. The samples were collected from different local markets of Dhaka metropolis. Muscle sample from dorsal portion of the selected fishes were examined for each specimen.

Table 2.1. List of the fish species used in this study

No of Fish	Scientific Name	English Name	Collection Place
1	*Poeciliasphenops*	Black Molly	Katabon, Dhaka
2	*Poeciliavelifera*	Yucatan Molly	Katabon, Dhaka
3	*Puntigrustetrazona*	Tiger Barb	Katabon, Dhaka
4	*Scatophagusargus*	Spotted Scat	Nature Aquatics, Uttara, Dhaka
5	*Carassiusauratus*	Goldfish	Katabon, Dhaka

2.2 Collection of samples

A total of 5 samples of aquarium fishes were collected from different markets of Dhaka metropolis. Fish samples were collected in sterilized box. Sampling was done during the periods May 2016 to December 2016. After sampling, the samples were transferred in the laboratory using icebox within required time.

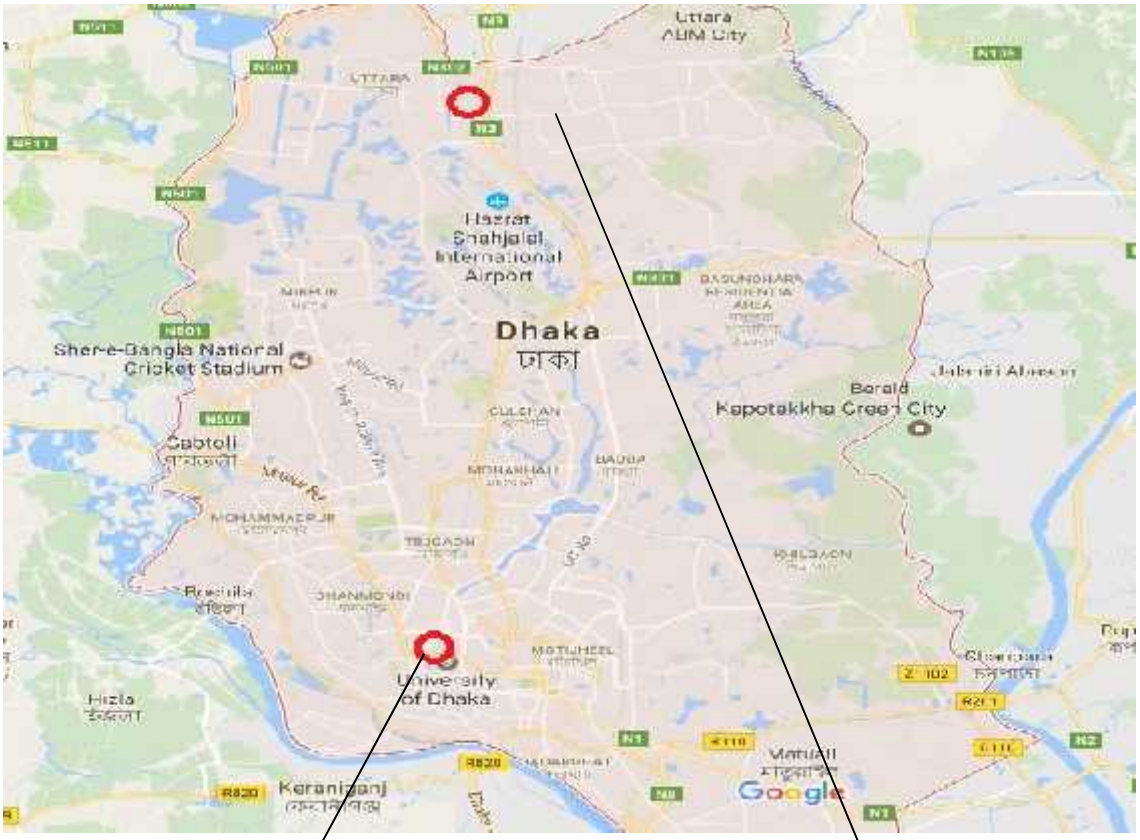


Fig.2.1.Map of Dhaka City

Fig. 2.2.Location of Katabon Market

Fig. 2.3. Location of Nature Aquatics

2.3 Transportation and preservation of samples and laboratory of investigation

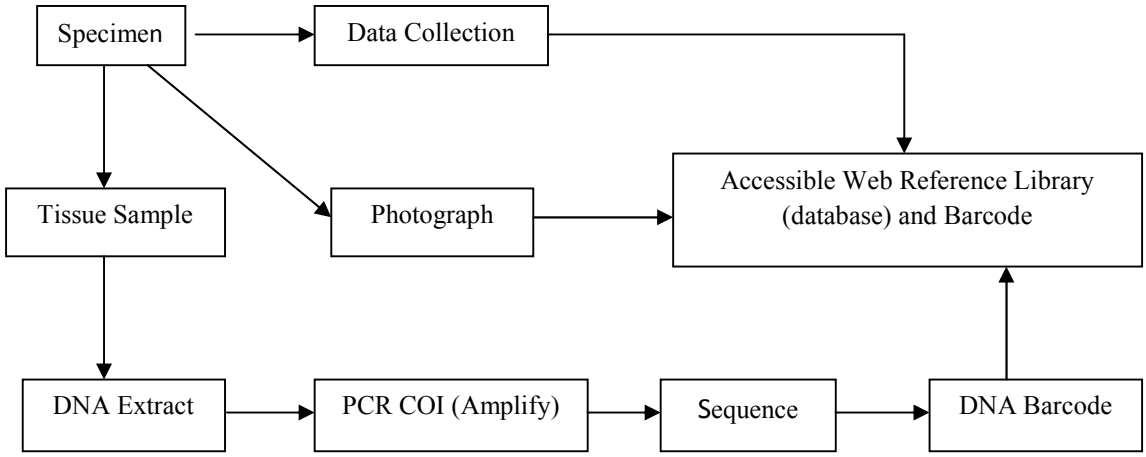
Samples were transported using icebox and stored at -20^{o}C and tissue sample collection was carried out in the Aquatic Laboratory of the Department of Fisheries, University of

Dhaka. Then samples were transmitted to the Department of Zoology, University of Dhaka for molecular activities (DNA extraction, DNA amplification by PCR, DNA visualization by Gel electrophoresis, purification of PCR products). DNA sequencing was done in the center of Advanced Research in science (CARS), University of Dhaka.

2.4 Molecular identification

Molecular identification basing on DNA barcoding is widely accepted method and is being applied in species identification which requires some steps.

Flow Chart 1.DNA barcoding process.Adapted from Floyd et al., (2010).



2.5 Isolation of DNA from fish muscle tissue

2.5.1 Chemical used for isolation of muscle DNA

All the solutions were made with de-ionized, sterile water.

a) 1 M Tris-HCl (pH 8.0, adjusted with concentrated HCl)

b) 0.5 M EDTA (pH 8.0, adjusted with concentrated NaOH)

c) TE buffer; pH 8.0

d) Phenol : Chloroform : Isoamyl alcohol (25:24:1 ; v/v)

e) 70% Ethanol

f) RNase solution (DNase free)

g) Liquid Nitrogen

h) SET isolation buffer (per 100 ml)

10% Sarcosyl 20 ml

5 M Nacl 02 ml

0.5 M EDTA 20 ml

50 mMTris-HCl05 ml

The volume was adjusted to 100 ml and stored at room temperature.

2.5.2 Required materials for agarose gel electrophoresis

a) Agarose

b) Ethidium bromide

c) 6x loading buffer

d) Electrophoresis buffer (TAE, pH 8.0)

e) DNA markers:

i. Gene Ruler^{TM} 1kb Plus DNA Ladder

ii. Gene Ruler^{TM} 1kb DNA Ladder

f) Gel casting tray, combs and power supply

Table 2.2. Random primers used in PCR and sequencing.

Primer Code	Sequence (5´-3´)	References
Fish F1	TCAACCAACCACAAAGACATTGGCAC	Ward et al., 2005
Fish F2	TCGACTAATCATAAAGATATCGGCAC	Ward et al., 2005
Fish R1	TAGACTTCTGGGTGGCCAAAGAATCA	Ward et al., 2005
Fish R2	ACTTCAGGGTGACCGAAGAATCAGAA	Ward et al., 2005

2.6 Genomic DNA isolation

DNA was isolated using the modified SDS method of Koh et al. (1999). The method is described below,

1. 250 mg freshly collected muscle taken in liquid nitrogen and grinded to fine powder using mortar and pestle.

2. 1ml of extraction buffer was added and grinded the muscle until it became homogenous paste.

3. The paste was transferred to an appendrof tube (2.0 ml) and incubated at 60^{o}C in a water bath for 15 minutes.

4. Then added 10µl (10mg/ml) proteinase K to breakdown protein skeleton and also added 100µl 1M DDT and followed by incubation at least 6 hours to dissolve the muscle tissue.

5. The sample were centrifuged at 4000 rpm for 20 min at room temperature to remove non-soluble debris

6. The supernatants were transferred to fresh appendrof tube and equal volume of phenol: Chloroform: Isoamyl alcohol (25:24:1) was added and mixed well by slow inversion, then centrifuged the tube at 6500 rpm for 10 minutes. Thus step was repeated for two times.

7. The supernatants were collected and 1/10 vol of 3M sodium acetate (pH 8.0) and 0.6 volume of chilled Isopropanol were added and mixed slowly precipitate DNA (Sambrook et al., 1987). Mixed the contents and kept into refrigeration overnight.

8. The samples were centrifuged for 10 min at 6500 rpm at room temperature. The supernatants were discarded carefully by using adjustable micropipette.

9. The pellet was washed with 70% ice-cold ethanol. The washing step was repeated at least 2-3 times. The pellets were air dried on a paper towel for about 1 hour.

10. The dried DNA was dissolved in 100µl of TE buffer and treated with RNase A and 10 µl of 1M DTT and incubated at 37^{o}C for 2 hours.

11. The samples was precipitated with isopropanol and sodium acetate and further washed with 70% ethanol and pure DNA pellet was re dissolved in 100 µl of TE buffer.

2.7 Estimation of quality and quantity of isolated DNA sample

Before PCR amplification it is important to know the quality and quantity of genomic DNA because different DNA extraction methods produced DNA of different purity. It is necessary to optimize the amount of DNA to achieve reproducibility and strong signal in PCR array. Excessive, as well as too little DNA may not produce clearly bands in the gel. Measurement of DNA concentration was carried out in this investigation by estimating the absorbance of DNA by spectrophotometer.

2.8Agarose Gel Electrophoresis

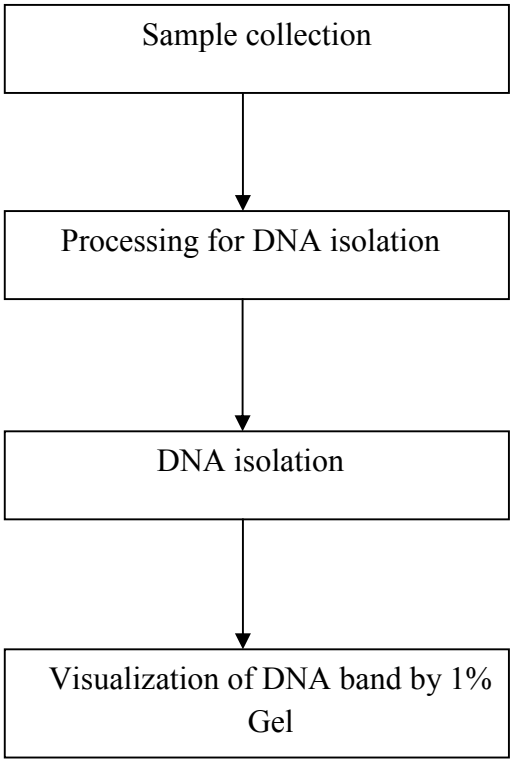
1. The Standard method used to separate, identify and purify DNA fragments through electrophoresis was followed according to the method described by Sharp et al. (1973).

2. 1.0 g of Agarose was heated to melt into 100 ml of TAE buffer, ethidium bromide was added (10µg/ml) and poured into gel casting tray fixed with appropriate combs.

3. After the gel was solidified it was placed into gel-running kit containing 1xTAE buffer.

4. Digested plant DNA solutions were loaded with 6x gel loading dye and electrophoresis was continued until DNA fragments were separated well.

Flow Chart 2.Flowchart of DNA isolation procedure.



2.9 Preparation of PCR Reaction Mixture

The following components were used to prepare PCR reaction The total volume of PCR reaction was for 5 samples.

Table 2.3.Components of PCR reaction mixture for 5 reactions.

Sl. No	Reagents	Amount per sample (µl)	Total
1	Sterile de-ionized distilled water	18.7	93.5
2	*Tag* buffer A 10x	2.5	12.5
3	Primer	1.0	5.0
4	dNTPs 10mM	0.25	1.25
5	*Tag* DNA Polymerase	0.05	0.25
6	Template DNA	2.5	12.5
	Total	**25**	**125**

During the experiment, PCR buffer, DNTPs, Primers and DNA sample solution were thawed from frozen stocks, mixed well by vortex and kept on ice.

2.10 PCR amplification

The optimum amplification cycle was as follows done by oil-free thermal cycler (Biometra, UNO 11).

No. of cycles	Steps	Temperature	Time
	Initial denaturation	94^{o}C	4 minutes
	Denaturation at	94^{o}C	45 seconds
40 Cycles	Annealing at	54^{o}C	1 minutes
	Extension at	72^{o}C	1 minutes
	Final extension at	94^{o}C	5 minutes

After completion of cycling programme, the reactions were held at 4^{o}C.

2.11 Electrophoresis of the Amplified DNA and documentation

The amplified products were separated electrophoretically on 1% agarose gel. The gel was prepared using 1.0g agarose powder containing ethidium bromide and 100ml 1xTAE buffer at 90 Volts for 30 minutes. DNA bands were observed on UV-transilluminator and photographed by a Gel Documentation system.

2.12 PCR product purification

- PCR product has transferred to 1.5ml appendrof.
- 5 times FADF buffer added and centrifuged at 1100 rpm for 1 minute.
- Flow through discarded and 750 ml buffer was added and centrifuged at 11000 rpm and liquid discarded.
- Again centrifuged at 14000 rpm for 10 minutes.

- Then 40µl of elution buffer added to the membrane centre of FSDF column and waited for 3 minutes for fully absorption.
- Centrifuged at full speed at 14000 rpm for 2 minutes to elute the DNA.
- Then the liquid further placed in column tube and centrifuged at 14000 rpm for 2 minutes.

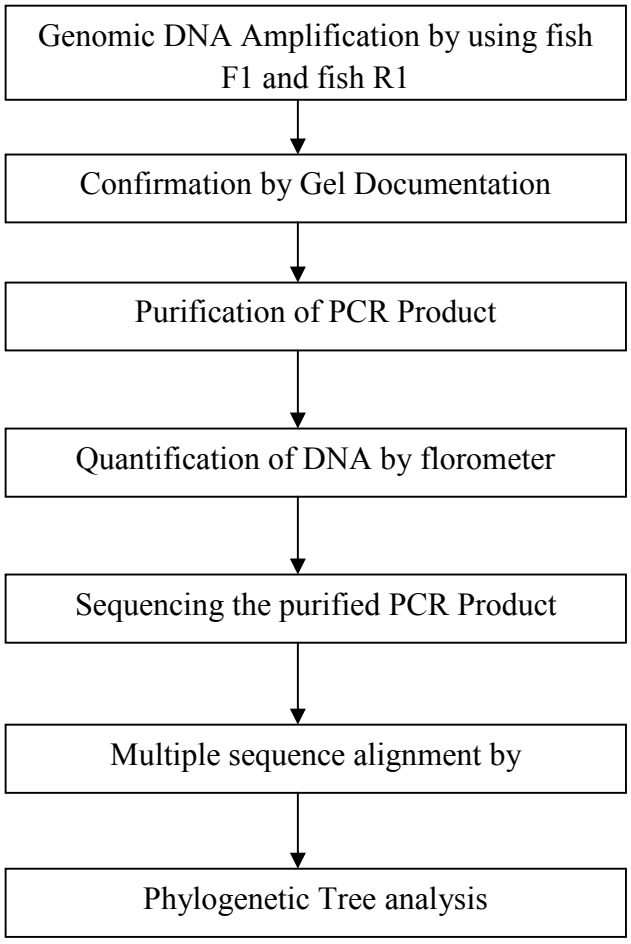
2.13 Quantification of DNA concentration

The quantification of DNA concentration is very important because very high as well as very low concentration of DNA may not give the band at the gel run in electrophoresis. Vortexed 200µl of working solution with Qubit^{TM} buffer for 3 minute and determine the concentration of the original sample by Fluorometer.

Table 2.4.Measurement of DNA concentration of 5 samples of aquarium fishes.

Sample ID	Concentration of DNA (ng/µl)
F3	16.5
F4	11.9
F5	17.2
F16	12.5
F17	13.8

Flow Chart 3.Working schedule after visualization of genomic DNA



2.14 Software used for sequence analysis

2.14.1 NCBI BLAST

In bioinformatics, BLAST for Basic Local Alignment Search Tool is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of proteins or the nucleotides of DNA sequences. A BLAST search enables a researcher to compare a query sequence with a library or database of sequences, and identify library sequences that resemble the query sequence above a certain threshold.

Different types of BLASTs are available according to the query sequences. For example, following the discovery of a previously unknown gene in the mouse, a scientist will typically perform a BLAST search of the human genome to see if humans carry a similar gene; BLAST will identify sequences in the human genome that resemble the mouse gene based on similarity of sequence. The BLAST algorithm and program were designed

by Stephen Altschul, Warren Gish, Webb Miller, Eugene Myers, and David J. Lipman at the National Institutes of Health and was published in the Journal of Molecular Biology in 1990.

BLAST is one of the most widely used bioinformatics programs for sequence searching. It addresses a fundamental problem in bioinformatics research. The heuristic algorithm it uses is much faster than other approaches, such as calculating an optimal alignment. This emphasis on speed is vital to making the algorithm practical on the huge genome databases currently available, although subsequent algorithms can be even faster.

Before BLAST, FASTA was developed by David J. Lipman and William R. Pearson in 1985. BLAST is more time-efficient than FASTA by searching only for the more significant patterns in the sequences, yet with comparative sensitivity. BLAST is also often used as part of other algorithms that require approximate sequence matching.

2.14.2 MEGA 7.0.14

Molecular Evolutionary Genetics Analysis(MEGA) is computer software for conductingstatisticalanalysis of molecular evolution and for constructingphylogenetic trees. It includes many sophisticated methods and tools forphylogenomicsandphylomedicine. It is licensedas proprietary freeware.

In version 7.0.14, we have now added facilities for building molecular evolutionary trees scaled to time (timetree), which are clearly needed by scientists as an increasing number of studies are reporting divergence times for species, strains, and duplicated gene (Kumer, 2011; Ward et al., 2013).

2.14.3 CLUSTALW

CLUSTALW is a general purpose multiple alignment program for DNA or protein. The program performs simultaneous of many nucleotide or amino acid sequences. It is typically run interactively, providing a menu an online help. CLUSTALW (Thomson et al. 1994) is the most widespread among many available algorithms. It can be accessed through web site or downloaded locally in personal computer. CLUSTALW must be searched in the NPS section.

Plate 1. Photograph of*Scatophagusargus*

Plate 2. Photograph of*Puntigrustetrazona*

Plate 3.Photograph of*Poeciliavelifera*

Plate 4.Photograph of*Poeciliasphenops*

Plate 5. Photograph of*Carassiusauratus*

Chapter 3

Results

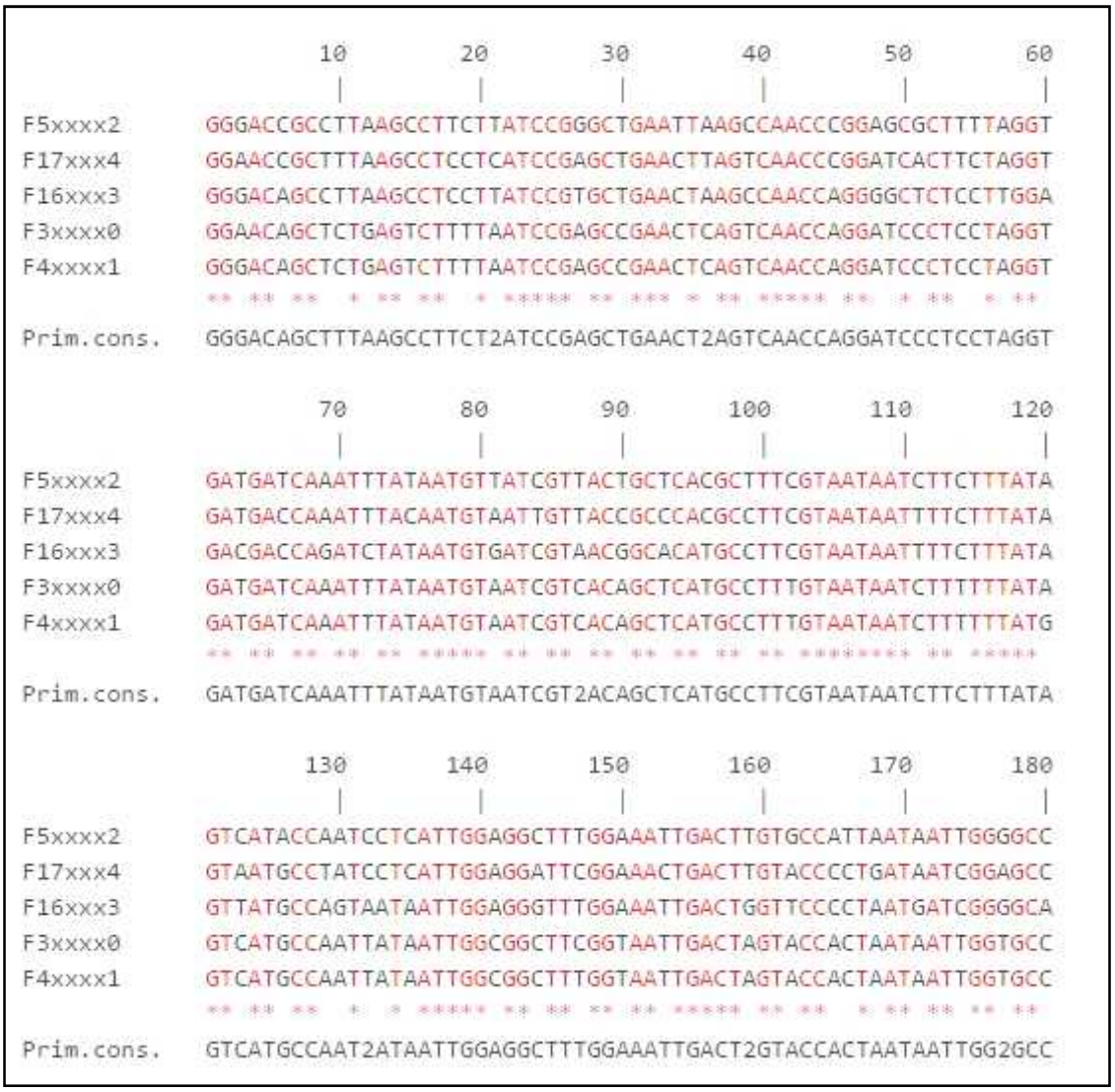
Full length DNA barcodes (551bp) were recovered using the Fish primers Fish F1 and Fish R1. DNA of five selected aquarium fish obtained from markets in Dhaka city was amplified using the primer set. Peak intensities and sequencing qualities of the generated barcodes were compared to the sequences downloaded from NCBI GeneBank.

Table 3.1.List of all identification results using the Gene Bank as search engine.

Sl. No	Description	Max Score	Total score	Query cover	E value	Identity	Accession
F3	***Poeciliasphenops***voucher KW11T075 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	1018	1018	100%	0.0	100%	KU568972.1
F4	***Poeciliavelifera*** voucher KW11T074 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	1018	1018	100%	0.0	100%	KU568973.1
F5	***Puntigrustetrazona***voucher KW11T020 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	1018	1018	100%	0.0	100%	KU569014.1
F16	***Scatophagusargus***haplotype ScarH1 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	1018	1018	100%	0.0	100%	KC774668.1
F17	***Carassiusauratus***voucher KW11-T309 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	1018	1018	100%	0.0	100%	KU568778.1

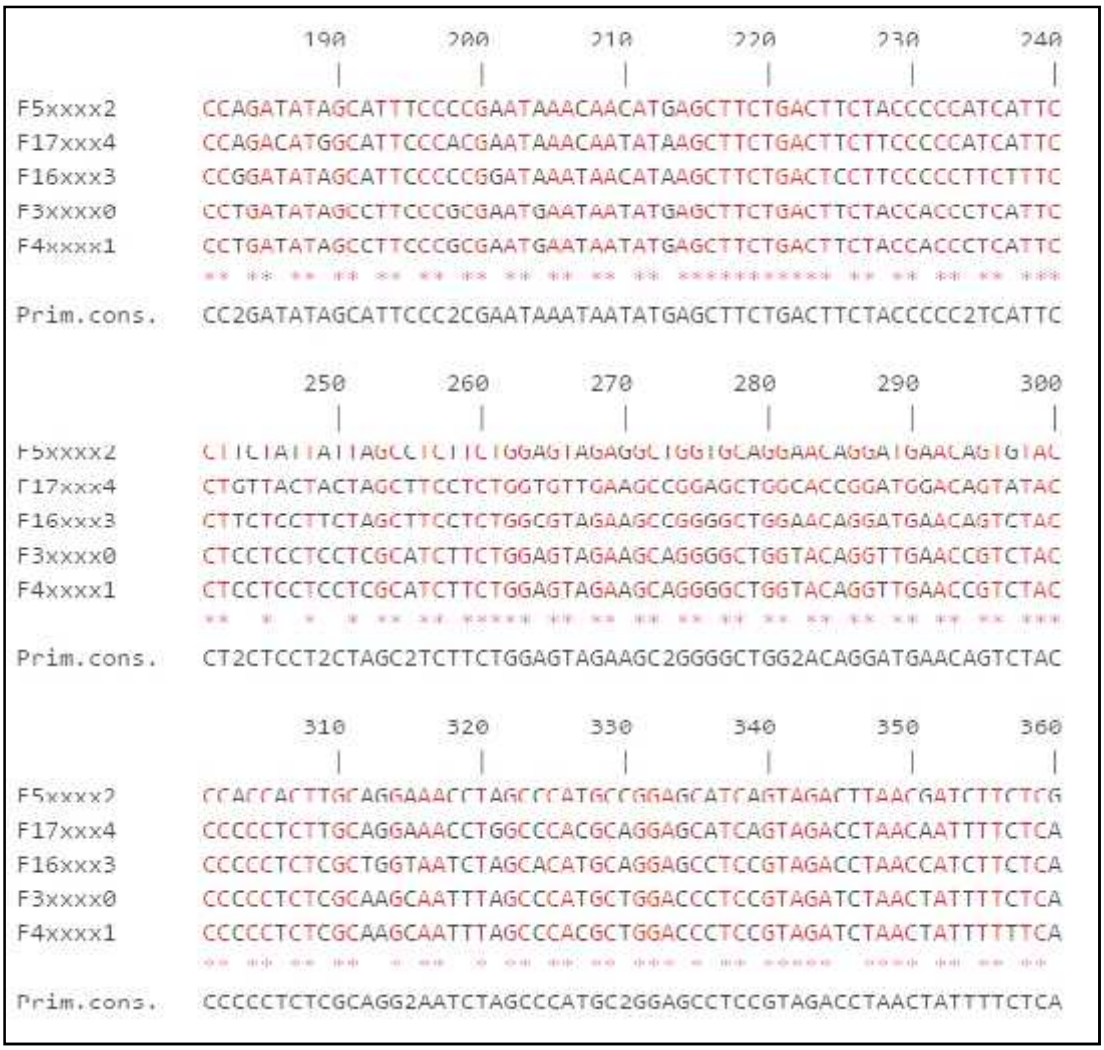
3.1 DNA sequence alignment

3.1.1 Multiple sequence alignment between selected 5 samples



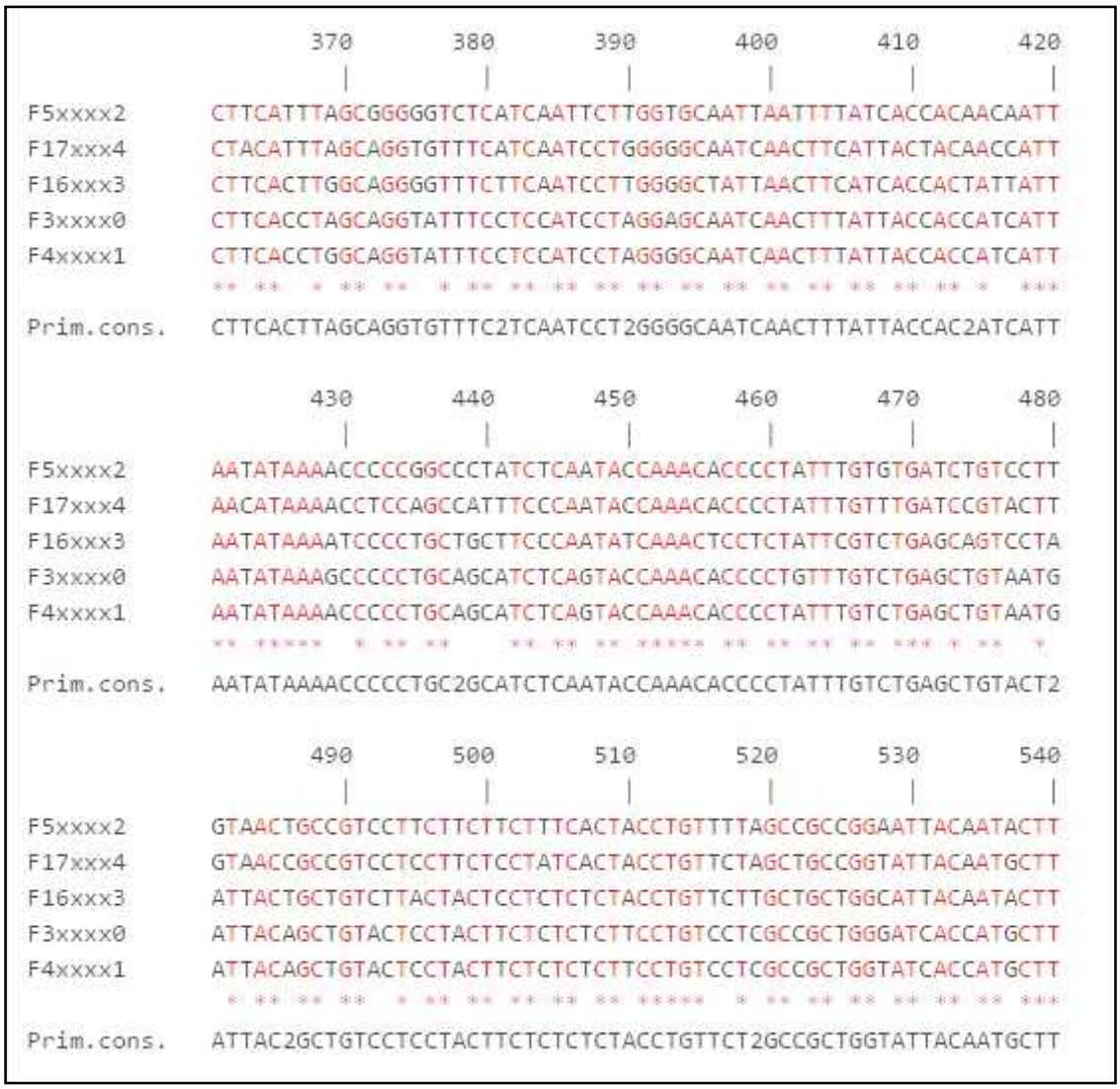
Indicates similarity **Indicates Polymorphic sites**

Fig. 3.1. Multiple sequence alignment of COI gene fragment of 5 aquarium fish species. Some representative polymorphic sites are indicated by nucleotide position.



Indicates similarity **Indicates Polymorphic sites**

Fig. 3.2. Multiple sequence alignment of COI gene fragment of 5 aquarium fish species. Some representative polymorphic sites are indicated by nucleotide position.



Indicates similarity ■ Indicates Polymorphic sites

Fig. 3.3. Multiple sequence alignment of COI gene fragment of 5 aquarium fish species. Some representative polymorphic sites are indicated by nucleotide position.

Comparison of COI gene sequence of selected 5 species of aquarium fishes indicates that they are different species because only 66.42% nucleotide bases (366bp out of 551bp) are similar among them. And 185bp remain different which is about 33.58%.

Alignment data :

Alignment length : 551

Identity (*) : 366 is 66.42 %

Strongly similar (:) : 0 is 0.00 %

Weakly similar (.) : 0 is 0.00 %

Different : 185 is 33.58 %

Sequence 0001 : F5xxxx2 (551 residues).

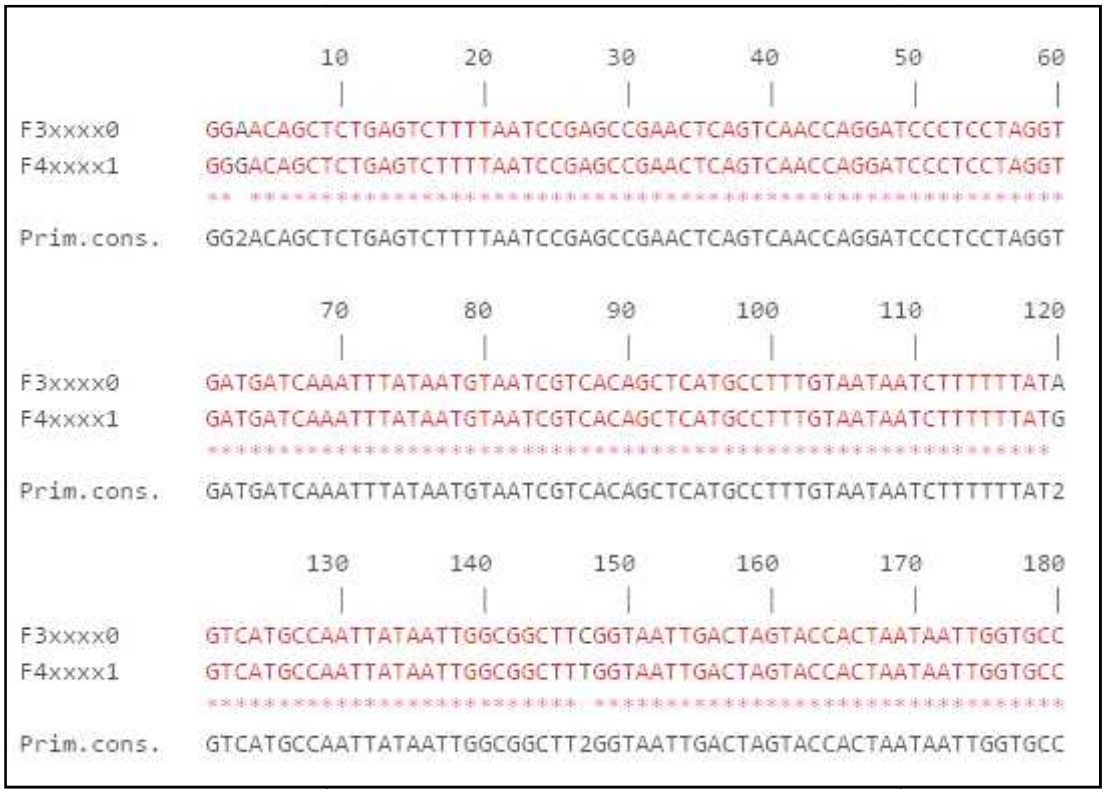
Sequence 0002 : F17xxx4 (551 residues).

Sequence 0003 : F16xxx3 (551 residues).

Sequence 0004 : F3xxxx0 (551 residues).

Sequence 0005 : F4xxxx1 (551 residues).

3.1.2 Alignment between F3 and F4 sample sequence



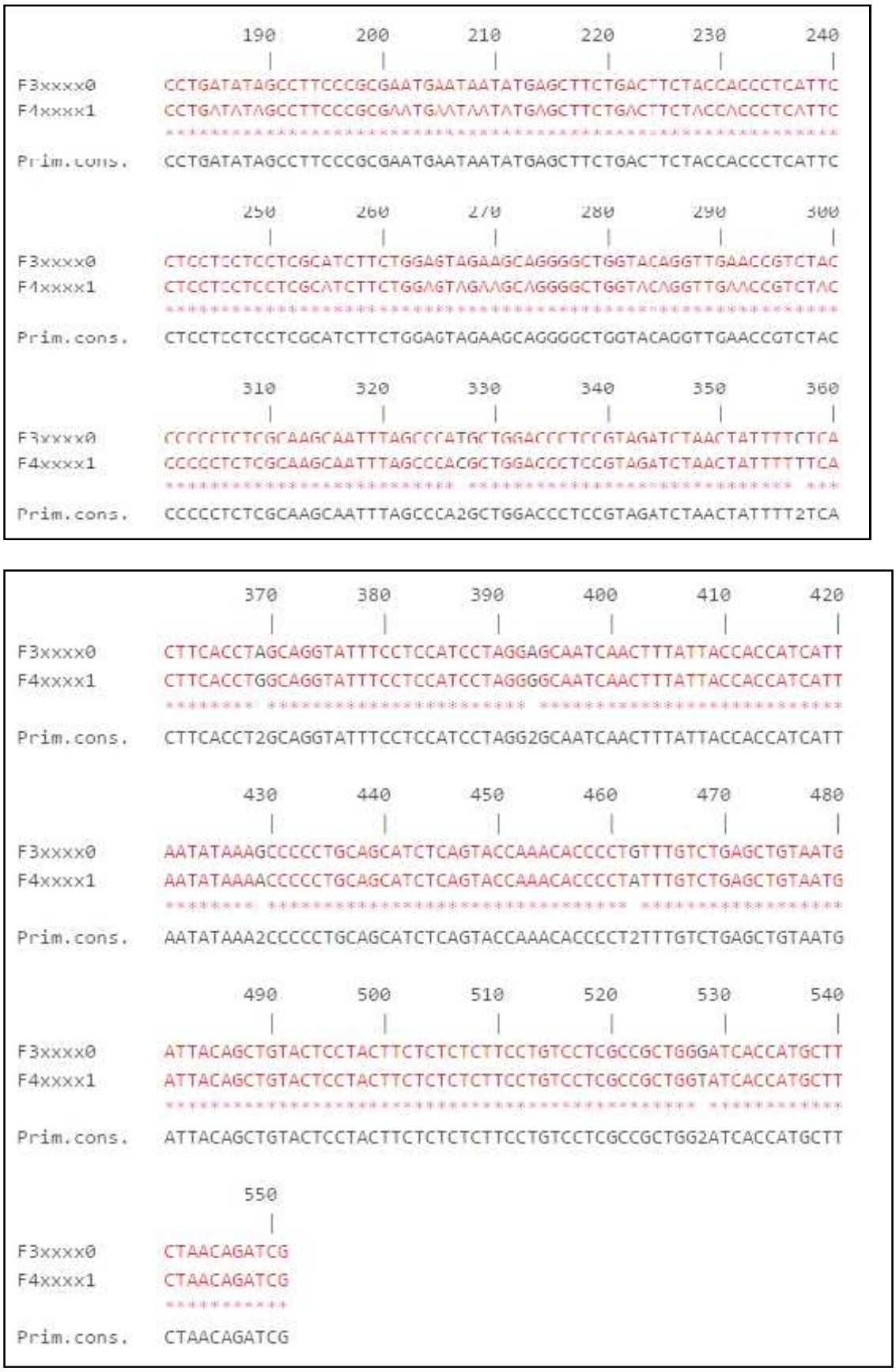


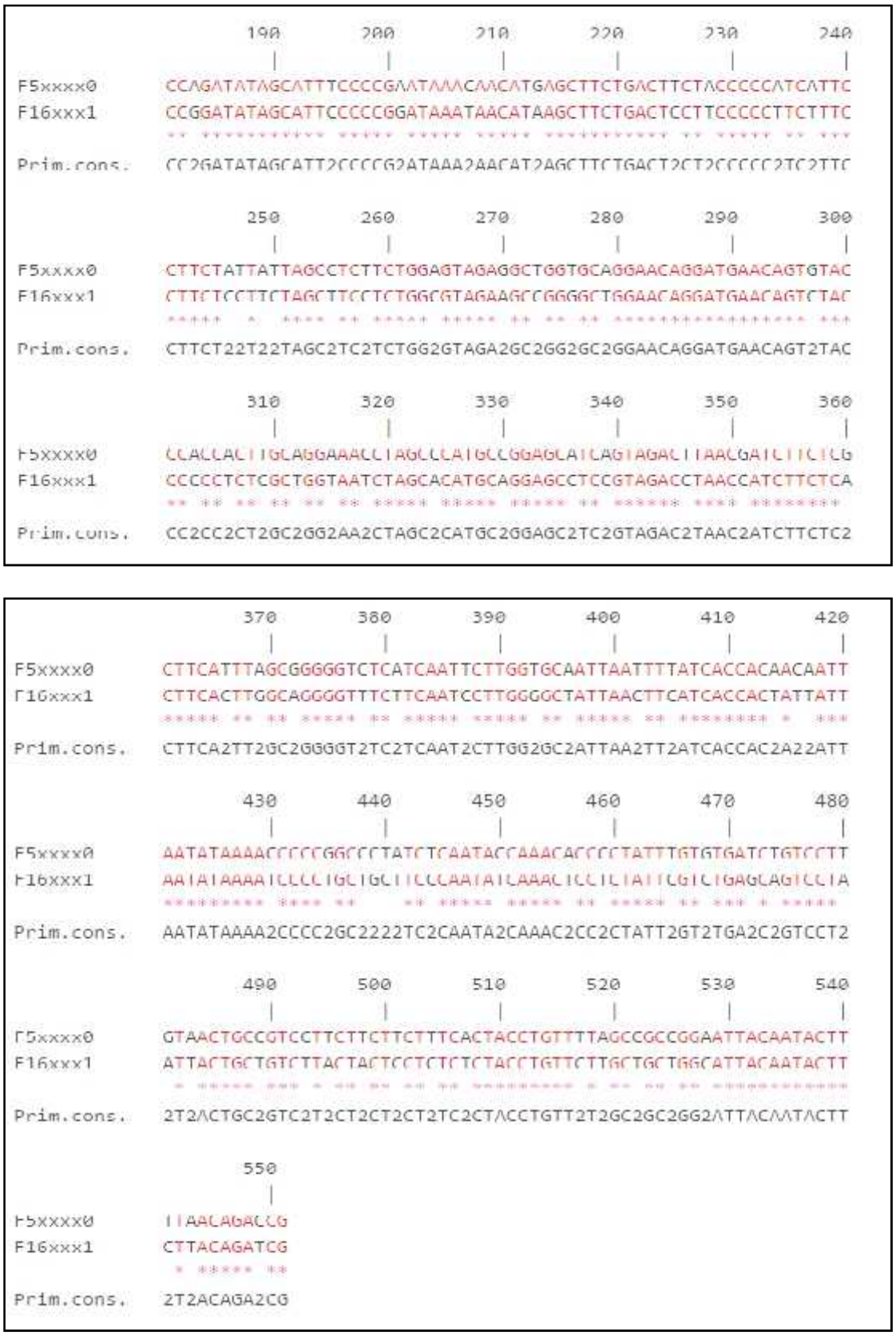
Fig. 3.4.Alignment of COI between F3 and F4 sample of fish spp.

Alignment data :
Alignment length : 551
Identity (*) : 541 is 98.19 %
Strongly similar (:) : 0 is 0.00 %
Weakly similar (.) : 0 is 0.00 %
Different : 10 is 1.81 %
Sequence 0001 : F3xxxx0 (551 residues).
Sequence 0002 : F4xxxx1 (551 residues)

541bp is similar out of 551bp among these fish species which is 98.19% and only 10bp remain different which shows a strong similarity between these two species which indicates they are from same genus and closely related species.

3.1.3 Alignment between F5 and F16 sample sequence

```
                    10        20        30        40        50        60
                     |         |         |         |         |         |
F5xxxx0         GGGACCGCCTTAAGCCTTCTTATCCGGGCTGAATTAAGCCAACCCGGAGCGCTTTTAGGT
F16xxx1         GGGACAGCCTTAAGCCTCCTTATCCGTGCTGAACTAAGCCAACCAGGGGCTCTCCTTGGA
                ***** ********** ******** ****** ********** ** ** **  * **

Prim.cons.      GGGAC2GCCTTAAGCCT2CTTATCCG2GCTGAA2TAAGCCAACC2GG2GC2CT22T2GG2

                    70        80        90       100       110       120
                     |         |         |         |         |         |
F5xxxx0         GATGA-CAAATTTATAATGTTATCGTTACTGCTCACGCTTTCGTAATAATCTTCTTTATA
F16xxx1         GACGACCAGATCTATAATGTGATCGTAACGGCACATGCCTTCGTAATAATTTTCTTTATA
                ** ** ** ** ******** ***** ** ** ** ** ********** *********

Prim.cons.      GA2GA2CA2AT2TATAATGT2ATCGT2AC2GC2CA2GC2TTCGTAATAAT2TTCTTTATA

                   130       140       150       160       170       180
                     |         |         |         |         |         |
F5xxxx0         GTCATACCAATCCTCATTGGAGGCTTTGGAAATTGACTTGTGCCATTAATAATTGGGGCC
F16xxx1         GTTATGCCAGTAATAATTGGAGGGTTTGGAAATTGACTGGTTCCCCTAATGATCGGGGCA
                ** ** *** *  * ******** ********************** **** ** *****

Prim.cons.      GT2AT2CCA2T22T2ATTGGAGG2TTTGGAAATTGACT2GT2CC22TAAT2AT2GGGGC2
```



Indicates similarity **Indicates Polymorphic sites**

Fig. 3.5.Alignment of COI between F5 and F16 sample of fish spp.

Alignment data :

Alignment length : 551

Identity (*) : 436 is 79.13 %

Strongly similar (:) : 0 is 0.00 %

Weakly similar (.) : 0 is 0.00 %

Different : 115 is 20.87 %

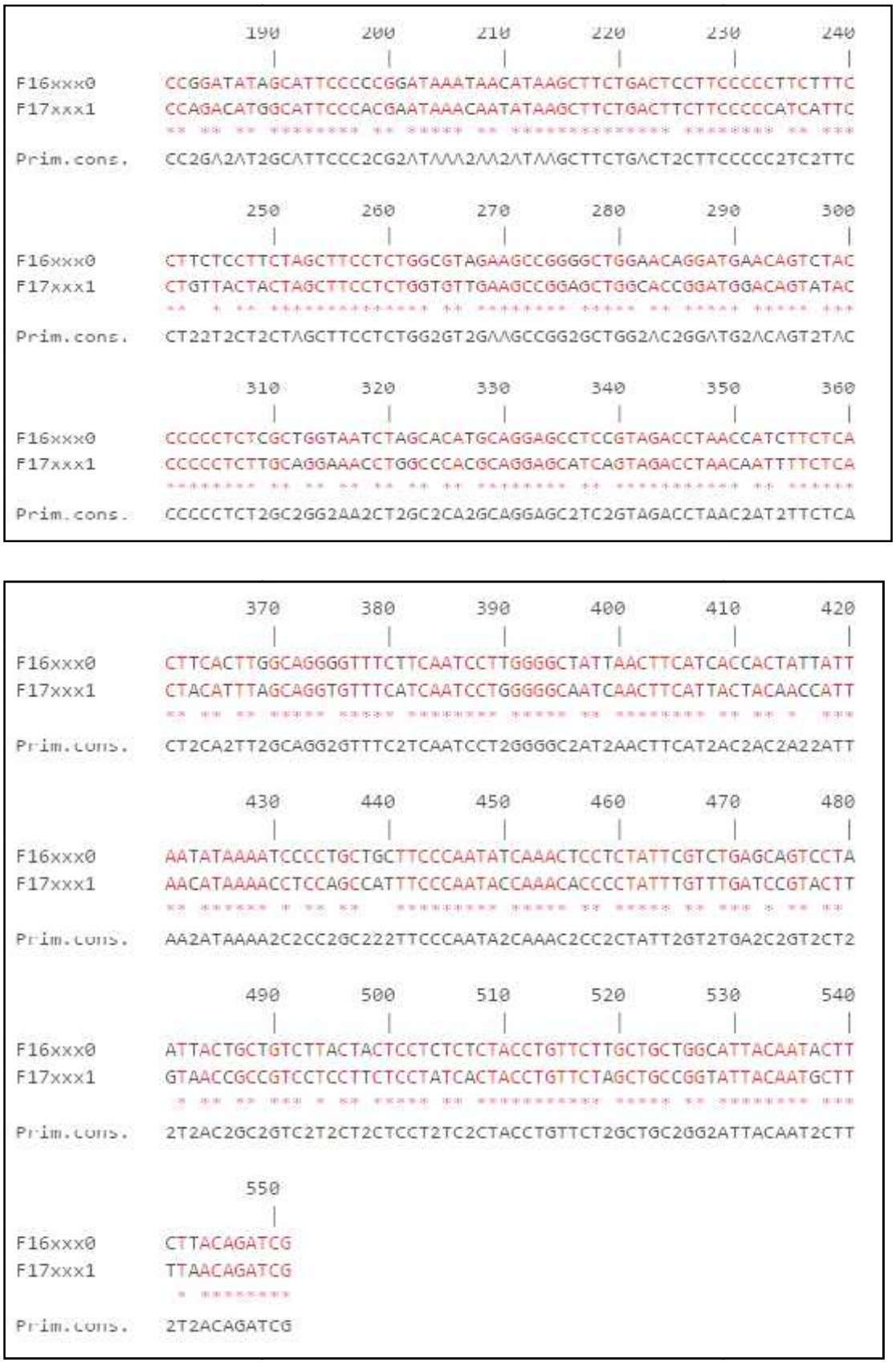
Sequence 0001 : F5xxxx0 (551 residues).

Sequence 0002 : F16xxx1 (551 residues).

There are 436bp is similar between these two species which is 79.13% and 20.87% (115bp) remain different which indicate they are distinct group of species.

3.1.4 Alignment between F16 and F17 sample sequence

```
                    10        20        30        40        50        60
                     |         |         |         |         |         |
F16xxx0      GGGACAGCCTTAAGCCTCCTTATCCGTGCTGAACTAAGCCAACCAGGGGCTCTCCTTGGA
F17xxx1      GGAACCGCTTTAAGCCTCCTCATCCGAGCTGAACTTAGTCAACCCGGATCACTTCTAGGT
             ** ** ** ********** ***** ******** ** ***** **  * ** ** **

Prim.cons.   GG2AC2GC2TTAAGCCTCCT2ATCCG2GCTGAACT2AG2CAACC2GG22C2CT2CT2GG2

                    70        80        90       100       110       120
                     |         |         |         |         |         |
F16xxx0      GACGACCAGATCTATAATGTGATCGTAACGGCACATGCCTTCGTAATAATTTTCTTTATA
F17xxx1      GATGACCAAATTTACAATGTAATTGTTACCGCCCACGCCTTCGTAATAATTTTCTTTATA
             ** ***** ** ** ***** ** ** ** ** ** ************************

Prim.cons.   GA2GACCA2AT2TA2AATGT2AT2GT2AC2GC2CA2GCCTTCGTAATAATTTTCTTTATA

                   130       140       150       160       170       180
                     |         |         |         |         |         |
F16xxx0      GTTATGCCAGTAATAATTGGAGGGTTTGGAAATTGACTGGTTCCCCTAATGATCGGGGCA
F17xxx1      GTAATGCCTATCCTCATTGGAGGATTCGGAAACTGACTTGTACCCCTGATAATCGGAGCC
             ** *****  *  * ******** ** ***** ***** ** ***** ** ***** **

Prim.cons.   GT2ATGCC22T22T2ATTGGAGG2TT2GGAAA2TGACT2GT2CCCCT2AT2ATCGG2GC2
```



Indicates similarity **Indicates Polymorphic sites**

Fig. 3.6.Alignment of COI between F16 and F15 sample of fish spp.

Alignment data :
Alignment length : 551
Identity (*) : 436 is 79.13 %
Strongly similar (:) : 0 is 0.00 %
Weakly similar (.) : 0 is 0.00 %
Different : 115 is 20.87 %
Sequence 0001 : F16xxx0 (551 residues).
Sequence 0002 : F17xxx1 (551 residues).

There are 436bp is similar between these two species which is 79.13% and 20.87% (115bp) remain different which indicate they are distinct group of species.

Table 3.2. Comparison between selected samples of COI sequences

Sl. No	Comparing Samples	Total Base Pair (Number)	Similar Base Pair (Number)	Similar Base Pair (Percentage)
1	F3-F4	551	541	98.19
2	F3-F5	551	426	77.31
3	F3-F16	551	434	78.77
4	F3-F17	551	428	77.68
5	F4-F5	551	426	77.31
6	F4-F16	551	437	79.31
7	F4-F17	551	428	77.68
8	F5-F16	551	436	79.13
9	F5-F17	551	457	82.94
10	F16-F17	551	436	79.13

Here are the similaritiesbetween species of the current study which shows about 77-80 % similarity among species except F3-F4 pair of sample that shows 98.19% similarity and indicates that they are the most closely related pair of species and comes from the same genus *Poecilia*.

Table 3.3. Genetic distance among selected species with outer group

		2	3	4	5	6	7	8	9	10	11	12
1. F3												
2. F4	0.019											
3. F5	0.304	0.305										
4. F16	0.275	0.267	0.274									
5. F17	0.297	0.296	0.296	0.271	0.300							
6. KU568778.1 Carassius auratus voucher KW11-7309	0.297	0.296	0.296	0.271	0.306	0.306						
7. KU9600[illegible].1 Pangasianodon hypophthalmus voucher KW11-7020	0.304	0.305	0.300	0.271	0.306	0.306						
8. KC774658.1 Salvelinus alpinus haplotype ScaH1	0.275	0.267	0.274	0.000	0.271	0.271	0.274					
9. KU568773.1 Poecilia reticulata voucher KW11-7071	0.000	0.019	0.304	0.275	0.297	0.297	0.304	0.275	0.019			
10. KU568724.1 Poecilia sphenops voucher KW11-6973	0.000	0.019	0.304	0.275	0.297	0.297	0.304	0.275	0.019			
11. KF468024.1 Cadra cautella isolate HC5 D-loop	1.685	1.663	1.782	1.785	1.612	1.612	1.782	1.786	1.663	1.663		
12. AY[illegible].1 Danio rerio estrogen receptor type alpha mRNA, partial sequence	2.022	2.020	2.039	2.041	2.000	2.000	2.039	2.051	2.020	2.022	2.000	

Genetic distance indicates the closeness of species being studied. Here, the distance between F3 and F4 sample is 0.019 which indicates that they are closely related. And the distance between F4 and F5 is 0.305 which is maximum among studied species that indicates that they are less close to each other. The distance between F3 and *Poeciliasphenops*is zero which means they are same species.

3.2 Phylogenetic Tree analysis

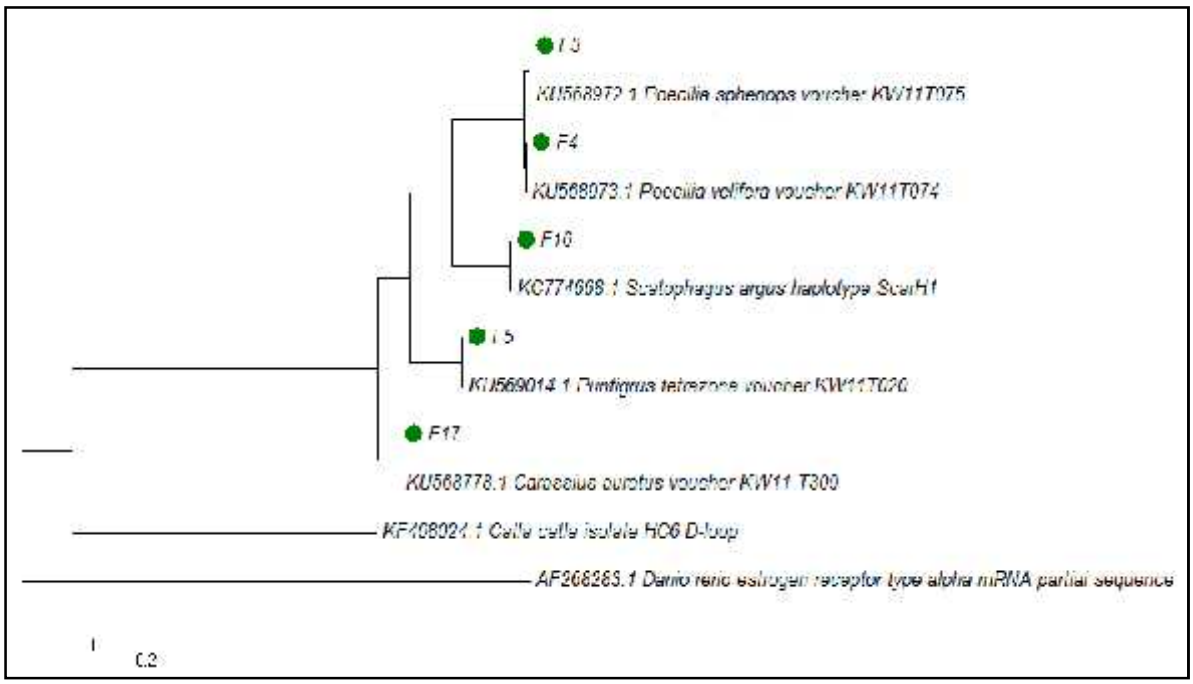


Fig. 3.7.Molecular Phylogenetic analysis by Maximum Likelihood method

The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest log likelihood (-3230.7859) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 12 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 541 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

From the phylogenetic tree it was shown that F3 is closely related with *Poeciliasphenops*, sample F4 is closely related with *Poeciliavelifera.* On the other hand F16 is closely related with *Scatophagusargus.*

The phylogenetic tree also indicates that F5 sample is closely related with *Puntigrustetrazona* while the taxonomic position of F17 shows similarity with *Carassiusauratus.*

Chapter 4

Discussion

DNA Barcoding is effective in identifying species and provided a straightforward identification system when a perfect match existed between the morphology-based taxonomy and genetic divergence.

Overall, this study demonstrated the ability of DNA barcoding to calibrate the current taxonomic resolution and to shed new light on the fish diversity. The application of COI sequence in forensics has already been investigated for reproducibility, heteroplasmy, mixed DNA samples, chemical treatment, environmental conditions and other factors showing consistent in which a great range of reference data exits (Dawnay et al., 2007).

Species identification search only if the species in the reference database has at least three barcoded specimens and identifies the query sequences if it matches the reference sequence within the conspecific distance of less than 2% (Tamura et al., 2007) or not exceeding 3% as suggested by Wong and Hanner (2008). Therefore, correct species labeling, morphological taxonomy and voucher documentation should be prioritized in case that reassessment of spurious data in necessary (Ward et al. 2005). Mislabeling is no unexpected since both of these species are genetically homologous (Na-Nakorn et al., 2002) and morphologically similar.

In this study, we sequenced the COI gene of five aquarium fishes which were collected from Uttara and Katabon market in Dhaka city and established a relationship among them. We found three Orders (Cyprinodontiformes, Cypriniformes and Perciformes) among these five samples which indicate the vast genetic divergence among aquarium fishes. As the aquarium fishes should have some particular characters like coloration, domestication, adaptation in a small environment, it is very important to know their genetic sequence to modify their characters due to customers demand and business issue. This study represents their genetic sequence as well as the relationship among species.

This study also shows the purity of species collected from the markets because the sequences of collected sample were matched with the sequences collected from the Gene Bank of respected species. Genetic similarity among these species may help to produce hybrid group of aquarium fishes with more attractive color as well as more adaptation

power which will increase market value, such as F3 (*Poeciliasphenops*) and F4 (*Poeciliavelifera*) are from same genus Poecilia which are very closely related species.

In the past, mainly morphological characters were used for inferring fish phylogenetic relationship to understand their speciation. In the case of aquarium fishes, it is difficult to differentiate the species because of the similarity in external morphology. Therefore, the reconstructed phylogenetic tree based on morphology were controversial due to the complex evolutionary changes in either morphological or physiological characters. Recent advances in molecular biology have changed this situation. Recently, the genetic analysis of mtDNA has conducted to resolve the controversial taxonomic problem (Suzuki et al. 2005; Erguden et al. 2000) and have proved that the molecular markers can facilitate the discrimination of morphologically similar species.

So, this present study avoided the morphological characters and used genetic sequences of COI as species identification tools to make the result more reliable and authentic. This study also showed the multiple alignment among 5 species as well as alignment among all possible pair (10 pair of 5 samples) of samples to show the genetic similarity and dissimilarity. Phylogenetic tree and genetic distance are also evaluated to indicate the genetically closeness of selected species of this study.

Chapter 5
Conclusion and Recommendation

5.1 Conclusion

The ornamental fish sector is a widespread and global component of international trade, fisheries, aquaculture and development. However, the scope of this sector is vast in our country, but we cannot go ahead because of unconsciousness, lack of knowledge, lack of care of government and no government institutions. This sector should be given priority with extra care because it may a lot of foreign exchange every year by exporting the native ornamental fishes that remain unused in our country.

DNA barcoding is emerging as an invaluable tool to regulatory agencies and fisheries managers for species authentication, food safety, conservation management as well as consumer health and support. The findings indicate that DNA barcoding is a powerful technique, accurately identifying samples regardless of sample source.

This study will influence other researchers to make them involved at this sector and will create interest among them and the government will be aware of the future of this sector and the business of aquarium fishes.

5.2 Recommendation

Probably this is the first attempt to study the taxonomy of aquarium fishes of Bangladesh at molecular level. There are several species of aquarium fish in Dhaka city which could not be possible to take under the present study. So further study is needed to take as many species as possible under investigation.

Moreover, only COI gene was amplified in the present study and species identification was done accordingly. Species identification may be based on other genes like 16S, rRNA, Cytochrome b.

References

Alam, R., Alam, J., Pattadar, S.N., Karim, R., Mahmud, S., 2016.trend of ornamental fish business in Barisal division, Bangladesh. International Journal of Fisheries and Aquatic Studies. 4(3), 263-266.

Chakraborty, M., Ghosh, S.K., 2014.An assessment of the DNA barcodes of Indian freshwater fishes. 537, 20–28.

Dasmahapatra, K.K., Mallet, J., 2006. DNA barcodes: Recent successes and future prospects. Heredity. 97, 254-255.

Dawnay, N., Ogden, R., Mcewin, R., Carvalho, G.R., Thorpe, R.S., 2007. Validation of the barcoding gene COI for use in forensic genetic species identification. Forensic Sci. Int. 173(1), 1-6.

Dugan, K.A., Lawrence, H.S., Hares, D.R., Fisher, C.L., Budowlf, B., 2002. An improved method for post-PCR purification for mtDNA sequence analysis. J. Foren. Sci. 47, 811-818.

Erguden, D., Gurlek, M., Yaglioglu, D., Turan, C., 2010. Genetic identification and taxonomic relationship of Mediterranean Mugilid Species based on Mitochondrial 16S rDNA sequence data. J. Ani. Veter. Adv. 9, 336-341.

Ferraris, J.R., Pinna, M.C.C., 1999. Higher-level names for catfishes (Actinopterygii: Ostariophysi: Siluriformes). Cali. Aca. Sci. 51, 1-17.

Hajibabaei, M., Smith, M.A., Janzen, D.H., 2006. A minimalist barcode can identify a specimen whose DNA is degraded. Mol. Ecol. Notes. 6, 959-964.

Hardman, M., 2005. The phylogenetic relationships among non-diplomystid catfishes as inferred from mitochondrial cytochrome b sequence; the search for the Ictalurid sister – taxon (Otophysi: Siluriformes). Mol. Phylogen. And Evol. 37, 700-720.

Hebert, P.D.N., Cywinska, A., Ball, S.L., Dewaard, J.R., 2003. Biological identification through DNA barcodes. Proceed. R. Soc. Lon. Ser. B. 270, 313-321.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptesfulgerator.* Proc. Natl. Acad. Sci. USA. 101, 14812-14817.

Hogg, I.D., Hebert, P.D.N., 2004. Biological identification of springtails (*Collembolahexapoda*) from the Canadian arctic, using mitochondrial DNA barcodes. Canadian J. Zool. 82, 749-754.

Holmes, B.H., Steinke, D., Ward, R.D., 2009. Identification of shark and ray fins using DNA barcoding. Fish Res. 95, 280-288.

Hubert, N., Hanner, R., Holm, E., Mandrak, N.E., Taylor, E., Burridge, M., Watkinson, D., Dumont, P., Curry, A., Bentzen, P., Zhang, J., April, J., Bernatchez, L., 2008.Identifying Canadian Freshwater Fishes through DNA barcodes.PLos ONE.3(6), e2490.

Islam, M.S., Alam M.A., 2005. Randomly amplified Polymorphic DNA analysis of four different population of the Indian major carp, *Labeorohita* (Hamilton). Journal of Applied Ichthyology. 20, 407-412.

Ivanova, N.V., Zemlak, T.S., Hanner, R.H., Hebert, P.D.N., 2007. Universal primers cocktails for fish DNA barcoding.Mole. Ecol. Not. Doi: 10.1111/j.1471-8286.2007.01748.

Jarman, S.N., Deagle, B.E., Gales, N.J., 2004. Group-specificpolymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples.Journal of Molecular Ecology.13 , 1313–1322.

Kartavtsev, Y.P., Jung, S.O., Lee, Y.M., Byeon, H.K., Lee, J.S., 2007. Complete mitochondrial genome of the bullhead torrent catfish: Genome description and phylogenetic considerations inferred from the Cty b and 16S rRNA genes. Gene. 396, 13-27.

Kim, Y.J., Kim, I.C., Lee, Y.S., Lee, W.O., Lee, J.S., 2003. The use and conservation in molecular phylogeny of fish mitochondrial DNAs in Korean waters. Korean J. Limmol. 36, 221-234.

Meusnier, L., Singer, G., Landry, J., Hickey, D., Hebert, P., Hajibabaei, M., 2008.A universal DNA mini-barcode for biodiversity analysis.BMC genomics.9, 214.

Okazaki, T., Jeon, S.R., Watanabe, M., Kitagawa, T., 1999.Geneticrelationships of Japanese and Korean bagrid catfishes inferred from mitochondrial DNA analysis. Zool. Sci. 16, 363-373.

Rasmussen, R.S., Morrissey, M.T., Hebert, P.D.N., 2009. DNA barcoding of commercially important salmon and trout species from North America.J. Agric. Foof Chem. 57, 8379-8385.

Ratnasingham, S., Hebert, P.D., 2007. BOLD: The barcode of life data system. Mol. Ecol. Not. 7, 355-364.

Smith, P.J., Mcveagh, M.S., Steinke, D., 2008. Application of DNA barcoding for the identification of smoked fish products. J. Fish. Biol. 72, 464-471.

Sriwattanarothai, N., Steinke, D., Ruenwongsa, P., Hanner, R., Oanijpan, B., 2010. Molecular and morphological evidence supports the species status of the Mahachaifighter(*Betta sp.*) and reveals new species of *Betta* from Thailand. J. Fish Biol. 77, 414-424

Steinke, D., Zemlak, T.S., Boutillier, J.A., Hebert, P.D.N., 2009a. DNA barcoding Pacific Canada's fishes. Mar. Biol. 156, 2641-2647

Steinke, D., Zemlak, T.S., Hebert, P.D., 2009. Barcoding memo: DNA-based identification for the ornamental fish trade. PLoS ONE 4(7), e6300.

Steinke, D., Zemlak, T.S., Hebert, P.D.N., 2009b. Barcoding memo: DNA-based identification for the ornamental fish trade. PLoS ONE 4: e6300

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA4) software version 4.0. Mol. Biol. Evol. 24, 1596-1599

Tavares, E.S., Baker, A.J., 2008. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. BMC.Evol. Biol. 8:81

Valdez-moreno, M., Ivanova, N.V., Elias-Gutierrez, M., Contreras-balderas, S., Hebert, P.D.N., 2009. Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes. J. Fish Biol. 74, 377-402.

Ward, R.D., Hanner, R., Hebert, P.D.N., 2009. The campaign to DNA barcode all fishes, FISH-BOL.J.Fish Biol. 74, 329-356

Ward, R.D., Holmes, B.H., White, W.T., Last, P.R., 2008. DNA barcoding Australian chondrichthyans results and potential uses in conservation. Marine Freshwater Res. 59, 57-71

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D., 2005. DNA barcoding Australia's fish species. Phil. Trans. R. Soc. B. Biol. Sci. 360, 1847-1857.

Wong, E.H.K., Hanner, R.H., 2008. DNA barcoding detects market substitution in North American seafood. Food Research Int. 41, 828-837

Zhang, J.B., Hanner, R., 2011. DNA barcoding is a useful tool for the identification of marine fishes from Japan. Biochemical Systematics and Ecology. 39, 31–42

Appendices

Appendix 1.

5 M Stock Solution of NaCl (100ml)

29.22 g of NaCl was dissolved in 75ml of distilled water and adjusted the volume to 100ml. The mixture was heated for 15 sec and stirred and store at 4^oC.

70% Ethanol (100ml)

30 ml water is added in 70 ml ethanol to prepare 100ml 70% ethanol.

TE (Tris-HCl EDTA) Buffer, pH 8.0 (100ml)

1 ml of 1 M Tris-HCl was added with 0.2 ml of 0.5 M EDTA and adjusted to 100 ml by sterile water and stored at 4^oC.

0.5 M Stock Solution of EDTA, pH 8.0 (100ml)

18.61 g of EDTA was added to 75 ml of distilled water and 2 g of NaOH was used to adjust the final pH 8.0. Then volume was adjusted to 100 ml by sterile water and stored at 4^oC.

1 M Stock Solution of Tris-HCl, pH 8.0 (100ml)

12.14 g of Trizma base was dissolved in 75 ml of distilled water and pH was adjusted at 8.0 by adding 5 ml of concentrated HCl. Then volume was adjusted to 100 ml by sterile water and stored at 4^oC.

3 M Sodium Acetate, pH 5.2 (100ml)

40.824 g of sodium acetate with 70 ml of ddH_2O and adjusted the final volume to 100 ml with ddH_2O and pH was maintained to 5.2

50x TAE Buffer, pH 8.3 (l liter)

242 Trizma base was dissolved into 900 ml of sterile water and 57 ml glacial acetic acid was added to the solution. Then 100 ml 0.5 M EDTA (pH 8.0) was added and pH was adjusted by HCl at 8.3. Final volume was adjusted to 1000 ml.

6x Loading Dye

6x loading Dye is required to load samples in gel electrophoresis for visualization.

Ethidium Bromide Solution

10 mg of ethidium bromide was mixed with 1 ml of water to make 1 ml solution and kept in a dark place. As it is carcinogenic, it's better to purchase directly from companies as 10 mg/ml solution.

Appendix 2.

Random primers used in PCR and sequencing

Primer Code	Sequence (5´-3´)	Reference
Fish F1	TCAACCAACCACAAAGACATTGGCAC	Ward et al., 2005
Fish F2	TCGACTAATCATAAAGATATCGGCAC	Ward et al., 2005
Fish R1	TAGACTTCTGGGTGGCCAAAGAATCA	Ward et al., 2005
Fish R2	ACTTCAGGGTGACCGAAGAATCAGAA	Ward et al., 2005

Components of PCR reaction mixture for 5 reactions

Sl. No	Reagents	Amount per sample (µl)	Total
1	Sterile de-ionized distilled water	18.7	93.5
2	*Tag* buffer A 10x	2.5	12.5
3	Primer	1.0	5.0
4	dNTPs 10mM	0.25	1.25
5	*Tag* DNA Polymerase	0.05	0.25
6	Template DNA	2.5	12.5
	Total	**25**	**125**

Table 6: Comparison between selected samples of COI sequences

Sl. No	Comparing Samples	Total Base Pair (Number)	Similar Base Pair (Number)	Similar Base Pair (Percentage)
1	F3-F4	551	541	98.19
2	F3-F5	551	426	77.31

3	F3-F16	551	434	78.77
4	F3-F17	551	428	77.68
5	F4-F5	551	426	77.31
6	F4-F16	551	437	79.31
7	F4-F17	551	428	77.68
8	F5-F16	551	436	79.13
9	F5-F17	551	457	82.94
10	F16-F17	551	436	79.13

Appendix 3.

All sequences of collected Sample and Real Species samples with Outer groups

>F3

GGAACAGCTCTGAGTCTTTTAATCCGAGCCGAACTCAGTCAACCAGGATCCCTCCTAGGTGATGATCAAATTTATAATG
TAATCGTCACAGCTCATGCCTTTGTAATAATCTTTTTTATAGTCATGCCAATTATAATTGGCGGCTTCGGTAATTGACTA
GTACCACTAATAATTGGTGCCCCTGATATAGCCTTCCCGCGAATGAATAATATGAGCTTCTGACTTCTACCACCCTCAT
TCCTCCTCCTCCTCGCATCTTCTGGAGTAGAAGCAGGGGCTGGTACAGGTTGAACCGTCTACCCCCCTCTCGCAAGCAA
TTTAGCCCATGCTGGACCCTCCGTAGATCTAACTATTTTCTCACTTCACCTAGCAGGTATTTCCTCCATCCTAGGAGCAA
TCAACTTTATTACCACCATCATTAATATAAAGCCCCCTGCAGCATCTCAGTACCAAACACCCCTGTTTGTCTGAGCTGT
AATGATTACAGCTGTACTCCTACTTCTCTCTCTTCCTGTCCTCGCCGCTGGGATCACCATGCTTCTAACAGATCG

>F4

GGGACAGCTCTGAGTCTTTTAATCCGAGCCGAACTCAGTCAACCAGGATCCCTCCTAGGTGATGATCAAATTTATAATG
TAATCGTCACAGCTCATGCCTTTGTAATAATCTTTTTTATGGTCATGCCAATTATAATTGGCGGCTTTGGTAATTGACTA
GTACCACTAATAATTGGTGCCCCTGATATAGCCTTCCCGCGAATGAATAATATGAGCTTCTGACTTCTACCACCCTCAT
TCCTCCTCCTCCTCGCATCTTCTGGAGTAGAAGCAGGGGCTGGTACAGGTTGAACCGTCTACCCCCCTCTCGCAAGCAA
TTTAGCCCACGCTGGACCCTCCGTAGATCTAACTATTTTTTCACTTCACCTGGCAGGTATTTCCTCCATCCTAGGGGCAA
TCAACTTTATTACCACCATCATTAATATAAAACCCCCTGCAGCATCTCAGTACCAAACACCCCTATTTGTCTGAGCTGT
AATGATTACAGCTGTACTCCTACTTCTCTCTCTTCCTGTCCTCGCCGCTGGTATCACCATGCTTCTAACAGATCG

>F5

GGGACCGCCTTAAGCCTTCTTATCCGGGCTGAATTAAGCCAACCCGGAGCGCTTTTAGGTGATGATCAAATTTATAATG
TTATCGTTACTGCTCACGCTTTCGTAATAATCTTCTTTATAGTCATACCAATCCTCATTGGAGGCTTTGGAAATTGACTT
GTGCCATTAATAATTGGGGCCCCAGATATAGCATTTCCCCGAATAAACAACATGAGCTTCTGACTTCTACCCCCATCAT
TCCTTCTATTATTAGCCTCTTCTGGAGTAGAGGCTGGTGCAGGAACAGGATGAACAGTGTACCCACCACTTGCAGGAA
ACCTAGCCCATGCCGGAGCATCAGTAGACTTAACGATCTTCTCGCTTCATTTAGCGGGGGTCTCATCAATTCTTGGTGC
AATTAATTTTATCACCACAACAATTAATATAAAACCCCCGGCCCTATCTCAATACCAAACACCCCTATTTGTGTGATCT
GTCCTTGTAACTGCCGTCCTTCTTCTTCTTTCACTACCTGTTTTAGCCGCCGGAATTACAATACTTTTAACAGACCG

>F16

GGGACAGCCTTAAGCCTCCTTATCCGTGCTGAACTAAGCCAACCAGGGGCTCTCCTTGGAGACGACCAGATCTATAAT
GTGATCGTAACGGCACATGCCTTCGTAATAATTTTCTTTATAGTTATGCCAGTAATAATTGGAGGGTTTGGAAATTGAC

TGGTTCCCCTAATGATCGGGGCACCGGATATAGCATTCCCCCGGATAAATAACATAAGCTTCTGACTCCTTCCCCCTTC
TTTCCTTCTCCTTCTAGCTTCCTCTGGCGTAGAAGCCGGGGCTGGAACAGGATGAACAGTCTACCCCCCTCTCGCTGGT
AATCTAGCACATGCAGGAGCCTCCGTAGACCTAACCATCTTCTCACTTCACTTGGCAGGGGTTTCTTCAATCCTTGGGG
CTATTAACTTCATCACCACTATTATTAATATAAAATCCCCTGCTGCTTCCCAATATCAAACTCCTCTATTCGTCTGAGCA
GTCCTAATTACTGCTGTCTTACTACTCCTCTCTCTACCTGTTCTTGCTGCTGGCATTACAATACTTCTTACAGATCG

> F17

GGAACCGCTTTAAGCCTCCTCATCCGAGCTGAACTTAGTCAACCCGGATCACTTCTAGGTGATGACCAAATTTACAATG
TAATTGTTACCGCCCACGCCTTCGTAATAATTTTCTTTATAGTAATGCCTATCCTCATTGGAGGATTCGGAAACTGACTT
GTACCCCTGATAATCGGAGCCCCAGACATGGCATTCCCACGAATAAACAATATAAGCTTCTGACTTCTTCCCCCATCAT
TCCTGTTACTACTAGCTTCCTCTGGTGTTGAAGCCGGAGCTGGCACCGGATGGACAGTATACCCCCCTCTTGCAGGAAA
CCTGGCCCACGCAGGAGCATCAGTAGACCTAACAATTTTCTCACTACATTTAGCAGGTGTTTCATCAATCCTGGGGGCA
ATCAACTTCATTACTACAACCATTAACATAAAACCTCCAGCCATTTCCCAATACCAAACACCCCTATTTGTTTGATCCG
TACTTGTAACCGCCGTCCTCCTTCTCCTATCACTACCTGTTCTAGCTGCCGGTATTACAATGCTTTTAACAGATCG

>KU568778.1 *Carassiusauratus* voucher KW11-T309

GGAACCGCTTTAAGCCTCCTCATCCGAGCTGAACTTAGTCAACCCGGATCACTTCTAGGTGATGACCAAATTTACAATG
TAATTGTTACCGCCCACGCCTTCGTAATAATTTTCTTTATAGTAATGCCTATCCTCATTGGAGGATTCGGAAACTGACTT
GTACCCCTGATAATCGGAGCCCCAGACATGGCATTCCCACGAATAAACAATATAAGCTTCTGACTTCTTCCCCCATCAT
TCCTGTTACTACTAGCTTCCTCTGGTGTTGAAGCCGGAGCTGGCACCGGATGGACAGTATACCCCCCTCTTGCAGGAAA
CCTGGCCCACGCAGGAGCATCAGTAGACCTAACAATTTTCTCACTACATTTAGCAGGTGTTTCATCAATCCTGGGGGCA
ATCAACTTCATTACTACAACCATTAACATAAAACCTCCAGCCATTTCCCAATACCAAACACCCCTATTTGTTTGATCCG
TACTTGTAACCGCCGTCCTCCTTCTCCTATCACTACCTGTTCTAGCTGCCGGTATTACAATGCTTTTAACAGATCG

>KU569014.1 *Puntigrustetrazona* voucher KW11T020

GGGACCGCCTTAAGCCTTCTTATCCGGGCTGAATTAAGCCAACCCGGAGCGCTTTTAGGTGATGATCAAATTTATAATG
TTATCGTTACTGCTCACGCTTTCGTAATAATCTTCTTTATAGTCATACCAATCCTCATTGGAGGCTTTGGAAATTGACTT
GTGCCATTAATAATTGGGGCCCCAGATATAGCATTTCCCCGAATAAACAACATGAGCTTCTGACTTCTACCCCCATCAT
TCCTTCTATTATTAGCCTCTTCTGGAGTAGAGGCTGGTGCAGGAACAGGATGAACAGTGTACCCACCACTTGCAGGAA
ACCTAGCCCATGCCGGAGCATCAGTAGACTTAACGATCTTCTCGCTTCATTTAGCGGGGGTCTCATCAATTCTTGGTGC
AATTAATTTTATCACCACAACAATTAATATAAAACCCCCGGCCCTATCTCAATACCAAACACCCCTATTTGTGTGATCT
GTCCTTGTAACTGCCGTCCTTCTTCTTCTTTCACTACCTGTTTTAGCCGCCGGAATTACAATACTTTTAACAGACCG

>KC774668.1 *Scatophagusargus* haplotype ScarH1

GGGACAGCCTTAAGCCTCCTTATCCGTGCTGAACTAAGCCAACCAGGGGCTCTCCTTGGAGACGACCAGATCTATAAT
GTGATCGTAACGGCACATGCCTTCGTAATAATTTTCTTTATAGTTATGCCAGTAATAATTGGAGGGTTTGGAAATTGAC
TGGTTCCCCTAATGATCGGGGCACCGGATATAGCATTCCCCCGGATAAATAACATAAGCTTCTGACTCCTTCCCCCTTC
TTTCCTTCTCCTTCTAGCTTCCTCTGGCGTAGAAGCCGGGGCTGGAACAGGATGAACAGTCTACCCCCCTCTCGCTGGT
AATCTAGCACATGCAGGAGCCTCCGTAGACCTAACCATCTTCTCACTTCACTTGGCAGGGGTTTCTTCAATCCTTGGGG
CTATTAACTTCATCACCACTATTATTAATATAAAATCCCCTGCTGCTTCCCAATATCAAACTCCTCTATTCGTCTGAGCA
GTCCTAATTACTGCTGTCTTACTACTCCTCTCTCTACCTGTTCTTGCTGCTGGCATTACAATACTTCTTACAGATCG

>KU568973.1 *Poeciliavelifera* voucher KW11T074

GGGACAGCTCTGAGTCTTTTAATCCGAGCCGAACTCAGTCAACCAGGATCCCTCCTAGGTGATGATCAAATTTATAATG
TAATCGTCACAGCTCATGCCTTTGTAATAATCTTTTTTATGGTCATGCCAATTATAATTGGCGGCTTTGGTAATTGACTA
GTACCACTAATAATTGGTGCCCCTGATATAGCCTTCCCGCGAATGAATAATATGAGCTTCTGACTTCTACCACCCTCAT
TCCTCCTCCTCCTCGCATCTTCTGGAGTAGAAGCAGGGGCTGGTACAGGTTGAACCGTCTACCCCCCTCTCGCAAGCAA
TTTAGCCCACGCTGGACCCTCCGTAGATCTAACTATTTTTTCACTTCACCTGGCAGGTATTTCCTCCATCCTAGGGGCAA

TCAACTTTATTACCACCATCATTAATATAAAACCCCCTGCAGCATCTCAGTACCAAACACCCCTATTTGTCTGAGCTGT
AATGATTACAGCTGTACTCCTACTTCTCTCTCTTCCTGTCCTCGCCGCTGGTATCACCATGCTTCTAACAGATCG

>**KU568972.1 *Poeciliasphenops* voucher KW11T075**

GGAACAGCTCTGAGTCTTTTAATCCGAGCCGAACTCAGTCAACCAGGATCCCTCCTAGGTGATGATCAAATTTATAATG
TAATCGTCACAGCTCATGCCTTTGTAATAATCTTTTTTATAGTCATGCCAATTATAATTGGCGGCTTCGGTAATTGACTA
GTACCACTAATAATTGGTGCCCCTGATATAGCCTTCCCGCGAATGAATAATATGAGCTTCTGACTTCTACCACCCTCAT
TCCTCCTCCTCCTCGCATCTTCTGGAGTAGAAGCAGGGGCTGGTACAGGTTGAACCGTCTACCCCCCTCTCGCAAGCAA
TTTAGCCCATGCTGGACCCTCCGTAGATCTAACTATTTTCTCACTTCACCTAGCAGGTATTTCCTCCATCCTAGGAGCAA
TCAACTTTATTACCACCATCATTAATATAAAGCCCCCTGCAGCATCTCAGTACCAAACACCCCTGTTTGTCTGAGCTGT
AATGATTACAGCTGTACTCCTACTTCTCTCTCTTCCTGTCCTCGCCGCTGGGATCACCATGCTTCTAACAGATCG

>KF468024.1 *Catlacatla* isolate HC6 D-loop

GGCATAAAATTAAGACCTAAAATTTAATTGTTTTAACATGAGAAACAGATTTTTCCCCAGAACATCGACCTCAAATTTT
TCCTTGAAGTAATCAACTAGGATCTTATTAAACCATTTTAATGTAGTAAGAGACCACCAACCTATTACATAAAGGAATA
TCATGCATGATAGAATCAGGGACAACAATCGTGGGGGTCGCACAATATGAACTATTACTGGCATCTGGTTCCTATTTC
AGGTCCATAAACTGTAAAATTCCCTCCCCAGATAATTATACTGGCATCTGATTAATGGTGTAGTACATACGCCTCGTTA
CCCACCATGCCGGGCATTCTTTTATATGCATAGGGTATCTCTTTCTGGTTTCCTTTCACTTGGCATCTCAGAGTGCAGGC
ACAAATGTTAATCAAGGTTGAACATTTTTCTTGTATGTGACTATATATATTAATTATTGAAAGACATAAGTTAAGCATT
GCATTCTCCTAGTTCAGGTGCATAACATATCTACTCCTTATTCAACTATACCTGCTATAATGCCCCCTTTGGTTTTTGCG
CGACAAACCC

>AF268283.1 *Daniorerio* estrogen receptor type alpha mRNA, partial sequence

GTGAGGGATGCAAAGCTTTCTTCAAGAGAAGCATTTAAGGTCACAATGACTATGTTTGTCCAGCGACCAACCAGTGCA
CTATTGACAGAAACCGTCGAAAGAGCTGCCAAGCATGCAGACTGCGCAAGTGTTATGAAGTAGGCATGATGAAAGGA
GGTATTCGTAAAGATCGCGGCGGGCGTTCTGTCAGGCGTGAGAGAAGAAGAAGCAGTAATGAAGATCGAGACAAGAG
CACCAGTGATCAGTGCAGCCGTGCTGGCGTGAGGACGACTGGCCCACAGGACAAGAGGAAGAAGCGCAGTGGTGGGG
TGGTCAGCACTTTATGCATGTCGCCTGACCAGGTGCTGCTGCTGCTGCTGGGGCTGAGCCACCCGCTGTCTGCTCACGA
CAGAAACACAGTCGGCCCTACACCGAGATCACCATGATGTCCCTGCTCACCAACATGGCTGACAAAGAACTCGTCCAC
ATGATCGCCTGGGCCAAGAAAGTGCCAGGATTCCAGGATCTGTCTCTGCATGACCAGGTTCAGTTGTTGGAGAGCTCT
TGGCTGGAGGTGTTGATGATTGGCCTCATATGGAGGTCCA